JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 7 (2019), Issue 4 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Viewpoints

Original Papers

Reviews

Corrigenda and Addendas

Viewpoint

Mobile Health Interventions: Exploring the Use of Common Relationship Factors

Emily R Grekin[1], PhD; Jessica R Beatty[2], PhD; Steven J Ondersma[2,3], PhD

[1]Department of Psychology, Wayne State University, Detroit, MI, United States

[2]Merrill-Palmer Skillman Institute, Wayne State University, Detroit, MI, United States

[3]Department of Psychiatry and Behavioral Neuroscience, Wayne State University, Detroit, MI, United States

Corresponding Author:
Emily R Grekin, PhD
Department of Psychology
Wayne State University
5057 Woodward Ave
Detroit, MI, 48202
United States
Phone: 1 313 577 2366
Fax: 1 313 577 7636
Email: grekine@wayne.edu

Abstract

The use of mobile health (mHealth) interventions has risen dramatically over the past two decades. It is important to consider mHealth intervention research within the broader therapy outcome literature. Among other key findings, this broader literature suggests that common relationship factors such as empathy, positive regard, and genuineness may play a critical role in therapy effectiveness. These findings raise intriguing questions for mobile interventions. For example, can mobile interventions incorporate aspects of common factors to augment their efficacy? Will the absence of relationship-based common factors make mobile interventions less effective? This viewpoint paper addresses these questions as well as related issues such as how to operationalize relationship qualities in the context of a mobile intervention and whether common relationship factors apply to computers or computerized narrators. The paper concludes by outlining a future research agenda guided by theory and empirical studies.

(JMIR Mhealth Uhealth 2019;7(4):e11245) doi:10.2196/11245

KEYWORDS

mobile health; mHealth; smartphone; empathy; mobile applications; therapeutic alliance

Mobile Health Interventions and the Use of Common Relationship Factors

Mobile health (mHealth) interventions have become increasingly prevalent in the scientific literature [1,2]. Currently, there are 7811 publications using the terms mobile intervention or e-intervention indexed on PsycINFO, including multiple meta-analyses within subareas of the field [3-5]. There are also more than 100,000 iPhone and Android apps specifically designed to target health-related behaviors [2]. Researchers working in this area often cite the potential of mHealth interventions to reach a large audience at low cost, regardless of barriers related to language, geographic location, or time. These factors make mHealth interventions uniquely applicable to nontreatment-seeking individuals, who may refuse extended, in-person treatment but accept a minimal, opportunistic intervention.

However, it is important to consider this research within the broader, person-delivered therapy outcome literature. Among other key findings, this literature suggests that common relationship factors such as empathy, alliance, positive regard, and genuineness play a critical role in therapy effectiveness and account for unique variance in treatment efficacy above and beyond specific therapeutic techniques [6]. Specifically, ratings of therapist-client relationship factors have been shown to predict therapy outcome across rater type (eg, client, therapist, observer), observed relationship characteristics (eg, empathy, genuineness, alliance, cohesion), patient characteristics (eg, age, gender, race, diagnosis), stage of therapy (eg, early, middle, late), and theoretical orientation [6,7].

These findings raise intriguing questions for mHealth interventions [1-4]. For example, will the absence of relationship-based common factors make mobile interventions less effective? Can mobile interventions incorporate aspects of

common factors to augment their efficacy? Do qualities such as empathy and positive regard matter in the context of a mobile intervention? This review will address these questions as well as related issues such as how to operationalize relationship qualities in the context of a mobile intervention. In particular, we will (1) review research suggesting that humans react to computers in social ways and respond positively to software using human-like relational agents, (2) describe studies directly testing the hypothesis that common factors increase mHealth intervention efficacy, and (3) outline a future research agenda guided by both comprehensive theory and empirical studies.

Responding to Computers in Social Ways

Literature from the field of human-computer interaction suggests that people automatically and unconsciously react to computers in social ways [8,9]. Much of the early work in this area was conducted by Nass and colleagues who, through a wide-ranging series of studies, found that human-computer interactions, in some ways, mirrored human-human interactions. For example, Nass and colleagues [10] assigned participants to work with a computer on an interactive tutoring task in which the computer presented and tested participants on a series of facts. After the task, participants were asked to evaluate the computer's performance. Participants completed the evaluation either (1) on the same computer that administered the task, (2) on a different computer in another room, or (3) on a paper-and-pencil questionnaire. Results showed that participants gave more positive evaluations when the computer asked about its own performance versus when participants completed the evaluation on a separate computer or on a paper-and-pencil questionnaire. Thus, participants appeared to apply social norms of politeness to the computer (despite denying that they did so in postexperimental interviews).

In a similar study, Moon [11] examined how norms of self-disclosure were applied to computers. Participants were asked a series of interview questions by a computer (eg, "What have you done in your life that you feel most guilty about?" and "What do you dislike about your physical appearance?"). In the no-reciprocity condition, the computer simply asked each question without presenting additional information. In the reciprocity condition, the computer preceded each question with information about itself (eg, "There are times when this computer crashes for reasons that are not apparent to its user. It usually does this at the most inopportune time, causing great inconvenience for the user. What have you done in your life that you feel most guilty about?"). Results showed that participants in the reciprocity condition provided more and longer disclosures than participants in the no-reciprocity condition. They also reported being more attracted to the computer.

Other studies suggest that humans respond positively to flattery from a computer. For example, Fogg and Nass [12] instructed participants to play a guessing game with a computer (similar to 20 questions). As part of the game, participants were asked to suggest guesses that might be useful to the computer in the future. They then received feedback about their suggestions from the computer (eg, "Your question makes an interesting and useful distinction. Great job!"). Participants in the sincere praise condition were told that feedback from the computer was directly related to their suggestions, participants in the flattery condition were told that computerized feedback was preprogrammed and unrelated to their suggestions, and participants in the generic feedback condition were given a neutral message ("Begin next round"). In reality, all feedback was preprogrammed and identical. Results showed that participants in the flattery condition reported more positive affect and gave higher ratings to the computer than participants in the generic feedback condition, even though they were told that computer feedback was unrelated to their responses. Moreover, responses from participants in the flattery and sincere praise conditions did not differ.

Other data indicate that humans automatically apply social categories (eg, gender, ethnicity, ingroup, and outgroup) to computers. For example, Nass and colleagues [13] asked Korean male participants to read a series of hypothetical scenarios in which they had to choose between a risky versus a safe course of action. Participants were then instructed to ask a computerized agent what course of action he would recommend and why. Afterward, participants were asked to rate the computerized agent and the quality of his arguments. In some cases, the computerized agent was Asian (ie, the same ethnicity as the participant), whereas in other cases, he was white (ie, a different ethnicity than the participant). Results showed that participants rated same-ethnicity agents as being more attractive, trustworthy, persuasive, and intelligent than different ethnicity agents. Participants also felt that the same-ethnicity agent's decision was closer to their own.

In a similar study, Nass and colleagues [14] examined whether humans could feel in-group bias toward a computer. In this study, participants were assigned to either a shared identity condition or a nonshared identity condition. In the shared identity condition, participants and their computer were referred to as the blue team. Participants were asked to wear a blue armband and to work with a computer that had a blue border around its monitor. Participants in this condition were reminded that they were dependent upon the computer. In the nonshared identity condition, participants wore a blue armband and were referred to as the blue person, whereas the computer had a green border and was referred to as the green computer. Participants in this condition were asked to focus on individual responsibility. After being assigned to an identity condition, participants worked with the computer on a desert survival problem. They then ranked their interaction with the computer along a variety of indices. Results showed that participants in the shared identity condition rated the computer as being more friendly, intelligent, and similar to themselves than did participants in the nonshared identity condition. They were also more likely to cooperate with the computer and conform to its suggestions.

Finally, data suggest that humans can feel ostracized by computers. For example, Zadro and colleagues [15] instructed participants to control the actions of an avatar who was playing a game of catch with 2 other avatars on a computer screen. Participants were told that, when they received the ball, they should click on 1 of the other 2 avatars to indicate where the

ball should go next. In the low ostracism condition, participants received the ball multiple times throughout the game. In the high ostracism condition, participants only received the ball once or twice at the beginning of the game. Data revealed that, compared with low ostracism participants, high ostracism participants experienced a host of negative feelings, including anger and lowered feelings of belonging, self-esteem, control, and meaningfulness. Moreover, these feelings were produced even when participants (1) knew that they were playing against a computer rather than another human and (2) were explicitly told that the other characters' actions were determined by a prewritten script.

Electronic Coaches and Relational Agents

Notably, although the studies described above demonstrate social responses to computers, the effect sizes reported in this literature have been small, suggesting that social reactions to computers, while consistently detectable, are smaller in magnitude than social reactions to actual humans. In addition, the extent to which these basic social reactions translate into therapeutic or long-term relationships is unclear (ie, we know that people apply social categories and in-group bias to computers, but can they also form therapeutic relationships with them?).

Relevant to this issue are findings from the electronic intervention (e-intervention) literature suggesting that computerized interventions may be more effective when coupled with human support. In particular, recent studies have shown that human electronic coaches (e-coaches; ie, individuals such as nurses, therapists, or research assistants who provide support and assistance throughout an intervention) can increase intervention effectiveness and adherence [16,17]. For example, Tate and colleagues [18] randomly assigned a group of overweight adults to 1 of 3 e-interventions. In the no counseling condition, participants attended a single group session in which they were given specific weight loss strategies. They were then taught to use an interactive website that provided weekly weight loss tips, prompts to report weight, recipes, and the potential to connect online with others trying to lose weight. In the automated feedback condition, participants used the website described above in addition to receiving automated, weekly, tailored feedback from a preprogrammed computer. In the human counseling group, participants used the interactive website and received regular, personalized emails from a trained, human counselor. Results showed that, at 3-month follow-up, the automated feedback and the human counseling groups had greater weight loss than the no-counseling group, and there was no difference in weight loss between the 2 counseling conditions. In contrast, at 6-month follow-up, the human counseling group had greater weight loss than both the automated feedback and the no-counseling conditions.

In a similar study, Gabriele and colleagues [19] assigned overweight adults to 1 of 3 weight loss intervention conditions: (1) a minimal support condition in which participants engaged with a Web-based weight loss program and were sent weekly lessons and feedback graphs; (2) a directive e-coach condition in which participants engaged with a Web-based online weight loss program and also received weekly emails from a directive coach who prescribed specific goals and plans; or (3) a nondirective e-coach condition in which participants engaged with a Web-based weight loss program and received weekly emails from a nondirective coach who allowed them to decide what goals to set and what strategies to follow. Results showed that females in the directive e-coach condition lost more weight, had greater increases in physical activity, and had greater changes in waist circumference than females in the nondirective or minimal support conditions.

Building upon these and other studies, Mohr and colleagues outlined the supportive accountability model, which describes how human support can enhance electronic health interventions [20]. According to this model, adherence to e-interventions is enhanced by coaches who are trustworthy, collaborative, able to provide patients with clear benefits and expertise, and explicit about expectations and accountability processes. Mohr and colleagues also hypothesize that the relationship between human support and e-intervention adherence is moderated by patient motivation and communication medium.

Notably, the supportive accountability model focuses exclusively on human support and does not address the degree to which e-interventions can be enhanced by support from nonhuman coaches, such as relational agents, or by purposeful inclusion of lifelike characteristics. Relational agents are "computational artifacts, such as animated, screen-based characters or social robots, that are designed to establish a sense of rapport, trust, and even therapeutic alliance with patients," by whatever means are appropriate [21]. A growing body of literature suggests that computerized relational agents are satisfying to work with, can provide support, and can help with a variety of diverse tasks [22,23]. For example, Bickmore and colleagues [24] developed an animated relational agent designed to help individuals find cancer-related clinical trials using the National Cancer Institute (NCI) database. Participants were 89 individuals with a cancer diagnosis and varying levels of health literacy. All participants were asked to search the NCI database for 1 clinical trial that met their needs and 1 clinical trial that met the needs of a hypothetical patient. Half of the participants were assigned to use the standard database search engine; the other half interacted with a relational agent who facilitated the search by asking questions, helping to narrow down search criteria, and explaining characteristics of identified clinical trials. The relational agent was an animated female who used synthetic speech and nonverbal behaviors (such as hand gestures, facial displays, gaze, and use of props). Results revealed that participants in the relational agent group were more satisfied and pleased and less frustrated with the search task than participants in the control group. In addition, participants with low health literacy in the relational agent group were significantly better at identifying clinical trials for a hypothetical patient than participants with low health literacy in the control group.

In a related study, Gardiner and colleagues [25] assigned 61 women to (1) a condition in which they interacted with a computerized relational agent who provided information on stress management, mindfulness, healthy eating, and physical activity or (2) a control condition in which they met for 60 min

with a technician who reviewed education sheets about stress management, mindfulness, healthy eating, and physical activity and were given a CD containing meditation and mindfulness exercises. Results showed that, compared with the control group, women who interacted with the computerized relational agent increased their fruit consumption and decreased their use of alcohol to cope with stress. They also made positive comments about their interactions with the relational agent, such as, "She relates to my stress" and "She helped me relax."

Chattaraman and colleagues [26] created a relational agent to help older adults navigate through a Web-based retail store. A total of 60 participants (mean age: 69 years) were assigned to purchase a set of clothing on a mock website. In addition, half of the participants were assisted by a relational agent (Gina) who interacted with them throughout the task. Results showed that the presence of a relational agent increased perceived social support, trust, and intentions to use the Web-based store. In addition, the effects of the agent on trust were mediated by perceived social support, and the effects of the agent on intentions to use the store were mediated by trust.

The effectiveness of relational agents has also been demonstrated by studies of social robots (ie, robots that interact with humans and exhibit social behaviors; [27,28]. Similar to computerized relational agents, social robots have demonstrated acceptability and usefulness [27,29]. They also tend to elicit social behaviors and anthropomorphization. For example, de Graaf and colleagues [27] conducted a qualitative study examining older adults' acceptance of an in-home social robot (Harvey, a 12-inch-tall rabbit with moving ears and blinking lights). The robot was designed to initiate at least three conversations per day with participants and alternated between 3 states: sleeping, alert, and engaged (ie, listening and talking). The robot was installed in each participant's home for three 10-day periods. Afterward, participants were interviewed about their experience, and their responses were coded for content. Participants tended to attribute human-like qualities to the robot (from de Graaf and colleagues [27]):

> *The rabbit itself was kind of sweet. If it was furry, I would stroke it.*
>
> *Because Harvey was Harvey, I talked to him as a male, and males do tend to get on your nerves from time to time...*

Participants also followed social rules, such as politeness, when interacting with the robot:

> *So whether it's a machine that talks to you or somebody who's going to stay, you have got to have some communication with them just out of sheer politeness and friendliness...*

All but one participant noted Harvey's potential for companionship:

> *I got used to the idea that it would greet me in the morning.*

Finally, studies from the intervention literature have shown that individuals are able to establish working alliances with relational agents and software programs. For example, Kiluk and colleagues [30] assessed working alliance in a sample of cocaine-dependent patients who were assigned to either treatment as usual (TAU: methadone maintenance plus regular sessions with a counselor) or TAU plus 7 sessions of a computerized cognitive behavioral intervention. Several times throughout the study, participants completed the Working Alliance Inventory (WAI), a measure designed to assess alliance with the therapist along 3 dimensions: task (therapist responsiveness to client needs), bond (mutual liking between therapist and client), and goal (extent to which therapy goals are agreed upon and attainable). In addition, participants who completed the computerized intervention were given an adapted version of the WAI (the WAI-Tech) designed to assess alliance with the computer program. Results showed that mean scores on the task and goal scales of the WAI-Tech were similar to (and sometimes higher than) mean scores on the task and goal scales on the WAI. In contrast, bond scores on WAI-Tech, while consistently above the neutral midpoint, were lower than bond scores on the WAI.

Strengthening the Effects of Relational Agents

As the literature on computerized relational agents has expanded, researchers have begun to focus on factors that strengthen their effects. In particular, some studies suggest that greater agent anthropomorphism and behavioral realism lead to high-quality social interaction. For example, Gong [31] asked undergraduates to work through a series of social dilemma scenarios with a computerized agent. The agents represented 4 levels of anthropomorphism, ranging from humanoid robot characters to actual human faces. After completing the task, participants rated the agent on competency, trust, homophily, and social judgment. Results showed that, as the agent became more anthropomorphic, ratings in all domains became more positive. Similarly, Lee and Nass [32] asked undergraduates to participate in a conformity experiment with 1 to 4 fictional participants whose opinions were represented with a text box, a stick figure with a speech bubble, or a fully animated figure with facial expressions, body movements, and a speech bubble. Although the text box condition unexpectedly elicited the most conformity, the animated character was rated as the most trustworthy, competent, and socially attractive.

Notably, some studies in this area have yielded null results [33,34]. Others have failed to control for agent attractiveness or have confounded anthropomorphism with modality; that is, rather than varying anthropomorphism *within* modality (ie, comparing faces or agents with varying levels of humanness), these studies compare text on the computer screen (the low anthropomorphic stimulus) with faces or agents (the high anthropomorphic stimulus [31]). It should also be noted that the effects of anthropomorphism may be moderated by individual difference variables such as need for social connection [35] or participant/agent ethnicity match [36]. Finally, some data suggest that when agents are too realistic (ie, when they have a near perfect human likeness), they can elicit negative reactions and cause discomfort (ie, the uncanny valley phenomenon [37,38]).

Another body of literature compares relational agents (animated figures whose speech and actions reflect computer algorithms) with avatars (animated figures whose speech and actions are controlled by a real person in real time). It is often assumed that avatars have more social influence than relational agents because they are controlled by real people (ie, the agency assumption). However, research testing this assumption has yielded mixed results, with some studies finding that avatars elicit more social behavior than agents [39-41] and others finding no difference between the 2 types of digital representations [42]. Recent meta-analytic data suggest that avatars do, in fact, have more influence over behavior than agents but that the effect of agency (ie, avatar vs agent) is moderated by several variables including task type (cooperative/competitive/neutral), level of immersion, subjective versus objective dependent variables, and whether the representation is actually controlled by a human [43].

Implications for Mobile Interventions

The findings reviewed above suggest that (1) humans automatically relate to computers/agents in social ways, (2) certain relational characteristics (anthropomorphism, agency, etc) may strengthen the social response to computers/agents, and (3) relational agents with human-like qualities can facilitate behavior change. These findings have important implications for mHealth/e-interventions and their therapeutic mechanisms. Specifically, they suggest that mobile interventions—particularly those with anthropomorphic agents or avatars—may activate social cognitions and expectations that may, in turn, affect intervention response. However, the degree to which these social reactions can be harnessed to improve mHealth or e-intervention efficacy is only beginning to be examined. In fact, only a small handful of studies have directly tested whether relational factors (eg, empathy, positive regard, humor, and genuineness) can increase the acceptability and/or efficacy of these interventions.

In 1 of the few studies directly examining this question, Bickmore and Picard [44] assigned 101 healthy adults to work with 1 of 3 exercise promotion programs: a relational program, a nonrelational program, or a control program. In all 3 programs, participants recorded their daily activity for 30 days. Participants in the relational program interacted with a computerized, relational agent who used social dialogue, empathic feedback, humor, and a variety of other relational behaviors. Participants in the nonrelational program interacted with a computerized, nonrelational agent who provided information about exercise in the absence of relational behaviors (she did not provide empathy, humor, dialogue, etc). Participants in the control condition did not interact with a computerized agent. Results showed that participants liked, trusted, and respected the relational agent more than the nonrelational agent. In addition, participants expressed more desire to continue working with the relational versus the nonrelational agent.

Similarly, Berry and colleagues [45] presented a healthy eating message to undergraduates using either text, a voice, a human actor, or a relational agent named GRETA. GRETA either (1) expressed emotion consistent with the message she was presenting (eg, smiling while talking about health benefits), (2) expressed emotion inconsistent with the message she was presenting (eg, looking concerned while talking about health benefits), or (3) did not express emotion (neutral condition). Participants rated evidence provided by the neutral version of GRETA as more convincing, more trustworthy, and of higher quality than the evidence provided by the emotional versions of GRETA. However, participants had the greatest recall for the healthy eating message that was presented by the consistent emotion version of GRETA, suggesting that emotionally consistent facial cues may aid in encoding and recall.

Other studies have focused specifically on empathy in relational agents. For example, Brave and colleagues [46] instructed 96 participants to play a game of blackjack with a computerized relational agent. At the end of each blackjack round, the agent made 1 comment about his/her performance and 1 comment about the participant's performance. A total of 2 primary variables were manipulated: the presence versus absence of empathic emotion and the presence versus absence of self-oriented emotion (the authors also manipulated the gender of the agent). When empathic emotion was present, the agent made empathic comments about the participant's performance after each round ("You won! That's wonderful!"). When self-oriented emotion was present, the agent made emotional comments about his/her own performance after each round ("The dealer beat me, I'm disappointed"). When empathic and/or self-oriented emotion were absent, the agent's comments were factual and did not contain emotion words (eg, "I won" or "The dealer beat you"). At the end of the game, participants rated the agent on a variety of dimensions. Similar to Bickmore and Picard [44] and Berry and colleagues [45], empathic agents were rated as more caring, likeable, trustworthy, and supportive than nonempathic agents. In contrast, self-oriented emotion had little effect on perceptions of the agent.

In another direct test of agent empathy, Ellis and colleagues [47] examined whether expressions of empathy from an animated relational agent improved the efficacy of a brief, motivational intervention for alcohol use. A total of 100 heavy-drinking undergraduates were randomly assigned to either a high or a low empathy version of the intervention. In the high empathy intervention, a relational agent used standard motivational interviewing techniques and made a series of personalized empathic reflections (eg, "You really like the way alcohol helps you to relax."). In the low empathy intervention, the agent used motivational interviewing strategies but did not make any empathic reflections. Intentions to reduce drinking were assessed both before and after the intervention, and a change score was calculated. Similar to previously reviewed studies, results showed that participants who worked with high empathy relational agents felt more supported and less criticized than participants who worked with low empathy relational agents. In addition, participants who worked with high empathy agents reported greater increases in intentions to reduce drinking over the course of the study than those who worked with low empathy agents. Thus, the presence of an empathic relational agent improved likeability and led to greater increases in intention to change alcohol use.

In sum, early studies imply that mHealth and e-interventions can be effective, not just by providing information and/or skills training but also by establishing a therapeutic *relationship* with

a client based on qualities such as respect and empathy. Although more research is clearly needed, existing data are promising and suggest the potential for improving computerized intervention outcomes.

Mobile Interventions as a Platform for Testing Relationship Factors

The studies reviewed above also highlight the methodological advantages of using mobile interventions as a platform for testing relational factors. In particular, computerized interventions facilitate testing of relationship factors using random assignment. To date, virtually all in-person common factors research has been correlational because of the practical and ethical barriers associated with manipulating common factors during in-person therapy (eg, therapists cannot reliably alter their levels of empathy and positive regard for clients in different study conditions). As a result, it is unclear whether client traits elicit reactions from therapists (eg, motivated clients may elicit more positive, empathic responses than unmotivated clients) or whether therapist behavior elicits reactions from clients (eg, empathic therapists may elicit more motivation from clients). In addition, it is unclear whether common factors are the cause or the result of a successful therapy outcome (eg, does empathy cause less substance use or does less substance use elicit more empathy?). Software, on the other hand, can be easily programmed to include (or not include) common factors such as reflections, statements of affirmation, humor, etc. Moreover, clients can be randomly assigned to different versions of a computer program (eg, a version with an empathic vs a nonempathic relational agent), with the knowledge that the computer will not be affected by the clients' behavior in undesired ways. Finally, mHealth interventions can reach large numbers of participants by reducing barriers associated with cost, transportation, and treatment-related stigma. These increased sample sizes allow researchers to examine moderators (ie, for whom and in what contexts do relational factors increase intervention effectiveness). Thus, by using random assignment, reaching large numbers of participants, and systematically manipulating the presence of relationship factors in mobile interventions, it is possible to examine associations between computerized interventions and common factors in a novel and effective way.

Future Research

Despite the widespread use of both mHealth interventions and relational agents, few studies have systematically examined ways in which relational factors affect the acceptability and efficacy of mobile interventions. There are also some notable gaps in the literature. For example, the ways in which relational factors have been operationalized, delivered, and analyzed has varied widely, making it difficult to generalize across studies. In addition, few studies have examined whether individual difference factors (eg, impulsiveness and loneliness), target behaviors (eg, substance use and weight loss), or contextual factors (eg, social support and impairment) moderate the relationship between relational factors and outcomes. Finally, studies have not examined whether intervention length (ie, single vs multiple session) moderates the effects of relational factors.

As the field moves forward, there are a multitude of potential investigative avenues to explore. However, the following research designs may be particularly fruitful in providing information and helping to make mHealth interventions more powerful:

1. Studies directly comparing mHealth interventions with and without relational factors using random assignment to condition. Few studies have attempted these direct comparisons. Those that have done so have examined widely varying target behaviors, intervention techniques, and relational factors, making it difficult to generalize across studies or draw firm conclusions.
2. Studies examining how to best operationalize relational factors in the context of mobile interventions. For example, what is the best way for a relational agent to express empathy? Are certain types of humor ineffective when expressed by a computerized agent? Can individuals perceive computerized agents as genuine? Although many studies have used relational agents, few have systematically examined ways to operationalize the common factors expressed by these agents.
3. Studies examining interactions between relational factors. For example, it is possible that expressions of empathy work best when they are delivered by highly realistic agents who use gestures and dynamic facial expressions. Similarly, it is possible that a participant/agent therapeutic alliance can only be established when the role of humans in developing the agent is emphasized.
4. Studies examining the degree to which computerized relational factors interact with individual difference variables. It is possible that specific traits or characteristics (eg, extraversion or loneliness) affect how individuals respond to computerized expressions of common factors. For example, individuals who are high in agreeableness may value empathy or humor within an mHealth intervention more than individuals who are low on these traits.
5. Studies comparing the effects of relational factors on single session versus more extended mHealth interventions. It is possible that certain relational factors (eg, empathy and genuineness) are more effective when delivered in extended interventions, whereas others (eg, humor) may be effective in brief and extended interventions.
6. Studies comparing interactions with real people with interactions with relational agents. There have been few direct comparisons between the use of an e-coach and the use of a computerized relational agent, and the degree to which relational agents can produce equivalent results as human e-coaches is unclear.

The above are but a few examples of how research using mobile interventions could evaluate the potential role of common factors in facilitating key outcomes such as engagement, retention, and efficacy. Although extensive research is needed in this field, it appears that incorporation of relational factors is a promising strategy that may make a meaningful difference in mHealth intervention efficacy.

Acknowledgments

This study was supported by National Institute on Alcohol Abuse and Alcoholism Grant R21 AA02366001A1 to ERG.

Conflicts of Interest

SJO reports part ownership of a company marketing authorable e-intervention software. The remaining authors declare no conflicts of interest.

References

1. Shingleton RM, Palfai TP. Technology-delivered adaptations of motivational interviewing for health-related behaviors: a systematic review of the current research. Patient Educ Couns 2016 Jan;99(1):17-35 [FREE Full text] [doi: 10.1016/j.pec.2015.08.005] [Medline: 26298219]
2. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
3. Boumparis N, Karyotaki E, Schaub MP, Cuijpers P, Riper H. Internet interventions for adult illicit substance users: a meta-analysis. Addiction 2017 Sep;112(9):1521-1532 [FREE Full text] [doi: 10.1111/add.13819] [Medline: 28295758]
4. 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;39(6):785-804. [doi: 10.1177/0145445515595198] [Medline: 26187164]
5. Spijkerman MP, Pots WT, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114 [FREE Full text] [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
6. Norcross JC, Wampold BE. Evidence-based therapy relationships: research conclusions and clinical practices. Psychotherapy (Chic) 2011 Mar;48(1):98-102. [doi: 10.1037/a0022161] [Medline: 21401280]
7. Norcross JC, Lambert MJ. Psychotherapy relationships that work II. Psychotherapy (Chic) 2011 Mar;48(1):4-8. [doi: 10.1037/a0022180] [Medline: 21401268]
8. Nass C, Moon Y. Machines and mindlessness: social responses to computers. J Soc Issues 2000 Jan;56(1):81-103 [FREE Full text] [doi: 10.1111/0022-4537.00153]
9. Reeves B, Nass C. The Media Equation: How People Treat Computers, Television, And New Media Like Real People And Places (csli Lecture Notes S). Stanford, CA: Center For The Study Of Language And Inf; 2019.
10. Nass C, Moon Y, Carney P. Are people polite to computers? Responses to computer-based interviewing systems. J Appl Soc Pyschol 1999 May;29(5):1093-1109. [doi: 10.1111/j.1559-1816.1999.tb00142.x]
11. Moon Y. Intimate exchanges: using computers to elicit self‐disclosure from consumers. J Consum Res 2000 Mar;26(4):323-339. [doi: 10.1086/209566]
12. Fogg B, Nass C. Silicon sycophants: the effects of computers that flatter. Int J Hum Comput Stud 1997 May;46(5):551-561. [doi: 10.1006/ijhc.1996.0104]
13. Justine C. Embodied Conversational Agents (mit Press). Cambridge, MA: The Mit Press; 2019.
14. Nass C, Fogg BJ, Moon Y. Can computers be teammates? Int J Hum Comput Stud 1996 Dec;45(6):669-678 [FREE Full text] [doi: 10.1006/ijhc.1996.0073]
15. Zadro L, Williams KD, Richardson R. How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. J Exp Soc Psychol 2004 Jul;40(4):560-567. [doi: 10.1016/j.jesp.2003.11.006]
16. Tate DF, Wing RR, Winett RA. Using internet technology to deliver a behavioral weight loss program. J Am Med Assoc 2001 Mar 7;285(9):1172-1177. [Medline: 11231746]
17. Tate DF, Jackvony EH, Wing RR. Effects of internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. J Am Med Assoc 2003 Apr 9;289(14):1833-1836. [doi: 10.1001/jama.289.14.1833] [Medline: 12684363]
18. Tate DF, Jackvony EH, Wing RR. A randomized trial comparing human e-mail counseling, computer-automated tailored counseling, and no counseling in an internet weight loss program. Arch Intern Med 2006 Aug;166(15):1620-1625. [doi: 10.1001/archinte.166.15.1620] [Medline: 16908795]
19. Gabriele JM, Carpenter BD, Tate DF, Fisher EB. Directive and nondirective e-coach support for weight loss in overweight adults. Ann Behav Med 2011 Apr;41(2):252-263 [FREE Full text] [doi: 10.1007/s12160-010-9240-2] [Medline: 21108032]
20. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
21. Bickmore T, Gruber A. Relational agents in clinical psychiatry. Harv Rev Psychiatry 2010;18(2):119-130. [doi: 10.3109/10673221003707538] [Medline: 20235777]

22. Sillice MA, Morokoff PJ, Ferszt G, Bickmore T, Bock BC, Lantini R, et al. Using relational agents to promote exercise and sun protection: assessment of participants' experiences with two interventions. J Med Internet Res 2018 Feb 07;20(2):e48 [FREE Full text] [doi: 10.2196/jmir.7640] [Medline: 29415873]
23. Scholten MR, Kelders SM, Van Gemert-Pijnen JE. Self-guided web-based interventions: scoping review on user needs and the potential of embodied conversational agents to address them. J Med Internet Res 2017 Dec 16;19(11):e383 [FREE Full text] [doi: 10.2196/jmir.7351] [Medline: 29146567]
24. Bickmore TW, Utami D, Matsuyama R, Paasche-Orlow MK. Improving access to online health information with conversational agents: a randomized controlled experiment. J Med Internet Res 2016;18(1):e1 [FREE Full text] [doi: 10.2196/jmir.5239] [Medline: 26728964]
25. Gardiner PM, McCue KD, Negash LM, Cheng T, White LF, Yinusa-Nyahkoon L, et al. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: A feasibility randomized control trial. Patient Educ Couns 2017 Sep;100(9):1720-1729. [doi: 10.1016/j.pec.2017.04.015] [Medline: 28495391]
26. Chattaraman V, Kwon W, Gilbert JE. Virtual agents in retail web sites: benefits of simulated social interaction for older users. Comput Human Behav 2012 Nov;28(6):2055-2066 [FREE Full text] [doi: 10.1016/j.chb.2012.06.009]
27. de Graaf MM, Allouch SB, Klamer T. Sharing a life with Harvey: Exploring the acceptance of and relationship-building with a social robot. Comput Human Behav 2015 Feb;43:1-14. [doi: 10.1016/j.chb.2014.10.030]
28. Young JE, Hawkins R, Sharlin E, Igarashi T. Toward acceptable domestic robots: applying insights from social psychology. Int J of Soc Robotics 2008 Nov 18;1(1):95-108. [doi: 10.1007/s12369-008-0006-y]
29. Kanda T, Sata R, Saiwaki N, Ishiguro H. A two month trial in an elementary school for long term interaction. IEEE Trans Robot 2007;23:962-971 [FREE Full text]
30. Kiluk BD, Serafini K, Frankforter T, Nich C, Carroll KM. Only connect: the working alliance in computer-based cognitive behavioral therapy. Behav Res Ther 2014 Dec;63:139-146. [doi: 10.1016/j.brat.2014.10.003]
31. Gong L. How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. Comput Human Behav 2008 Jul;24(4):1494-1509. [doi: 10.1016/j.chb.2007.05.007]
32. Lee EJ, Nass C. Experimental tests of normative group influence and representation effects in computer-mediated communication: when interacting via computers differs from interacting with computers. Human Comm Res 2002 Jul;28(3):349-381. [doi: 10.1111/j.1468-2958.2002.tb00812.x]
33. Nowak KL. The influence of anthropomorphism and agency on social judgment in virtual environments. J Comput Mediat Commun 2004;9:00-00. [doi: 10.1111/j.1083-6101.2004.tb00284.x]
34. Nowak KL, Biocca F. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. Presence (Camb) 2003 Oct;12(5):481-494. [doi: 10.1162/105474603322761289]
35. Powers KE, Worsham AL, Freeman JB, Wheatley T, Heatherton TF. Social connection modulates perceptions of animacy. Psychol Sci 2014 Oct;25(10):1943-1948 [FREE Full text] [doi: 10.1177/0956797614547706] [Medline: 25193944]
36. Pratt JA, Hauser K, Ugray Z, Patterson O. Looking at human–computer interface design: effects of ethnicity in computer agents. Interact Comput 2007 Jul;19(4):512-523. [doi: 10.1016/j.intcom.2007.02.003]
37. Mori M, MacDorman K, Kageki N. The uncanny valley [from the field]. IEEE Robot Automat Mag 2012 Jun;19(2):98-100. [doi: 10.1109/MRA.2012.2192811]
38. Tinwell A, Grimshaw M, Nabi D, Williams A. Facial expression of emotion and perception of the Uncanny Valley in virtual characters. Comput Human Behav 2011 Mar;27(2):741-749 [FREE Full text] [doi: 10.1016/j.chb.2010.10.018]
39. Guadagno R, Blascovich J, Bailenson J, McCall C. Virtual humans and persuasion: the effects of agency and behavioral realism. Media Psychol 2007;10:22. [doi: 10.1080/15213260701300865]
40. Hoyt CL, Blascovich J, Swinth KR. Social inhibition in immersive virtual environments. Presence (Camb) 2003 Apr;12(2):183-195. [doi: 10.1162/105474603321640932]
41. Morkes J, Kernal HK, Nass C. Effects of humor in task-oriented human-computer interaction and computermediated communication: a direct test of SRCT theory. Hum Comput Interact 2009 Nov 12;14(4):395-435. [doi: 10.1207/S15327051HCI1404_2]
42. Appel J, von der Pütten A, Krämer NC, Gratch J. Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. Adv Hum Comput Interact 2012;2012:1-10. [doi: 10.1155/2012/324694]
43. Fox J, Ahn S, Janssen J, Yeykelis L, Segovia K, Bailenson J. Avatars versus agents: a meta-analysis quantifying the effect of agency on social influence. Hum Comput Interact 2015:A [FREE Full text] [doi: 10.1080/07370024.2014.921494]
44. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. ACM Trans Comput Hum Interact 2005 Jun 01;12(2):293-327. [doi: 10.1145/1067860.1067867]
45. Berry DC, Butler LT, de Rosis F. Evaluating a realistic agent in an advice-giving task. Int J Hum Comput Stud 2005 Sep;63(3):304-327. [doi: 10.1016/j.ijhcs.2005.03.006]
46. Brave S, Nass C, Hutchinson K. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int J Hum Comput Stud 2005 Feb;62(2):161-178 [FREE Full text] [doi: 10.1016/j.ijhcs.2004.11.002]

47. Ellis JD, Grekin ER, Beatty JR, McGoron L, LaLiberte BV, Pop DE, et al. Effects of narrator empathy in a computer delivered brief intervention for alcohol use. Contemp Clin Trials 2017 Oct;61:29-32. [doi: 10.1016/j.cct.2017.07.014] [Medline: 28732758]

Abbreviations

e-coach: electronic coach
e-intervention: electronic intervention
mHealth: mobile health
NCI: National Cancer Institute
TAU: treatment as usual
WAI: Working Alliance Inventory

Edited by G Eysenbach; submitted 06.06.18; peer-reviewed by M Schaub, A Kassavou, T Bickmore; comments to author 06.10.18; revised version received 06.11.18; accepted 12.11.18; published 15.04.19.

Please cite as:
Grekin ER, Beatty JR, Ondersma SJ
Mobile Health Interventions: Exploring the Use of Common Relationship Factors
JMIR Mhealth Uhealth 2019;7(4):e11245
URL: http://mhealth.jmir.org/2019/4/e11245/
doi:10.2196/11245
PMID:30985281

Original Paper

The Comparative Effectiveness of Mobile Phone Interventions in Improving Health Outcomes: Meta-Analytic Review

Qinghua Yang[1], MA, PhD; Stephanie K Van Stee[2], MA, PhD

[1]Department of Communication Studies, Bob Schieffer College of Communication, Texas Christian University, Fort Worth, TX, United States

[2]Department of Communication and Media, University of Missouri–St. Louis, St. Louis, MO, United States

Corresponding Author:
Qinghua Yang, MA, PhD
Department of Communication Studies
Bob Schieffer College of Communication
Texas Christian University
2805 South University Drive
Moudy South, Room 315
Fort Worth, TX, 76129
United States
Phone: 1 9176804053
Email: q.yang@tcu.edu

Abstract

Background: As mobile technology continues expanding, researchers have been using mobile phones to conduct health interventions (mobile health—mHealth—interventions). The multiple features of mobile phones offer great opportunities to disseminate large-scale, cost-efficient, and tailored messages to participants. However, the interventions to date have shown mixed results, with a large variance of effect sizes (Cohen *d*=−0.62 to 1.65).

Objective: The study aimed to generate cumulative knowledge that informs mHealth intervention research. The aims were twofold: (1) to calculate an overall effect magnitude for mHealth interventions compared with alternative interventions or conditions, and (2) to analyze potential moderators of mHealth interventions' comparative efficacy.

Methods: Comprehensive searches of the *Communication & Mass Media Complete*, *PsycINFO*, *Web of Knowledge*, *Academic Search Premier, PubMed* and *MEDLINE* databases were conducted to identify potentially eligible studies in peer-reviewed journals, conference proceedings, and dissertations and theses. Search queries were formulated using a combination of search terms: "intervention" (Title or Abstract) AND "health" (Title or Abstract) AND "*phone*" OR "black-berr*" (OR mHealth OR "application*" OR app* OR mobile OR cellular OR "short messag*" OR palm* OR iPhone* OR MP3* OR MP4* OR iPod*) (Title or Abstract). Cohen *d* was computed as the basic unit of analysis, and the variance-weighted analysis was implemented to compute the overall effect size under a random-effects model. Analysis of variance–like and meta-regression models were conducted to analyze categorical and continuous moderators, respectively.

Results: The search resulted in 3424 potential studies, the abstracts (and full text, as necessary) of which were reviewed for relevance. Studies were screened in multiple stages using explicit inclusion and exclusion criteria, and citations were evaluated for inclusion of qualified studies. A total of 64 studies were included in the current meta-analysis. Results showed that mHealth interventions are relatively more effective than comparison interventions or conditions, with a small but significant overall weighted effect size (Cohen *d*=0.31). In addition, the effects of interventions are moderated by theoretical paradigm, 3 engagement types (ie, changing personal environment, reinforcement tracking, social presentation), mobile use type, intervention channel, and length of follow-up.

Conclusions: To the best of our knowledge, this is the most comprehensive meta-analysis to date that examined the overall effectiveness of mHealth interventions across health topics and is the first study that statistically tested moderators. Our findings not only shed light on intervention design using mobile phones, but also provide new directions for research in health communication and promotion using new media. Future research scholarship is needed to examine the effectiveness of mHealth interventions across various health issues, especially those that have not yet been investigated (eg, substance use, sexual health), engaging participants using social features on mobile phones, and designing tailored mHealth interventions for diverse subpopulations to maximize effects.

(JMIR Mhealth Uhealth 2019;7(4):e11244) doi:10.2196/11244

KEYWORDS

meta-analysis; mobile phones; mHealth; intervention study

Introduction

Background

As mobile technology continues to expand and mobile phones become ubiquitous, with 73% of Americans actively using mobile phones [1] and approximately 7 billion mobile phone subscriptions worldwide [2], mobile phones are being increasingly used to conduct health interventions (mobile health—mHealth—interventions) to improve health conditions [3]. Studies have found several advantages of mHealth interventions compared with traditional approaches. Given the number of mobile phone users, mHealth interventions have the potential to engage a large group of people at a relatively lower cost, making public health interventions more feasible and impactful [3-5]. As many mHealth intervention platforms (eg, text-based, apps) have become available to benefit interventions in differing ways [1], mHealth can now address many issues including a limited workforce, finances, and accessing difficult-to-reach groups [6]. Overall, the use of mobile phones as a part of health interventions has become an effective tool to potentially prevent and treat health issues [6].

Despite the promises of mHealth interventions, previous mHealth interventions have yielded conflicting results and inconsistent effect sizes (ESs), ranging from Cohen *d* of −0.62 [7] to 1.65 [8]. Such large variance in ESs makes a comprehensive meta-analysis with moderator analyses imperative to provide a clearer picture of the effectiveness of mHealth interventions. However, previous syntheses either focused on a specific health issue [9,10] or were constrained by small sample sizes and a lack of moderator analyses [3,4], leaving what factors account for variance in ESs unanswered.

To fill the gap and provide insights into the effectiveness of harnessing mobile technology for health interventions, we aim to do the following: (1) calculate an overall effect magnitude for mHealth interventions compared with alternative interventions or conditions, and (2) analyze potential theoretical, methodological, and demographic moderators of mHealth interventions' comparative effectiveness. To achieve these goals, meta-analysis was implemented as the research method, which enables researchers to conduct a sophisticated synthesis of quantitative research literature [11]. Applying this innovative approach, we identify the overall effect of mHealth interventions and significant moderators to provide guidance for mHealth intervention design and implementation.

Mobile Phone Use and Mobile Health

With the growing technological culture, mobile phones have become the most popular and widespread personal technology, with 95% of the Americans owning a cell phone of some kind and 77% owning a smartphone [12]. Mobile technologies also include personal digital assistant phones (eg, BlackBerry), portable media players (eg, MP3- and MP4-players), and handheld computers (eg, iPad). As the use of mobile technology has increased, public health professionals have begun to take advantage of the multiple platforms provided by mobile phones to serve a wide variety of purposes, such as physical activity, weight loss, smoking cessation, mental health, and chronic disease management.

"The use of mobile computing and communication technologies in health care and public health" [10] is referred to as mHealth, a rapidly expanding branch within eHealth. There are many effective strategies that utilize mobile phones to promote public health. In general, mobile phones have been used to share information about public health as it is economical, sustainable, and effective [13]. To benefit mobile phone users, public health professionals have begun to utilize mobile phone capabilities for prevention, management, and treatment of health issues [14]. Mobile phones have been primarily applied for health purposes through short message service (SMS) and app features of these technological devices. In particular, text messages sent through mobile phones were found as a simple and efficient option for health services, which could be affordably used to send tailored health messages and reminders to improve delivery to patients [15-17]. Mobile apps are also widely used to promote public health, which can integrate a variety of built-in interactive features. Therefore, this allows for the potential to target heterogeneous audiences to address specific needs with diverse outcomes [15]. These apps offer great potential for dynamic engagement of patients and providers in health care and an innovative approach to improve health outcomes [18].

Mobile Health Interventions

Health interventions have utilized mobile technology in a variety of ways to increase health knowledge, promote health education, and change health behaviors. With the increasing use of mHealth interventions, several approaches have emerged for utilizing mobile technologies to implement health interventions. A primary approach to incorporating mobile phones into health interventions is focusing on the voice and text features that have achieved significant improvements in compliance with medication adherence, asthma symptoms, glycated hemoglobin (HbA_{1c}) levels, stress levels, smoking cessation, and self-efficacy [19]. Leveraging the advantages of mHealth interventions has improved public health outcomes in these areas as it has helped individuals become more aware and take accountability for their health issues. Other approaches to mHealth interventions include self-monitoring techniques and real-time surveillance features of the technology [20]. It is increasingly popular to utilize mobile phones to track individuals' daily activity and provide reminders and motivational text-based messages to continue progression during the intervention [21]. As smartphones have become more popular, another approach is the social networking component that allows users to interact and share information using social media [22]. With the booming marketplace and the engaging features of mobile apps, mHealth interventions have been increasingly based on mobile apps to deliver health information

and modify health behavior and have significant influence on youths' health outcomes [1].

Mobile phones have become a source of interactive communication, providing numerous advantages in conducting health interventions, including widespread use of mobile technology across various socioeconomic groups, few geographical constraints compared with other media, and cost-effectiveness to reach a diverse and large population [2,23]. A unique and noteworthy advantage of mHealth interventions is the ability to target underdeveloped or underserved areas due to the enabling resources provided by mobile phones [14,24]. A review of mobile phone-based health interventions for noncommunicable disease management in sub-Saharan Africa reported that using apps on cellular phones can improve physical and mental health outcomes [14]. Incorporating mobile phones within public health education and promotion can be beneficial for a wide range of situations and populations.

Mixed Effects of Mobile Health Interventions and Research Questions

Despite mHealth interventions' great potential to be superior to health interventions using traditional approaches, empirical research to date has generated mixed results in terms of the efficacy of mHealth interventions. For example, one successful mHealth intervention led to lower levels of perceived stress for intervention participants relative to a waitlist control group (Cohen d=1.02) at 6-month follow-up [25]. However, other mHealth interventions performed no better or worse than comparison conditions. For example, Cho et al [26] found that an mHealth intervention was less successful than a control condition (Cohen d=−0.24) at improving fasting plasma glucose levels of patients with type 2 diabetes. Such inconsistent findings make the efficacy of mHealth interventions in improving health outcomes unclear and indicate the necessity for a meta-analytic review.

To assess the overall comparative effect of mHealth interventions, the first research question (*RQ1*) was posed as follows: What is the overall effect magnitude for mHealth interventions compared with alternative interventions or conditions in improving health outcomes?

To take a close examination of the large variation of ESs in the efficacy of mHealth interventions, the first potential moderating variable is the theoretical framework applied in these studies. Theory enriches and provides a roadmap for research practices [5]. Therefore, it is plausible to assume that theory-based mHealth interventions may be more efficacious than their counterparts. A wide variety of theories have been applied in the mHealth interventions to date, including behavioral theories (eg, health belief model, HBM [27]; theory of planned behavior, TPB [28]), cognitive theories (eg, social cognitive theory, SCT [29]), or behavioral and cognitive theories (eg, cognitive behavioral therapy, CBT [30]). However, which theoretical paradigm works best in mHealth interventions remains unclear, especially when previous meta-analyses documented that the use of theory to guide intervention development does not significantly moderate ESs [1].

Besides investigating the theoretical framework, previous research [31-33] suggests that health topic, intervention designs (eg, control group design, length of intervention and follow-up), and participants' features (eg, age, gender, and health conditions) could moderate the effects of health interventions. Specifically, a meta-analysis on health interventions using social networking sites reported that studies using a *true* control condition without giving any intervention had a significantly higher weighted mean ES than studies giving an alternative intervention to the control group [33]. To examine the comparative effectiveness instead of the absolute effectiveness of mHealth interventions, it is crucial to take into consideration the control group design, including regular treatment [34], print-version [35] or computer-version [36] interventions, less intensive version of mHealth interventions [37], or interventions combining multiple channels [38]. Therefore, these methodological and demographic variables suggested by previous studies will be analyzed as potential moderators.

In addition, several mobile-phone-related features will also be analyzed as potential moderators. First, mobile phones have been applied in health interventions through different strategies. Some studies only used SMS [38,39] or mobile apps [40,41], whereas others combined both SMS and mobile apps [42,43]. Second and relatedly, there are not only interventions that applied mobile phone as the only channel [7,17] but also those that combined mobile phone with either face-to-face communication [39,44], another type of media [8,37], or both [34,45]. Although some researchers suggested that unimodal interventions could provide participants with more exposure and be easier to manage [46,47], leading to higher effectiveness, others advocated for multimodal interventions, which are more likely to engage participants and therefore function better in health promotion [48] or reported no difference between them [1]. Furthermore, Sama et al proposed a typology of 8 types of mobile phone engagement—changing personal environment, facilitating social support, goal setting, progress tracking, reinforcement tracking, self-monitoring, social presentation, and social referencing [18]. However, the types of engagement that improve the effectiveness of mHealth interventions remain unclear and will be examined in this study.

RQ2: Is the comparative effectiveness of mHealth interventions moderated by (1) theoretical paradigm, (2) health topics, (3) types of engagement, (4) mobile use type, (5) intervention channel, (6) control group design, (7) length of intervention, (8) length of follow up, and (9) participants' features?

Methods

Literature Search

To provide a clear picture of mHealth interventions' effectiveness in improving health outcomes, comprehensive searches of the *Communication & Mass Media Complete*, *PsycINFO*, *Web of Knowledge*, *Academic Search Premier*, *PubMed* and *MEDLINE* databases were conducted to identify potentially eligible studies in peer-reviewed journals and conference proceedings as well as dissertations and theses, which have been published through December 31, 2017. Search queries were formulated using a combination of search terms:

"intervention" (Title or Abstract) AND "health" (Title or Abstract) AND "*phone*" OR "black-berr*" (OR mHealth OR "application*" OR app* OR mobile OR cellular OR "short messag*" OR palm* OR iPhone* OR MP3* OR MP4* OR iPod*) (Title or Abstract). We retrieved 3506 studies from the databases, the abstracts (and full text, as necessary) of which were reviewed for relevance. Studies were screened in multiple stages using explicit inclusion and exclusion criteria (see Figure 1). We also screened the primary research articles included in a systematic review of mHealth-focused systematic reviews (n=546; [49]) to evaluate them for inclusion.

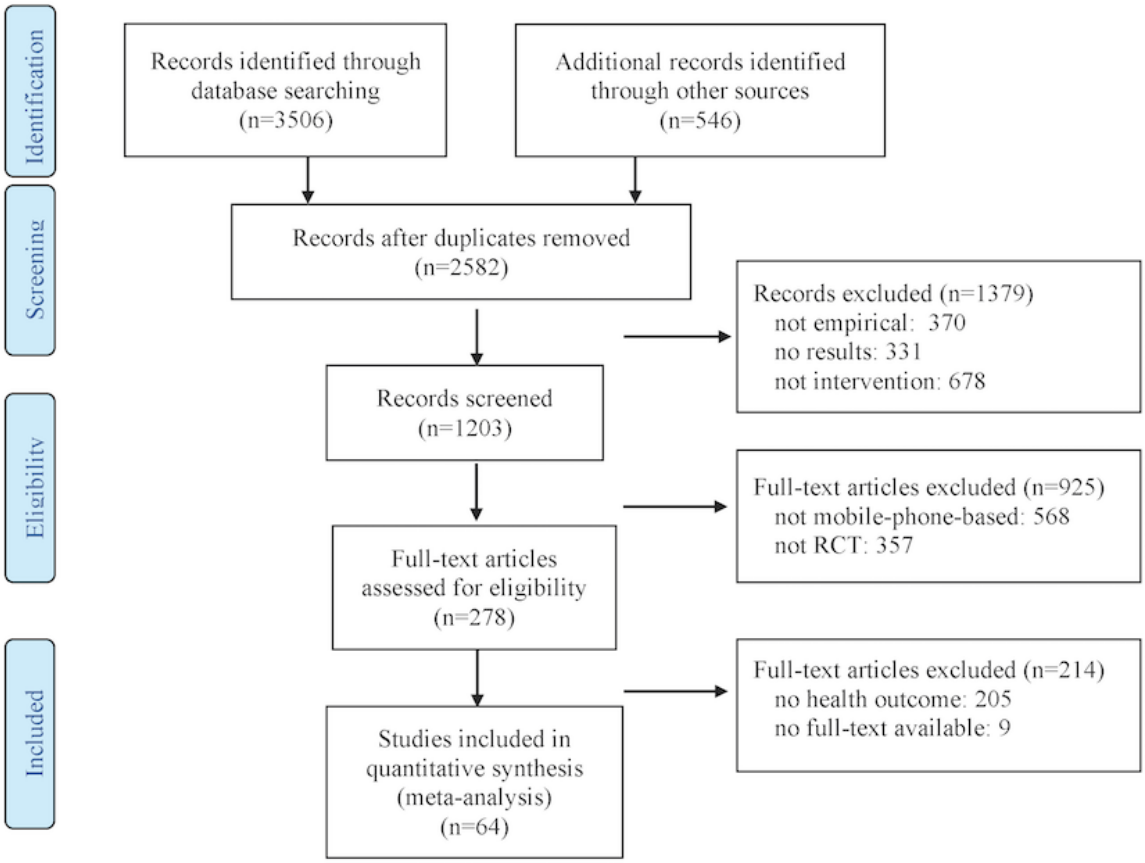
Overview of Meta-Analysis

As generally recommended in the meta-analysis methodological literature [50,51], Cohen *d* was computed as the basic unit of analysis for the meta-analytic review. The statistical analyses were based on methods proposed by Hedges and Olkin [52]. As publication bias may exist when the publication status depends on the statistical significance of study results [53], multiple analytic approaches were implemented to check for publication bias. First, a funnel plot was used to examine whether ESs from smaller studies show more variability than those from larger studies. Given that the funnel plot interpretation was open to subjectivity, Rosenthal's Fail-safe *N* and Duval and Tweedie's Trim and Fill method were also applied to provide statistical evidence of publication bias.

The current meta-analysis used the variance-weighted analysis [52]: the overall weighted ES was computed by weighting the unbiased ES (*d*) by the inverse of its associated variance (Wi=1/Vi). The overall homogeneity of ESs was tested using *Q* statistics to determine whether all effects were from the same population. When *Q* statistics are significant, the ESs are not from the same population, and the overall ES should be computed under the random-effects models (REMs), which incorporates between-studies uncertainty in the computation [54]. Otherwise, fixed-effects model (FEM) would be used.

In the moderator analysis, analysis of variance–like categorical models were conducted to analyze categorical moderators (eg, health topic, mobile use type) using mixed-effects models (MEM), as FEM with categorical moderator assumes that all studies in 1 subgroup share a common ES, whereas the MEM allows true variation of effects within subgroups of studies [55]. The same approach was applied when using meta-regression modeling to analyze continuous moderators (eg, length of follow-up and participants' age). In the cases where moderator analyses were statistically significant under the MEM, posthoc analysis was conducted for pairwise comparison using Tukey contrasts with adjusted *P* value. The analyses were conducted using *Metafor* and *Multcomp* package in R software (R Foundation for Statistical Computing).

Figure 1. Summary of selection process used in this study. Interventions using mobile phones only for data collection or making phone calls were excluded in this meta-analysis. RCT: randomized controlled trial.



Results

Study Description

A total of 64 studies were included in the current meta-analysis (see Multimedia Appendix 1), with 142 ESs computed following Schmidt and Hunter's approach [56]. Among the 64 studies, 47 were based on at least one theory, including SCT [7,8,26,30,35,37,39,40,57-73], transtheoretical model [17,38,58,74-78], self-regulation theory [39,68,79], self-determination theory [42,43], HBM [45,57,66,80], and theory of reasoned action and/or TPB [7,8,40,60]. The meta-analyzed mHealth interventions focused on 5 topics, namely mental health [25,36,45,48,56,81-84], nutrition and weight status [17,34,35,38,39,42,57,69,74-76], physical activity [7,8,26,37,56,57,74-76,78,79,84,85], health-related quality of life and well-being [47,61,73,86], and chronic disease management [8,23,26,40,44,58,62,63,72,75,77,79,87-99]. The categorization of health topics was based on Healthy People 2020 [100]. There were originally 8 categories in the code book (ie, 1=tobacco use, 2=mental health, 3=a nutrition and weight status, 4=physical activity, 5=sexual health, 6=health-related quality of life and well-being, 7=HIV/AIDS, 8=chronic disease management). However, as no study focused on tobacco use, sexual health, or HIV/AIDS, these 3 categories were excluded from analyses. There were 10,296 (N=10,296) participants across included studies.

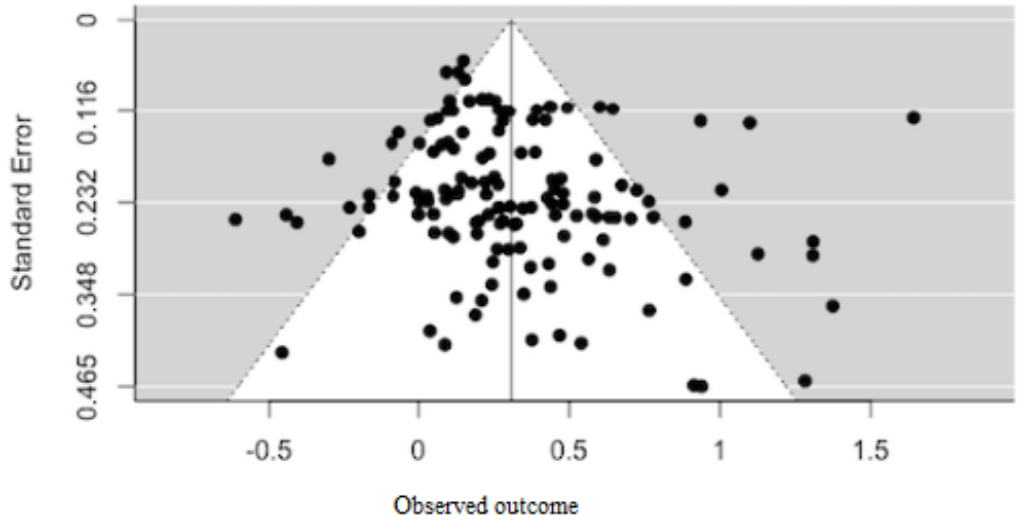
Publication Bias

Publication bias may exist when publication status depends on the statistical significance of study results [52]. We have applied multiple techniques to check for potential publication bias. First, a funnel plot can be used to examine whether ESs from smaller studies show more variability than those from larger studies. As shown in Figure 2, the funnel plot of ESs seems to be generally symmetric, which is consistent with the Regression Test for funnel plot asymmetry (z=1.51, P=.13) and provides evidence for the absence of publication bias. Rosenthal's Fail-safe N was 17,539, which is much larger than the tolerance level (5k +10=660), and no study was found missing for symmetry using Duval and Tweedie's Trim and Fill Method, which further confirmed the absence of publication bias.

Overall Analysis

Estimated under the FEM, the Q statistic was significant (Q_{total} (df=141)=467.01, P<.001), indicating that the ESs were not homogeneous, and mean ES was estimated under the REM using Restricted Maximum Likelihood Estimation method. Under the REM, the sample weighted mean for standardized mean difference was 0.31 (95% CIs; 0.25, 0.36), which is a small ES [101], but statistically significant (P<.001). In other words, there was a statistically significant mean difference between the mHealth intervention and control groups according to the overall analysis. I^2, an index representing the ratio of true heterogeneity to total variance across observed ESs, is 71.87%, indicating large between-study variance [102]. Similarly, Birge ratio, another index to quantify the magnitude of heterogeneity (computed as Q /df=467.01/141=3.31), is larger than 1 (the ratio when all the variance comes from sampling error), indicating large between-study heterogeneity. Sampling error variance (S_e^2=.0135) only accounted for 18.91% of the total variance (S^2=.0714), suggesting the presence of a moderator. Therefore, the moderators proposed in *RQ2* were analyzed.

Figure 2. Funnel plot of effect sizes to check publication bias for this study.



Moderator Analyses

Moderator analyses were conducted by analyzing *theoretical paradigm* (1=no theory, 2=behavioral theory, 3=cognitive theory, 4=behavioral and cognitive theories combined), *health topic* (1=mental health, 2=nutrition and weight status, 3=physical activity, 4=health-related quality of life and well-being, 5=chronic disease management), *eight types of engagement* [18], *mobile use type* (1=text messages, 2=mobile app, 3=combined), *intervention channel* (1=mobile phone only, 2=mobile phone combined with other type of media, 3=mobile phone combined with face-to-face communication, 4=mobile phone combined with other type of media and face-to-face communication), *control group design* (1=no intervention, eg, waiting list group; 2=intervention based on interpersonal communication [no media involved, eg, counseling at clinic], 3=intervention using other type of media than mobile phone, eg, website; 4=intervention using mobile phone, eg, text messages about general health information instead of targeted health behavior; 5=intervention with multiple features), and *participants' health condition*: (1=general healthy adults, 2=population at risk) as categorical moderators respectively. Moreover, *participants' mean age*, *percentage of female participants* (to examine the influence of gender), *length of the intervention*, and *length of follow-up* were analyzed as continuous moderators.

Theoretical Paradigm

Under MEM, theoretical paradigm was significant as a moderator ($Q_{between}$ (*df*=3)=11.01, *P*=.01). Posthoc pairwise comparison indicated that mHealth interventions based on cognitive and behavioral theories combined (*d*=0.45, *P*<.001) had the highest weighted mean ES among the 4 categories and was significantly higher than interventions not indicating a theory (*d*=0.28, *P*<.001) or interventions applying behavioral (*d*=0.23, *P*<.001) both at .05 level or cognitive theory only (*d*=0.20, *P*<.001) at .01 level.

Health Topic

Under MEM, health topic was not a significant moderator ($Q_{between}$ (*df*= 4)=1.63, *P*=.80), with mHealth interventions on physical activity showing the lowest weighted mean ES (*d*=0.24; SE=0.06; 95% CIs 0.11, 0.36; *P*<.001) and interventions on nutrition and weight status showing the highest ES (*d*=0.36; SE=.08; 95% CIs 0.21, 0.52; *P*<.001), which however are not significantly different from each other. The weighted mean ESs across all 5 topics were significantly larger than zero.

Types of Engagement

Among the 8 types of engagement proposed by Sama et al [18], only changing personal environment ($Q_{between}$ (*df*=1)=9.44, *P*=.002), reinforcement tracking ($Q_{between}$ (*df*=1)=10.24, *P*=.001), and social presentation or announcement ($Q_{between}$ (*df*=1)=6.42, *P*=.01) were significant moderators. Specifically, the weighted mean ES of the mHealth interventions with the function of changing personal environment (*d*=0.76; SE=0.15; 95% CIs 0.47,1.05; *P*<.001) was significantly higher than that of the interventions without this feature (*d*= 0.29; SE=0.03; 95% CIs 0.24, 0.35; *P*<.001) at .01 level (*z*=3.11). Similarly, the weighted mean ES of the mHealth interventions with the reinforcement tracking function (*d*=0.43; SE=0.05; 95% CIs 0.33, 0.53; *P*<.001) was significantly higher than that of the interventions without this function (*d*=0.24; SE=0.03; 95% CIs 0.18, 0.30; *P*<.001) at .01 level (*z*=3.18). However, the weighted mean ES of the mHealth interventions with the social presentation or announcement function (*d*=0.04; SE=0.08; 95% CIs -0.13, 0.20; *P*=.65) was significantly lower than that of the interventions without (*d*=0.33; SE=.03; 95% CIs 0.27, 0.38; *P*<.001) at .05 level (*z*=2.53).

Mobile Use Type

How the mobile phone was applied in the intervention was found as a significant moderator ($Q_{between}$ (*df*=2)=17.35, *P*<.001). Post hoc analysis indicated that the weighted mean ES of interventions combining SMS and mobile apps (*d*=0.59; SE=.12; 95% CIs 0.36, 0.83; *P*<.001) was significantly higher than the ESs of those using only SMS (*d*=0.30; SE=0.04; 95% CIs 0.23, 0.38; *P*<.001) or mobile apps (*d*=0.22; SE=.03; 95% CIs 0.16, 0.29; *P*<.001; Z_{SMS}=3.37, Z_{App}=4.16).

Intervention Channel

The channel through which mHealth interventions were implemented turned out to be another significant moderator ($Q_{between}$ (*df*=3)=12.56, *P*=.006). Pairwise comparison indicated that the weighted mean ES of interventions combining mobile phone with another type of media (*d*=0.39; SE=.05 95% CIs 0.30, 0.49; *P*<.001) was significantly higher than that of interventions using mobile phone only (*d*=0.19; SE=.04; 95% CIs 0.12, 0.26; *P*<.001; *z*=2.63) or interventions combining mobile phone with face-to-face communication (*d*=0.18; SE=.04; *K*=14; 95% CIs 0.10, 0.25; *P*<.001; *z*=2.74).

Control Group Design

Nonsignificant differences between-study variance under MEM ($Q_{between}$ (*df*=4)=5.21, *P*=.27) indicated that control group design did not show a significant difference in the comparative effectiveness of the interventions.

Participants' Age, Gender, and Health Condition

The majority of mHealth interventions in the current sample were conducted with at-risk populations, except for 6 studies [7,42,71,61,78,85]. However, whether the participants were healthy or at-risk populations was not a significant moderator of the ESs ($Q_{between}$ (*df*=1)=2.31, *P*=.13). Neither participants' age ($Q_{between}$ (*df*= 1)=.99, *P*=.32) nor gender ($Q_{between}$ (*df*= 1)=1.94, *P*=.16) was a significant moderator.

Length of Intervention and Follow-Up

Intervention length was a nonsignificant moderator ($Q_{between}$ (*df*=1)=0.57, *P*=.45); however, when the follow-up measures were conducted, it did moderate the ESs ($Q_{between}$ (*df*=1)=13.46, *P*<.001). Length of follow-up ranged from immediate [57] to 9 months later [71], with an average follow-up period being 2.48 weeks (SD=6.31). The weighted mean ES was significant immediately after the intervention (*d*=0.27, *P*<.001), and increased by .015 for each additional week of follow-up (*P*<.001).

Discussion

Overall Effects

Findings from this meta-analysis indicated that mHealth interventions are significantly more effective than comparison conditions at improving health outcomes (d=0.31; 95% CIs 0.25, 0. 36), which is consistent with previous meta-analyses focusing on specific health issues [9,103,104]. In particular, mHealth interventions have been significantly more effective than comparison conditions for physical activity (Hedge g=0.54; 95% CIs 0.17, 0.91] [9] and led to significant improvement in diabetes management (mean 0.5% reduction in HbA_{1c}) [104]. As it relates to SMS, text interventions were more effective for antiretroviral therapy adherence than control conditions (OR 1.39, CI 1.18, 1.64) [103]. Our finding related to the relative effectiveness of mHealth interventions shows not only consistency with previous research but extends current research by examining the effects across health contexts. In addition to the significant overall effect of mHealth interventions, several moderators were identified, which help explain the mechanisms behind the variance in mHealth interventions' efficacy.

Effects of Engagement

Findings indicated several statistically significant moderators of mHealth intervention effects. Of these, 1 such moderator is the type of engagement (ie, changing personal environment, reinforcement tracking, and social presentation or announcement). mHealth interventions that included features for changing one's personal environment and/or reinforcement tracking exhibited larger relative effects than mHealth interventions without those features. Changing a person's personal environment directly enables people to engage in the desired behavior or a behavior that affects the desired health outcome (eg, soothing sounds for meditating, which may help with stress) [18]. Resources that allow users to immediately engage in the desired behavior (or a behavior that affects the desired health outcome) may help to reduce barriers that may otherwise prevent them from engaging in healthy behaviors that improve health outcomes. Theories of health behavior change (eg, HBM and TPB) address the important role that perceived or actual barriers play, indicating that reducing perceived or actual barriers can enhance the likelihood of positive behavior change. Alternatively, mHealth interventions that included the feature of social presentation or announcement exhibited smaller relative effects than mHealth interventions without that feature. This finding is counterintuitive, as one might expect that a social presentation or announcement feature would have greater positive effects on health outcomes, as providing information to others about one's accomplishments may increase motivation [105]. A possible explanation for this is that social presentation or announcement may serve as a distraction to participants. Previous research has found that features intended to draw in or engage audiences may actually distract them from the position advocated by messages [106]. Alternatively, the lower relative effects of mHealth interventions with a social presentation or announcement feature may be due largely to the 1 study with this feature that contributed the largest negative ES in the meta-analysis [7].

Effects of Intervention Channel

mHealth interventions using both SMS and a mobile app were relatively more effective than interventions using either SMS or a mobile app. SMS were typically used to collect data about participants' behavior [45], and/or provide reinforcement for desired behaviors [67,85]. Furthermore, mHealth interventions that included other media were relatively more effective than interventions that included face-to-face components. Many of the mHealth interventions that included additional media channels used websites, emails, and/or print materials. Interventions with face-to-face communication as an additional channel typically used in-person health care or counseling or group workshops. mHealth interventions incorporating other media may be more effective because they provide additional content exposure, use complementary strategies, and/or drive people to the mobile intervention components, which is helpful in delivering messages to users with a variety of media use habits [107], especially for those whose preferred medium is not mobile phone.

Length of Follow-Up

mHealth interventions with a longer follow-up exhibited greater effects than those with a shorter follow-up. Previous research indicates that length of follow-up can serve as a moderator of relative intervention effects [31]; however, in previous research, greater effectiveness was shown for shorter-term follow up [33], which is the opposite of the findings of this study. The finding may reflect the fact that some mHealth intervention studies used apps that are freely available, commercial apps [68], which participants could continue using after the intervention period. This finding highlights the promise of mHealth interventions in promoting positive long-term health outcomes. The easy accessibility and cost efficiency of mobile features may help prevent diminishing intervention effects compared with those that use other types of new media [33].

Theoretical and Practical Implications

Our moderator analyses found that studies based on both cognitive and behavioral theories were more effective than those based on no theory or behavioral or cognitive theory only. Cognitive theories, such as SCT [29] or self-regulation theory [108], mainly focus on psychological factors and inner thoughts. Although they are related to behavior, there is still a gap between intention and actual behavior due to internal and external barriers [28], and how to bridge this gap and eventually trigger behavioral change are not well-addressed. Given this, scholars have critiqued the use of traditional psychosocial theories, which could fail to maximize the opportunities offered by mHealth interventions and called for closer assessment of each mHealth program and user engagement [109]. On the other hand, perceived and actual barriers are key variables in behavioral theories, such as HBM [27] or TPB [28], which acknowledge and emphasize reducing these barriers to achieve positive behavioral change. However, behavior is rooted in cognitions, which are considered precursors to behavioral change; positive behavioral change that is lacking a strong cognitive foundation may not have longitudinal effects. Therefore, the complementarity of cognitive and behavioral theories may explain the large ESs of studies that applied both types of

theories [62,93]. The convergence of cognitive and behavioral theories has observed success in CBT [110] and deserves further investigation by health communication researchers.

In terms of mHealth intervention design, given that tailored communication is effective in promoting health behavior change and health outcomes [31], it is important to enable users to change their personal environment using mobile phones and provide personalized reinforcement messages based on users' progress on health outcomes. According to the Elaboration Likelihood Model [111], personally relevant messages are more likely to increase personal involvement and trigger central route message processing, which would achieve stable persuasive effects over time. Alternatively, health researchers and professionals should take caution when incorporating social presentation or announcement features to engage participants and to avoid distraction. More research examining how to strategically integrate social engagement features [18] into mobile phones without affecting message exposure [46,47] to improve health outcomes is needed.

When designing mobile interventions, it is worth considering combining both SMS and app features, which are available on most mobile phones. SMS is easier to implement, whereas apps could afford multimedia interactive features; such features would complement each other to potentially maximize user engagement and health outcomes. According to channel complementary theory [103], people use multiple sources, which serve different niches and present unique information, to acquire information in certain health topics. Therefore, health researchers and professionals could also take advantage of other types of media when designing mHealth interventions, especially internet, through which several studies have achieved high efficacy [8,25,36]. Finally, due to the easy accessibility and cost efficiency of mobile phones, our finding also highlights the promise of mHealth interventions in achieving long-term health effects. Thus, researchers aiming to improve health outcomes over a long period could base interventions on existing low-cost mobile apps or free SMS to enable sustained use of the target mobile features and consequently maintain positive effects.

Limitations and Future Research

Despite this study's pioneering efforts, several limitations should be noted. First, although this study started from a comprehensive literature search and included more studies than previous reviews ([3,4,10], the sample size of several specific categories remains small. For instance, none of the studies meeting inclusion criteria focused on reducing substance use, promoting sexual health, or preventing HIV/AIDS. Among the health issues in the current meta-analysis, only 4 studies focused on health-related quality of life [47,61,72,112]. In the same vein, both the social presentation [7,48,61,68] and social referencing [7,8,48,85] engagement types have been applied in only 4 studies. The comparatively small sample sizes of the randomized controlled trials (RCTs)—attributed to the recency of mHealth interventions and the long implementing and publishing process—could limit the reliability of statistical results in specific categories and increase the likelihood of chance differences. As such, more empirical evidence is needed to have a more reliable estimation of the moderating effects in mHealth interventions, especially in understudied areas.

Second, in an attempt to be as comprehensive as possible, this study included not only published studies but also conference papers and unpublished work, which makes the study vulnerable to inclusion of lower quality studies, a general limitation for meta-analytic research [54]. However, only 2 included studies contributing 5 ESs were dissertations; all other included studies appeared in peer-reviewed journal articles.

Finally, despite efforts to include only the most relevant RCT studies, this study is susceptible to the "apples and oranges" critique concerning the comparability of studies, a common concern for meta-analytic reviews [54]. Comparability refers to all the studies included in one meta-analysis examining the same constructs or relationships. Due to the relatively small sample size in some categories, such as the variety of specific media that were used in combination with mobile phones in interventions, they were grouped in a larger category for moderator analyses.

Conclusions

mHealth interventions have become increasingly common in recent years, given the ubiquity of mobile phone ownership, providing the possibility to tailor health messages, cost efficiency of mobile delivery channels, and opportunity for large-scale dissemination. To the best of our knowledge, this is the most comprehensive meta-analysis to date that examined the overall effectiveness of mobile phone interventions across health topics and is the first study across health behaviors that statistically tested moderators. By analyzing 64 studies, we found that mHealth interventions have a small but significant weighted mean ES (Cohen d=0.31), which is moderated by theoretical paradigm, engagement types (ie, changing personal environment, reinforcement tracking, and social presentation), mobile use type, intervention channel, and length of follow-up. Our findings not only shed light on intervention design using mobile phones, but also provide new directions for research in health communication and promotion using new media. Future research is needed to examine the effectiveness of mHealth interventions across various health issues, especially those that have not yet been fully investigated (eg, substance abuse and sexual health), engaging participants using social features on mobile phones, and designing tailored mHealth interventions for diverse subpopulations to maximize effects.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of studies included in the meta-analysis.

[PDF File (Adobe PDF File), 140KB - mhealth_v7i4e11244_app1.pdf]

References

1. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. JAMA Pediatr 2017 May 01;171(5):461-469. [doi: 10.1001/jamapediatrics.2017.0042] [Medline: 28319239]
2. Voth EC, Oelke ND, Jung ME. A theory-based exercise app to enhance exercise adherence: a pilot study. JMIR Mhealth Uhealth 2016 Jun 15;4(2):e62 [FREE Full text] [doi: 10.2196/mhealth.4997] [Medline: 27307134]
3. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
4. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013 Jan;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
5. Glanz K, Rimer B, Viswanath K, editors. Health Behavior And Health Education: Theory, Research, And Practice. San Francisco: Jossey-Bass; 2008.
6. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. Telemed J E Health 2014 Jan;20(1):75-82 [FREE Full text] [doi: 10.1089/tmj.2012.0328] [Medline: 24205809]
7. Du H, Venkatakrishnan A, Youngblood GM, Ram A, Pirolli P. A group-based mobile application to increase adherence in exercise and nutrition programs: a factorial design feasibility study. JMIR Mhealth Uhealth 2016 Jan;4(1):e4 [FREE Full text] [doi: 10.2196/mhealth.4900] [Medline: 26772910]
8. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons with prediabetes. J Med Internet Res 2015 Oct;17(10):e240 [FREE Full text] [doi: 10.2196/jmir.4897] [Medline: 26499966]
9. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012 Nov;14(6):e161 [FREE Full text] [doi: 10.2196/jmir.2171] [Medline: 23171838]
10. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
11. Noar SM. In pursuit of cumulative knowledge in health communication: the role of meta-analysis. Health Commun 2006 Jul;20(2):169-175. [doi: 10.1207/s15327027hc2002_8] [Medline: 16965254]
12. Pew Research Center. 2018. Mobile fact sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-02-02] [WebCite Cache ID 75ttWf3Il]
13. Brinkel J, Krämer A, Krumkamp R, May J, Fobil J. Mobile phone-based mHealth approaches for public health surveillance in Sub- Saharan Africa: a systematic review. IJERPH 2014 Nov 12;11(11):11559-11582. [doi: 10.3390/ijerph111111559] [Medline: 25396767]
14. Opoku D, Stephani V, Quentin W. A realist review of mobile phone-based health interventions for non-communicable disease management in sub-Saharan Africa. BMC Med 2017 Feb 06;15(1):24 [FREE Full text] [doi: 10.1186/s12916-017-0782-z] [Medline: 28162090]
15. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013 May 21;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
16. Montes JM, Medina E, Gomez-Beneyto M, Maurino J. A short message service (SMS)-based strategy for enhancing adherence to antipsychotic medication in schizophrenia. Psychiatry Res 2012 Dec 30;200(2-3):89-95. [doi: 10.1016/j.psychres.2012.07.034] [Medline: 22901437]
17. Shaw R. Duke University. 2012. A mobile health intervention to sustain recent weight loss (Doctoral dissertation, Duke University) URL: https://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/5866/Shaw_duke_0066D_11609.pdf?sequence=1 [accessed 2019-02-02] [WebCite Cache ID 75ttkDc2k]
18. Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. JMIR Mhealth Uhealth 2014 May;2(2):e19 [FREE Full text] [doi: 10.2196/mhealth.3088] [Medline: 25099179]
19. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemedicine and e-Health 2009;15(3). [doi: 10.1089/tmj.1008.0099]
20. Tate EB, Spruijt-Metz D, O'Reilly G, Jordan-Marsh M, Gotsis M, Pentz MA, et al. mHealth approaches to child obesity prevention: successes, unique challenges, and next directions. Transl Behav Med 2013 Dec;3(4):406-415 [FREE Full text] [doi: 10.1007/s13142-013-0222-3] [Medline: 24294329]

21. Martin SS, Feldman DI, Blumenthal RS, Jones SR, Post WS, McKibben RA, et al. mActive: a randomized clinical trial of an automated mHealth intervention for physical activity promotion. J Am Heart Assoc 2015 Nov 09;4(11) [FREE Full text] [doi: 10.1161/JAHA.115.002239] [Medline: 26553211]
22. Maher C, Ferguson M, Vandelanotte C, Plotnikoff R, De Bourdeaudhuij I, Thomas S, et al. A web-based, social networking physical activity intervention for insufficiently active adults delivered via Facebook app: randomized controlled trial. J Med Internet Res 2015 Jul;17(7):e174 [FREE Full text] [doi: 10.2196/jmir.4086] [Medline: 26169067]
23. Quinn CC, Sareh PL, Shardell ML, Terrin ML, Barr EA, Gruber-Baldini AL. Mobile diabetes intervention for glycemic control: impact on physician prescribing. J Diabetes Sci Technol 2014 Mar;8(2):362-370 [FREE Full text] [doi: 10.1177/1932296813514503] [Medline: 24876589]
24. Krishnan A, Cravero C. A multipronged evidence-based approach to implement mHealth for underserved HIV-infected populations. Mobile Media Commun 2017 Mar 23;5(2):194-211. [doi: 10.1177/2050157917692390]
25. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: a randomized controlled trial. J Med Internet Res 2016 Jan 27;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline: 26818683]
26. Cho JH, Lee HC, Lim DJ, Kwon HS, Yoon KH. Mobile communication using a mobile phone with a glucometer for glucose control in Type 2 patients with diabetes: as effective as an Internet-based glucose monitoring system. J Telemed Telecare 2009 Mar;15(2):77-82. [doi: 10.1258/jtt.2008.080412] [Medline: 19246607]
27. Rosenstock IM. Historical origins of the health belief model. Health Educ Monogr 1974;2:328-335. [doi: 10.1177/109019817400200403]
28. Ajzen I. The theory of planned behavior. Organ Behav Hum Dec 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
29. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
30. Beck J. Cognitive Behavior Therapy: Basics and Beyond. New York, NY: The Guilford Press; 2011.
31. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
32. Snyder LB, Hamilton MA, Mitchell EW, Kiwanuka-Tondo J, Fleming-Milici F, Proctor D. A meta-analysis of the effect of mediated health communication campaigns on behavior change in the United States. J Health Commun 2004;9(Suppl 1):71-96. [doi: 10.1080/10810730490271548] [Medline: 14960405]
33. Yang Q. Are social networking sites making health behavior change interventions more effective? A meta-analytic review. J Health Commun 2017 Mar;22(3):223-233. [doi: 10.1080/10810730.2016.1271065] [Medline: 28248623]
34. Oh B, Cho B, Han MK, Choi H, Lee MN, Kang H, et al. The effectiveness of mobile phone-based care for weight control in metabolic syndrome patients: randomized controlled trial. JMIR Mhealth Uhealth 2015;3(3):e83 [FREE Full text] [doi: 10.2196/mhealth.4222] [Medline: 26293568]
35. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009 Jan;11(1):e1 [FREE Full text] [doi: 10.2196/jmir.1100] [Medline: 19141433]
36. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry 2013;13:49 [FREE Full text] [doi: 10.1186/1471-244X-13-49] [Medline: 23391304]
37. Newton J, Marker A, Allen H, Machtmes R, Han H, Johnson W, et al. Parent-targeted mobile phone intervention to increase physical activity in sedentary children: randomized pilot trial. JMIR Mhealth Uhealth 2014 Nov 10;2(4):e48 [FREE Full text] [doi: 10.2196/mhealth.3420] [Medline: 25386899]
38. Patrick K, Norman GJ, Davila EP, Calfas KJ, Raab F, Gottschalk M, et al. Outcomes of a 12-month technology-based intervention to promote weight loss in adolescents at risk for type 2 diabetes. J Diabetes Sci Technol 2013 May 01;7(3):759-770 [FREE Full text] [doi: 10.1177/193229681300700322] [Medline: 23759410]
39. de Niet J, Timman R, Bauer S, van den Akker E, Buijks H, de Klerk C, et al. The effect of a short message service maintenance treatment on body mass index and psychological well-being in overweight and obese children: a randomized controlled trial. Pediatr Obes 2012 Jun;7(3):205-219. [doi: 10.1111/j.2047-6310.2012.00048.x] [Medline: 22492669]
40. Irvine AB, Russell H, Manocchia M, Mino DE, Cox GT, Morgan R, et al. Mobile-Web app to self-manage low back pain: randomized controlled trial. J Med Internet Res 2015 Jan 02;17(1):e1 [FREE Full text] [doi: 10.2196/jmir.3130] [Medline: 25565416]
41. Mira JJ, Navarro I, Botella F, Borrás F, Nuño-Solinís R, Orozco D, et al. A Spanish pillbox app for elderly patients taking multiple medications: randomized controlled trial. J Med Internet Res 2014 Apr 04;16(4):e99 [FREE Full text] [doi: 10.2196/jmir.3269] [Medline: 24705022]
42. Kerr DA, Harray AJ, Pollard CM, Dhaliwal SS, Delp EJ, Howat PA, et al. The connecting health and technology study: a 6-month randomized controlled trial to improve nutrition behaviours using a mobile food record and text messaging support in young adults. Int J Behav Nutr Phys Act 2016 Apr 21;13:52 [FREE Full text] [doi: 10.1186/s12966-016-0376-8] [Medline: 27098449]

43. McGillicuddy JW, Gregoski MJ, Weiland AK, Rock RA, Brunner-Jackson BM, Patel SK, et al. Mobile health medication adherence and blood pressure control in renal transplant recipients: a proof-of-concept randomized controlled trial. JMIR Res Protoc 2013;2(2):e32 [FREE Full text] [doi: 10.2196/resprot.2633] [Medline: 24004517]
44. Wayne N, Perez DF, Kaplan DM, Ritvo P. ealth coaching reduces HbA1c in type 2 diabetic patients from a lower-socioeconomic status community: a randomized controlled trial. J Med Internet Res 2015 Oct 05;17(10):e224 [FREE Full text] [doi: 10.2196/jmir.4871] [Medline: 26441467]
45. Bhati S. George Mason University. 2015. The effect of the sleep support for moms intervention on postpartum sleep and depressive symptoms: A pilot randomized controlled trial URL: https://search.proquest.com/docview/1617971471?pq-origsite=summon [accessed 2019-02-02] [WebCite Cache ID 75tu6q4mA]
46. Rote AE, Klos LA, Brondino MJ, Harley AE, Swartz AM. The efficacy of a walking intervention using social media to increase physical activity: a randomized trial. J Phys Act Health 2015 Jun;12(Suppl 1):S18-S25. [doi: 10.1123/jpah.2014-0279] [Medline: 25599378]
47. Lee M, Wu H, Lin J, Tan T, Chan P, Chen Y. Development and evaluation of an E-health system to care for patients with bladder pain syndrome/interstitial cystitis. Int J Urol 2014 Apr;21(Suppl 1):62-68 [FREE Full text] [doi: 10.1111/iju.12336] [Medline: 24807502]
48. Lappalainen P, Kaipainen K, Lappalainen R, Hoffrén H, Myllymäki T, Kinnunen M, et al. Feasibility of a personal health technology-based psychological intervention for men with stress and mood problems: randomized controlled pilot trial. JMIR Res Protoc 2013 Jan;2(1):e1 [FREE Full text] [doi: 10.2196/resprot.2389] [Medline: 23611946]
49. 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]
50. Hunter J, Schmidt F. Methods of Meta-Analysis: Correcting error and bias in research findings. Thousand Oaks, CA: SAGE Publications; 2004.
51. Rosenthal R. Meta-Analytic Procedures for Social Research. Newbury Park, CA: Sage; 1991.
52. Hedges LV, Olkin I. Statistical Methods for Meta-Analysis. Orlando, FL: Academic; 1985.
53. Sutton AJ. Publication bias. In: Cooper BS, Hedges LV, Valentine JC, editors. The Handbook of Research Synthesis and Meta-Analysi. New York, NY: Russell Sage Foundation; 2009:435-452.
54. Raudenbush SW. Analyzing effect sizes: random-effects models. In: Cooper BS, Hedges LV, Valentine JC, editors. The Handbook of Research Synthesis and Meta-Analysis. New York, NY: Russell Sage Foundation; 2009:295-316.
55. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to Meta-Analysis. UK: John Wiley & Sons; 2009.
56. Schmidt FL, Hunter JE. Comparison of three meta-analysis methods revisited: an analysis of Johnson, Mullen, and Salas (1995). J Appl Psychol 1999;84(1):144-148. [doi: 10.1037/0021-9010.84.1.144]
57. Choi J, Lee JH, Vittinghoff E, Fukuoka Y. mHealth physical activity intervention: a randomized pilot study in physically inactive pregnant women. Matern Child Health J 2016 May;20(5):1091-1101. [doi: 10.1007/s10995-015-1895-7] [Medline: 26649879]
58. Abraham AA, Chow WC, So HK, Yip BH, Li AM, Kumta SM, et al. Lifestyle intervention using an internet-based curriculum with cell phone reminders for obese Chinese teens: a randomized controlled study. PLoS One 2015 May;10(5):e0125673 [FREE Full text] [doi: 10.1371/journal.pone.0125673] [Medline: 25946465]
59. Baron JS, Hirani S, Newman SP. A randomised, controlled trial of the effects of a mobile telehealth intervention on clinical and patient-reported outcomes in people with poorly controlled diabetes. J Telemed Telecare 2017 Feb;23(2):207-216. [doi: 10.1177/1357633X16631628] [Medline: 26880694]
60. Chau JP, Lee DT, Yu DS, Chow AY, Yu W, Chair S, et al. A feasibility study to investigate the acceptability and potential effectiveness of a telecare service for older people with chronic obstructive pulmonary disease. Int J Med Inform 2012 Oct;81(10):674-682. [doi: 10.1016/j.ijmedinf.2012.06.003] [Medline: 22789911]
61. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. J Am Med Assoc 2015 Sep;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
62. Cobb NK, Poirier J. Effectiveness of a multimodal online well-being intervention: a randomized controlled trial. Am J Prev Med 2014 Jan;46(1):41-48 [FREE Full text] [doi: 10.1016/j.amepre.2013.08.018] [Medline: 24355670]
63. Fountoulakis S, Papanastasiou L, Gryparis A, Markou A, Piaditis G. Impact and duration effect of telemonitoring on HbA1c, BMI and cost in insulin-treated Diabetes Mellitus patients with inadequate glycemic control: a randomized controlled study. Hormones (Athens) 2015 Jul;14(4):632-643 [FREE Full text] [doi: 10.14310/horm.2002.1603] [Medline: 26188234]
64. Frederix I, Hansen D, Coninx K, Vandervoort P, Vandijck D, Hens N, et al. Medium-term effectiveness of a comprehensive internet-based and patient-specific telerehabilitation program with text messaging support for cardiac patients: randomized controlled trial. J Med Internet Res 2015 Jul 23;17(7):e185. [doi: 10.2196/jmir.4799] [Medline: 26206311]
65. Istepanian RS, Zitouni K, Harry D, Moutosammy N, Sungoor A, Tang B, et al. Evaluation of a mobile phone telemonitoring system for glycaemic control in patients with diabetes. J Telemed Telecare 2009 Apr;15(3):125-128 [FREE Full text] [doi: 10.1258/jtt.2009.003006] [Medline: 19364893]

66. Johnson KB, Patterson BL, Ho YX, Chen Q, Nian H, Davison CL, et al. The feasibility of text reminders to improve medication adherence in adolescents with asthma. J Am Med Inform Assoc 2016 May;23(3):449-455 [FREE Full text] [doi: 10.1093/jamia/ocv158] [Medline: 26661717]
67. Kamal AK, Shaikh Q, Pasha O, Azam I, Islam M, Memon AA, et al. A randomized controlled behavioral intervention trial to improve medication adherence in adult stroke patients with prescription tailored Short Messaging Service (SMS)-SMS4Stroke study. BMC Neurol 2015 Oct 21;15:212 [FREE Full text] [doi: 10.1186/s12883-015-0471-5] [Medline: 26486857]
68. Kirwan M, Vandelanotte C, Fenning A, Duncan MJ. Diabetes self-management smartphone application for adults with type 1 diabetes: randomized controlled trial. J Med Internet Res 2013 Nov;15(11):e235 [FREE Full text] [doi: 10.2196/jmir.2588] [Medline: 24225149]
69. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12 [FREE Full text] [doi: 10.7326/M13-3005] [Medline: 25402403]
70. Nguyen B, Shrewsbury VA, O'Connor J, Steinbeck KS, Hill AJ, Shah S, et al. Two-year outcomes of an adjunctive telephone coaching and electronic contact intervention for adolescent weight-loss maintenance: the Loozit randomized controlled trial. Int J Obes (Lond) 2013 Mar;37(3):468-472. [doi: 10.1038/ijo.2012.74] [Medline: 22584456]
71. Quinn CC, Clough SS, Minor JM, Lender D, Okafor MC, Gruber-Baldini A. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. Diabetes Technol Ther 2008 Jun;10(3):160-168. [doi: 10.1089/dia.2008.0283] [Medline: 18473689]
72. Youl PH, Soyer HP, Baade PD, Marshall AL, Finch L, Janda M. Can skin cancer prevention and early detection be improved via mobile phone text messaging? A randomised, attention control trial. Prev Med 2015 Feb;71:50-56. [doi: 10.1016/j.ypmed.2014.12.009] [Medline: 25524612]
73. Zairina E, Abramson MJ, McDonald CF, Li J, Dharmasiri T, Stewart K, et al. Telehealth to improve asthma control in pregnancy: a randomized controlled trial. Respirology 2016 Jul;21(5):867-874. [doi: 10.1111/resp.12773] [Medline: 27037722]
74. Allman-Farinelli M, Partridge SR, McGeechan K, Balestracci K, Hebden L, Wong A, et al. A mobile health lifestyle program for prevention of weight gain in young adults (TXT2BFiT): nine-month outcomes of a randomized controlled trial. JMIR Mhealth Uhealth 2016 Jun 22;4(2):e78 [FREE Full text] [doi: 10.2196/mhealth.5768] [Medline: 27335237]
75. Hebden L, Cook A, van der Ploeg HP, King L, Bauman A, Allman-Farinelli M. A mobile health intervention for weight management among young adults: a pilot randomised controlled trial. J Hum Nutr Diet 2013 Aug 29;27(4):322-332. [doi: 10.1111/jhn.12155]
76. Ramachandran A, Snehalatha C, Ram J, Selvam S, Simon M, Nanditha A, et al. Effectiveness of mobile phone messaging in prevention of type 2 diabetes by lifestyle modification in men in India: a prospective, parallel-group, randomised controlled trial. Lancet Diabetes Endocrinol 2013 Nov;1(3):191-198. [doi: 10.1016/S2213-8587(13)70067-6] [Medline: 24622367]
77. Holmen H, Torbjørnsen A, Wahl AK, Jenum AK, Småstuen MC, Arsand E, et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-year results from the Norwegian randomized controlled trial RENEWING HEALTH. JMIR Mhealth Uhealth 2014 Dec;2(4):e57 [FREE Full text] [doi: 10.2196/mhealth.3882] [Medline: 25499872]
78. Torbjørnsen A, Jenum AK, Småstuen M, Årsand E, Holmen H, Wahl A, et al. A low-intensity mobile health intervention with and without health counseling for persons with type 2 diabetes, part 1: Baseline and short-term results from a randomized controlled trial in the Norwegian part of RENEWING HEALTH. JMIR mHealth uHealth 2014;2(4). [doi: 10.2196/mheatlh.3535] [Medline: 25499592]
79. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for IMproving FITness and increasing physical activity among young people: the AIMFIT pragmatic randomized controlled trial. J Med Internet Res 2015 Aug 27;17(8):e210 [FREE Full text] [doi: 10.2196/jmir.4568] [Medline: 26316499]
80. Arora S, Peters AL, Burner E, Lam CN, Menchine M. Trial to examine text message-based mHealth in emergency department patients with diabetes (TExT-MED): a randomized controlled trial. Ann Emerg Med 2014 Jun;63(6):745-54.e6. [doi: 10.1016/j.annemergmed.2013.10.012] [Medline: 24225332]
81. Proudfoot J, Clarke J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
82. Reid SC, Kauer SD, Hearps SJ, Crooke AH, Khor AS, Sanci LA, et al. A mobile phone application for the assessment and management of youth mental health problems in primary care: a randomised controlled trial. BMC Fam Pract 2011;12:131 [FREE Full text] [doi: 10.1186/1471-2296-12-131] [Medline: 22123031]
83. Constant D, de Tolly K, Harries J, Myer L. Mobile phone messages to provide support to women during the home phase of medical abortion in South Africa: a randomised controlled trial. Contraception 2014 Sep;90(3):226-233. [doi: 10.1016/j.contraception.2014.04.009] [Medline: 24850188]

84. Liguori S, Stacchini M, Ciofi D, Olivini N, Bisogni S, Festini F. Effectiveness of an app for reducing preoperative anxiety in children: a randomized clinical trial. JAMA Pediatr 2016 Aug 01;170(8):e160533. [doi: 10.1001/jamapediatrics.2016.0533] [Medline: 27294708]
85. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007 Apr;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
86. Karhula T, Vuorinen AL, Rääpysjärvi K, Pakanen M, Itkonen P, Tepponen M, et al. Telemonitoring and mobile phone-based health coaching among finnish diabetic and heart disease patients: randomized controlled trial. J Med Internet Res 2015 Jun 17;17(6):e153 [FREE Full text] [doi: 10.2196/jmir.4059] [Medline: 26084979]
87. Bell AM, Fonda SJ, Walker MS, Schmidt V, Vigersky RA. Mobile phone-based video messages for diabetes self-care support. J Diabetes Sci Technol 2012 Mar 01;6(2):310-319 [FREE Full text] [doi: 10.1177/193229681200600214] [Medline: 22538140]
88. Bobrow K, Farmer AJ, Springer D, Shanyinde M, Yu LM, Brennan T, et al. Mobile phone text messages to support treatment adherence in adults with high blood pressure (SMS-text adherence support [StAR]): a single-blind, randomized trial. Circulation 2016 Feb 09;133(6):592-600 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.017530] [Medline: 26769742]
89. Cingi C, Yorgancioglu A, Cingi CC, Oguzulgen K, Muluk NB, Ulusoy S, et al. The "physician on call patient engagement trial" (POPET): measuring the impact of a mobile patient engagement application on health outcomes and quality of life in allergic rhinitis and asthma patients. Int Forum Allergy Rhinol 2015 Jun;5(6):487-497. [doi: 10.1002/alr.21468] [Medline: 25856270]
90. DeVito Dabbs A, Song MK, Myers BA, Li R, Hawkins RP, Pilewski JM, et al. A randomized controlled trial of a mobile health intervention to promote self-management after lung transplantation. Am J Transplant 2016 Dec;16(7):2172-2180 [FREE Full text] [doi: 10.1111/ajt.13701] [Medline: 26729617]
91. Gamito P, Oliveira J, Lopes P, Brito R, Morais D, Silva D, et al. Executive functioning in alcoholics following an mHealth cognitive stimulation program: randomized controlled trial. J Med Internet Res 2014 Apr 17;16(4):e102 [FREE Full text] [doi: 10.2196/jmir.2923] [Medline: 24742381]
92. Hsu WC, Lau KH, Huang R, Ghiloni S, Le H, Gilroy S, et al. Utilization of a cloud-based diabetes management program for insulin initiation and titration enables collaborative decision making between healthcare providers and patients. Diabetes Technol Ther 2016 Feb;18(2):59-67 [FREE Full text] [doi: 10.1089/dia.2015.0160] [Medline: 26645932]
93. Levy N, Moynihan V, Nilo A, Singer K, Bernik LS, Etiebet MA, et al. The mobile insulin titration intervention (MITI) for insulin adjustment in an urban, low-income population: randomized controlled trial. J Med Internet Res 2015 Jul 17;17(7):e180 [FREE Full text] [doi: 10.2196/jmir.4716] [Medline: 26187303]
94. Markowitz JT, Cousineau T, Franko DL, Schultz AT, Trant M, Rodgers R, et al. Text messaging intervention for teens and young adults with diabetes. J Diabetes Sci Technol 2014 Sep;8(5):1029-1034 [FREE Full text] [doi: 10.1177/1932296814540130] [Medline: 25172879]
95. Petrella RJ, Stuckey MI, Shapiro S, Gill DP. Mobile health, exercise and metabolic risk: a randomized controlled trial. BMC Public Health 2014 Oct 18;14:1082 [FREE Full text] [doi: 10.1186/1471-2458-14-1082] [Medline: 25326074]
96. Newton KH, Wiltshire EJ, Elley CR. Pedometers and text messaging to increase physical activity: randomized controlled trial of adolescents with type 1 diabetes. Diabetes Care 2009 May;32(5):813-815 [FREE Full text] [doi: 10.2337/dc08-1974] [Medline: 19228863]
97. Yoon KH, Kim HS. A short message service by cellular phone in type 2 diabetic patients for 12 months. Diabetes Res Clin Pract 2008 Feb;79(2):256-261. [doi: 10.1016/j.diabres.2007.09.007] [Medline: 17988756]
98. Kim CS, Park SY, Kang JG, Lee SJ, Ihm SH, Choi MG, et al. Insulin dose titration system in diabetes patients using a short messaging service automatically produced by a knowledge matrix. Diabetes Technol Ther 2010 Aug;12(8):663-669. [doi: 10.1089/dia.2010.0031] [Medline: 20615108]
99. Liu CT, Yeh YT, Lee TI, Li YC. Observations on online services for diabetes management. Diabetes Care 2005 Oct 25;28(11):2807a-2808a. [doi: 10.2337/diacare.28.11.2807a] [Medline: 16249564]
100. Office of Disease Prevention Health Promotion. 2020 Topics and Objectives – Objectives A–Z URL: http://www.healthypeople.gov/2020/topicsobjectives2020/default [WebCite Cache ID 70Xn9FeUr]
101. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum; 1988.
102. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Br Med J 2003 Sep 6;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
103. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014 Feb;9(2):e88166 [FREE Full text] [doi: 10.1371/journal.pone.0088166] [Medline: 24505411]
104. Liang X, Wang Q, Yang X, Cao J, Chen J, Mo X, et al. Effect of mobile phone intervention for diabetes on glycaemic control: a meta-analysis. Diabet Med 2011 Apr;28(4):455-463. [doi: 10.1111/j.1464-5491.2010.03180.x] [Medline: 21392066]

105. Foster D, Linehan C, Kirman B, Lawson S, James G. Motivating physical activity at work: using persuasive social media for competitive step counting. 2010 Presented at: 14th International Academic MindTrek Conferencenvisioning Future Media Environments; 2010; Tampere, Finland. [doi: 10.1145/1930488.1930510]
106. Kang Y, Cappella J, Fishbein M. The attentional mechanism of message sensation value: interaction between message sensation value and argument quality on message effectiveness. Commun Monogr 2006 Dec;73(4):351-378. [doi: 10.1080/03637750601024164]
107. Dutta-Bergman MJ. Complementarity in consumption of news types across traditional and new media. J Broadcast Electron Media 2004 Mar;48(1):41-60. [doi: 10.1207/s15506878jobem4801_3]
108. Bandura A, Simon KM. The role of proximal intentions in self-regulation of refractory behavior. Cogn Ther Res 1977 Sep;1(3):177-193. [doi: 10.1007/bf01186792]
109. Bull S, Ezeanochie N. From Foucault to Freire through Facebook: toward an integrated theory of mHealth. Health Educ Behav 2016 Dec;43(4):399-411. [doi: 10.1177/1090198115605310] [Medline: 26384499]
110. Brewin CR. Theoretical foundations of cognitive-behavior therapy for anxiety and depression. Annu Rev Psychol 1996 Feb;47:33-57. [doi: 10.1146/annurev.psych.47.1.33] [Medline: 8624137]
111. Petty RE, Cacioppo JT. Attitudes and Persuasion: Classic and Contemporary Approaches. Boulder, Colorado: Westview Press; 1981.
112. Lee H, Chae D, Cho S, Kim J, Yoo R. Influence of a community-based stretching intervention on the health outcomes among Korean-Chinese female migrant workers in South Korea: A randomized prospective trial. Jpn J Nurs Sci 2016 Dec 15;14(4):277-287. [doi: 10.1111/jjns.12156] [Medline: 27981744]

Abbreviations

CBT: cognitive behavioral therapy
ES: effect size
FEM: fixed-effects model
HbA $_{1c}$: glycated hemoglobin
HBM: health belief model
MEM: mixed-effects model
mHealth: mobile health
RCT: randomized controlled trial
REM: random-effects model
RQ: research question
SCT: social cognitive theory
SMS: short message service
TPB: theory of planned behavior

Edited by G Eysenbach; submitted 03.07.18; peer-reviewed by B Fjeldsoe, I Mircheva; comments to author 06.10.18; revised version received 01.12.18; accepted 10.12.18; published 03.04.19.

Please cite as:
Yang Q, Van Stee SK
The Comparative Effectiveness of Mobile Phone Interventions in Improving Health Outcomes: Meta-Analytic Review
JMIR Mhealth Uhealth 2019;7(4):e11244
URL: https://mhealth.jmir.org/2019/4/e11244/
doi:10.2196/11244
PMID:30942695

Review

Effects of Mobile Health Including Wearable Activity Trackers to Increase Physical Activity Outcomes Among Healthy Children and Adolescents: Systematic Review

Birgit Böhm[1], MSc, PhD; Svenja D Karwiese[1], MSc; Harald Böhm[2], MSc, PhD; Renate Oberhoffer[1], MD

[1]Institute of Preventive Pediatrics, Technical University of Munich, Munich, Germany

[2]Orthopaedic Hospital for Children, Behandlungszentrum Aschau GmbH, Aschau im Chiemgau, Germany

Corresponding Author:
Birgit Böhm, MSc, PhD
Institute of Preventive Pediatrics
Technical University of Munich
Georg-Brauchle-Ring 60/62
Munich, 80992
Germany
Phone: 49 89 289 ext 24571
Fax: 49 89 289 24572
Email: birgit.boehm@tum.de

Abstract

Background: Children and adolescents do not meet the current recommendations on physical activity (PA), and as such, the health-related benefits of regular PA are not achieved. Nowadays, technology-based programs represent an appealing and promising option for children and adolescents to promote PA.

Objective: The aim of this review was to systematically evaluate the effects of mobile health (mHealth) and wearable activity trackers on PA-related outcomes in this target group.

Methods: Electronic databases such as the Cochrane Central Register of Controlled Trials, PubMed, Scopus, SPORTDiscus, and Web of Science were searched to retrieve English language articles published in peer-reviewed journals from January 2012 to June 2018. Those included were articles that contained descriptions of interventions designed to increase PA among children (aged 6 to 12 years) only, or adolescents (aged 13 to 18 years) only, or articles that include both populations, and also, articles that measured at least 1 PA-related cognitive, psychosocial, or behavioral outcome. The interventions had to be based on mHealth tools (mobile phones, smartphones, tablets, or mobile apps) or wearable activity trackers. Randomized controlled trials (RCTs) and non-RCTs, cohort studies, before-and-after studies, and cross-sectional studies were considered, but only controlled studies with a PA comparison between groups were assessed for methodological quality.

Results: In total, 857 articles were identified. Finally, 7 studies (5 with tools of mHealth and 2 with wearable activity trackers) met the inclusion criteria. All studies with tools of mHealth used an RCT design, and 3 were of high methodological quality. Intervention delivery ranged from 4 weeks to 12 months, whereby mainly smartphone apps were used as a tool. Intervention delivery in studies with wearable activity trackers covered a period from 22 sessions during school recess and 8 weeks. Trackers were used as an intervention and evaluation tool. No evidence was found for the effect of mHealth tools, respectively wearable activity trackers, on PA-related outcomes.

Conclusions: Given the small number of studies, poor compliance with accelerometers as a measuring instrument for PA, risk of bias, missing RCTs in relation to wearable activity trackers, and the heterogeneity of intervention programs, caution is warranted regarding the comparability of the studies and their effects. There is a clear need for future studies to develop PA interventions grounded on intervention mapping with a high methodological study design for specific target groups to achieve meaningful evidence.

(JMIR Mhealth Uhealth 2019;7(4):e8298) doi:10.2196/mhealth.8298

KEYWORDS

children; adolescent; mHealth; fitness tracker; physical activity; physical fitness

Introduction

Background

Physical inactivity is an increasing public health problem among children and adolescents worldwide [1-3]. Only a minority meets the global recommendations of the World Health Organization (WHO) on physical activity (PA) for health [1,4-7]. Thus, young people aged 5 to 17 years should perform at least, in total, 60 min of moderate- to vigorous-intensity PA (MVPA) daily, including vigorous activities on at least 3 days per week [8]. Physical inactivity increases the risk of noncommunicable diseases [9], already for primary school children [10], and represents the fourth-largest risk factor for mortality in the world [11].

Intervention strategies for health promotion, especially in children, must start early to grow healthy into adulthood because health-related attitudes and behavior patterns develop in early childhood, which are often maintained up to adolescence and adult age (12 to 14 years) [12]. Regular PA makes a significant contribution to the positive development of health in childhood and youth [4]. Numerous health benefits of regular PA are plentiful in this age [13], which persist into adulthood, such as positive effects on fitness, body fat, and blood pressure [14]. The dose-response relations observed in observational studies indicate that the more PA, the greater the health benefit. To achieve substantive health benefits, PA should be of at least a moderate intensity. Vigorous intensity activities may provide even greater benefit. Aerobic-based activities had the greatest health benefit, other than for bone health, in which case high-impact weight-bearing activities were required [15]. However, nowadays, an inactive everyday life is already ubiquitous in a young age [16,17].

The development of effective interventions to encourage active lifestyles among children and adolescents is one opportunity to address the lack of PA in this population group [18]. Use of technology-based interventions makes it more interesting [19]. Some preliminary data suggest that wearable activity trackers may have the potential to increase activity levels through self-monitoring and goal setting in the short term [14]. Mobile health (mHealth) and wearable activity trackers represent 2 of these innovations.

To date, a standardized definition of mHealth is not established. This is demonstrated in the fact that the terms *mHealth*, *electronic health (ehealth)* and *telehealth* are frequently used interchangeably [20]. The definition of mHealth as medical and public health practice supported by mobile devices, such as mobile phones, and other wireless devices is taken by WHO [21].

The use of smartphones among young people has increased in recent years. One-fifth of Germans aged 6 to 7 years use a smartphone. From the age of 12 years and above, the usage is over 80%. Tablets are most commonly used by those aged 12 to 13 years (43%) [22]. This trend can also be seen in the United States [23,24] and also in developing countries [25] where smartphones are used more than any other modern technology. Therefore, mobile devices and apps may be an effective strategy for promoting PA in this target group.

Furthermore, there is an increasing interest in commercial wearable devices that track health- and fitness-related activities and promote PA [26]. On the basis of their growing availability, popularity, and widespread adoption, they also offer a creative solution for children and adolescents to get moving in a playful way [23]. Currently, there are no data available on how many children and adolescents use wearable activity trackers. However, wearables are increasingly gaining importance as smart gadgets, and manufacturers are always looking for new apps to increase their sales [24]. Fitness trackers, such as *Garmin vívofit jr.* and *Jawbone UP*, are used to promote PA among this target group [27,28]. The former was developed for children and also involves the parents by setting tasks and defining rewards [27].

To date, several reviews have mainly or partially focused on PA outcomes of mHealth tools, in which mostly studies with adults were examined [29-32]. A large volume of PA research with tools of mHealth has primarily focused on weight control (eg, to prevent obesity) [33,34], on the treatment of diseases (eg, chronic diseases such as diabetes mellitus) [35,36], or on improving medication adherence [37,38].

The number of reviews on the issue of mHealth on PA is still low, compared with reviews using wearable activity trackers to increase PA. To date, wearable activity trackers have been primarily studied to examine their ability, validity, and reliability to estimate PA [39,40]. There are also some feasibility studies of these devices for children [41,42]. Frequently, they are used in an intervention as an evaluation tool to measure PA levels objectively [39,43]. However, little is known about the effectiveness of these devices as a tool for promoting PA outcomes, whether as a single strategy or in combination with others. Until now, healthy children and adolescents seem to play a tangential role in this area of research. The review by Lewis et al [44] reported some initial evidence that wearable activity trackers can increase PA, but only studies with adults have been taken into account. However, children and adolescents have a high affinity to new technologies and use them in their daily lives.

Objectives

To date, no review evaluated tools of mHealth as well as wearable activity trackers that promote PA and increased PA behavior in healthy children and adolescents. Therefore, there is a research need to evaluate systematically whether tools of mHealth or wearable activity trackers are appropriate and effective in promoting and changing PA in this target group. The results are important for informing and supporting future PA interventions in young people. Moreover, it has the potential to contribute to the development of public health guidelines relating to the role of these tools in PA and health promotion.

Therefore, the aim of this systematic review was to examine the effectiveness of interventions that use tools of mHealth, respectively wearable activity trackers, to promote and change PA among children and/or adolescents. As the health status may

confound the effectiveness of interventions this review considered only healthy children and adolescents.

Methods

Treatment Objectives

The definitions and inclusion criteria used are described below.

Mobile Health Tools

Mobile phones, smartphones, tablets, and apps on these devices were considered as tools in the field of mHealth, which aimed to promote PA to increase PA levels among healthy children and adolescents. With regard to the widespread usage, the field of mHealth is already strongly focused on these devices, especially on smartphones, to create promising interventions for the youth [45-50].

Wearable Activity Trackers

Wearable activity trackers were defined as an electronic device with the following features: designed to be worn on the user's body; uses accelerometers, altimeters, or other sensors to track the wearer's movements or biometric data or both; and can provide feedback via the monitor display or through a partnering app to elicit continual self-monitoring of activity behavior [44,51]. This definition eliminates pedometers and accelerometers that do not supply automated feedback to the wearer [44]. Systems with feedback were included in the definition, as self-monitoring resulted in significantly more activity compared with a no-feedback condition [52-54].

Data Source and Search Strategy

A systematic literature search was conducted to find out relevant articles in 5 electronic databases: the Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, Scopus, SPORTDiscus, and Web of Science. The search was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. For most entries, text words and synonyms were used, marked with (tw), plus Medical Subject Headings (MeSH), marked with (MeSH), for major records with relevant keywords without any limitations. The MeSH search was only conducted in CENTRAL and PubMed databases (Multimedia Appendix 1). Accordingly, text words were oriented on entry terms of MeSH headings to cover a comparable search spectrum in the other data banks this was especially necessary in the German data bank, SPORTDiscus, as these conceptually, for example, put another focus (Multimedia Appendix 2).

Search strategies for the various databases contained search strings in 4 main areas: population, treatment method, treatment objective, and outcome variable. Terms for mHealth and wearable activity trackers were adapted from previous reviews [14,38] and entry terms from the MeSH heading *Fitness Trackers*, which were introduced in PubMed in 2017.

The search was carried out based on article title, abstract, and keywords in all 5 databases. In the case of PubMed, the terms were entered into the search box using PubMed's search field tags (tw) for the text words and their synonyms and (mh) to search the MeSH headings. In all databases, each individual term was scanned first. Then, the Boolean operator *OR* was used within the individual areas and, subsequently, the operator *AND* to combine the 4 search fields. Finally, the available limits in the various databases were selected.

The reference list of thematically related review articles was also searched for potentially useful sources. Checking the bibliographies of identified studies is a generally used approach to identify additional relevant studies for potential inclusion in systematic reviews [55].

Selection Criteria of Studies

Inclusion Criteria

1. Published in peer-reviewed journals in English. Studies in press were included if they had a unique digital object identifier.
2. Published from the beginning of January 2012 to the end of June 2018.
3. Focused on children and/or adolescents.
4. Included healthy participants (including underweight, overweight, and obese without any reported dysfunction).
5. Specifically examined the use of at least 1 mHealth tool or of a wearable device within an intervention to promote PA (even if it was only 1 component of the whole intervention).
6. Measured at least 1 PA-related variable as the outcome (in this, connection was not defined as a restriction regarding the types of PA-related outcomes, which could be cognitive [ie, PA knowledge and PA self-efficacy], psychosocial [ie, PA intention, social support to PA, and stage of change], or behavioral [ie, energy expenditure, step counts, or observed or self-reported PA level], or physical fitness).

Overall, randomized controlled trials (RCTs) and non-RCTs, cohort studies, before-and-after studies, and cross-sectional studies were considered. If the study design was not clearly stated but contained in their description characteristics of one of the included study designs, it was included. In addition, studies based on an experimental design were checked according to this criterion and upon fulfillment were included. If there were multiple publications from mHealth or wearable activity tracker interventions, only the study with the PA outcome(s) or the most recent publication with PA outcome(s) was included.

We oriented the relatively strict search years at the time when consumer wearable activity trackers with proofed validity and reliability entered the market following the review of Everson et al [39]. In addition, the results of Ridgers et al [14] were taken into account, as well as the review on mobile phone interventions published in 2017 [56].

Exclusion Criteria

1. Conference proceedings, book chapters, dissertations, pilot studies, and systematic reviews.
2. Studies where the main mHealth component was not mobile, eg, Web- or email-based), that did not evaluate at least 1 mobile aspect of assessment or intervention delivery, or where there were no indications of mobile platform compatibility (eg, the app used on a desktop computer does not run equally on a tablet or smartphone).

3. Studies that used wearable activity trackers only to evaluate an intervention.
4. Articles examining the validity or feasibility of mHealth tools or wearable activity trackers were excluded if they did not evaluate these as technologies to measure participants' PA-related outcome(s).
5. Studies where participants had additional reported dysfunctions.
6. Studies focusing on weight control or loss without any PA measurement.

Study Selection and Data Extraction

Following a standard protocol, 2 authors (SS and BB) independently screened studies for eligibility based on the title, abstract, and full text. Uncertainty was discussed involving a third author (RO), and any disagreement was resolved by consensus.

All search results were exported into EndNote X7.7.1 (Thomson Reuters). Information about each paper was extracted by BB and SS independently for quality assurance. Screening of all entries took place in 4 steps: First, duplicate references were removed. Then, all titles were screened, and additional entries, which did not match the MeSH terms and text words that lead to a different content, were removed. Entries were left in the database if the context was not fully clear from the title. After that, abstracts of the remaining articles were screened. If there was any doubt in the information in the abstract, the full article was retrieved to ensure that no relevant entries were lost. In the end, full text articles were retrieved for further assessment, if the eligibility criteria had been fulfilled or suggested that the article was a potential study for this review. All the remaining entries were reviewed for final inclusion. Figure 1 contains the excluded criteria for decision making of the selection of potential useful articles used in the examination of the abstracts and full texts.

The data from the selected intervention studies were extracted with regard to the following information: (first) author; year of publication; study design; country in which the intervention was carried out; place of recruitment; number of study participants and their characteristics (age, gender, and body mass index [BMI]); types of tools used (in the field of mHealth or wearable activity trackers); intervention description and duration; time of measuring and measuring instruments of PA-related outcome(s); and key findings on PA or physical fitness. In the case of a controlled trial (CT) with comparison between the intervention group (IG) and control group (CG), further data were also extracted: task or program of CG and differences in PA levels between the IG and CG.

Furthermore, studies were distinguished by their intervention field: (1) tools of mHealth or (2) wearable activity trackers. A PRISMA flow diagram presents the summary of the study selection process (Figure 2). Databases used were Cochrane Central Register of Controlled Trials (n=136), PubMed (n=88), Scopus (n=216), SPORTDiscus (n=199), and Web of Science (n=211). At abstract screening, records could be excluded for at least 1 of these reasons (the first exclusion criterion was always counted, even if several apply): (1) no study with use of at least one mHealth tool or a wearable activity tracker to promote PA (n=84); (2) no target age groups (n=46); (3) no healthy participants (n=1); (4) no original study (n=15); (5) no measurement of at least one PA variable (n=2); or (6) not published in a peer-reviewed journal (n=1). At the full-text article screening, studies were excluded for the following reasons: (1) no study with use of at least one mHealth tool or a wearable activity tracker to promote PA (n=17); (2) without PA measurement (n=5); (3) duplicate publication (n=1); or (4) not selected study design (n=10). Additionally, 1 full text could not be obtained, and for 5 studies, it was only revealed in the full text that the participants did not meet the age criteria groups (n=3), were not healthy (n=1), or it was a previous publication (n=1).

Figure 1. Flowchart of the decision-making process for levels of evidence, based on study design and methodological quality. CT: controlled trial; RCT: randomized controlled trial.

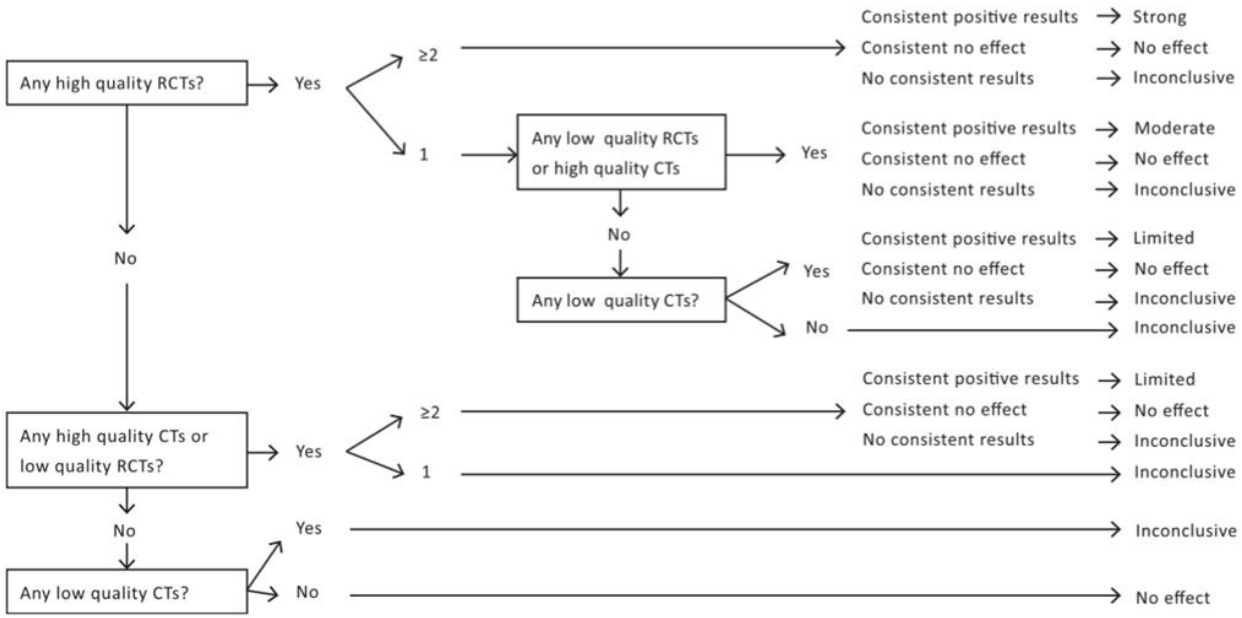
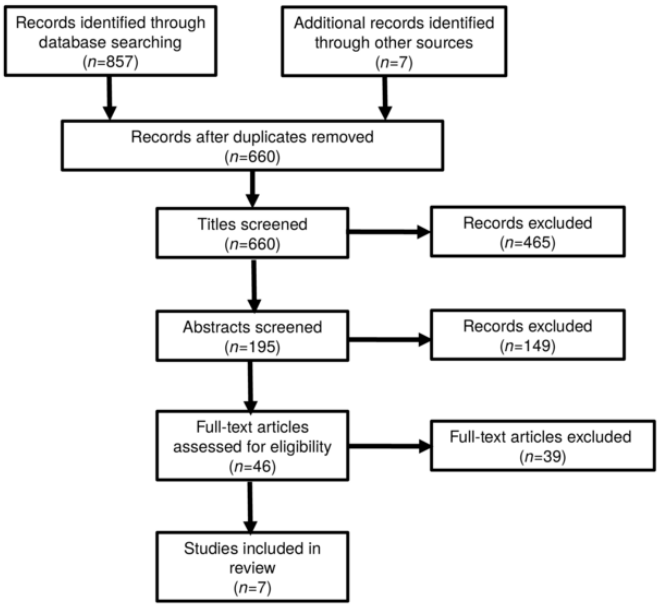


Figure 2. Process of identification and selection of included studies. Databases used were Cochrane Central Register of Controlled Trials (n=136), PubMed (n=88), Scopus (n=216), SPORTDiscus (n=199), and Web of Science (n=211).



Publication Bias

A sufficiently large number of studies will be included in this review (including some with high subscriber numbers), funnel-plot analysis will be implemented. This is used as a test for publication bias and similar systematic errors [57].

Criteria of Methodological Quality

Only controlled studies with PA comparison among groups were examined with regard to their methodological quality. It is essential to take into account a possible biasing influence on estimates of intervention effectiveness [58]. By means of the tool by the *Cochrane Handbook for Systematic Reviews of Interventions*, systematic reviewers are provided help to select the adequate criteria for evaluating the possible bias in a specific field of study. These standards were primarily developed for medical and health science studies [59]. On the basis of these recommendations [59] as well as previous systematic reviews that can be applied to this research field [60-62], a total of 9 criteria must be satisfied to maintain high methodological quality (Multimedia Appendix 3). One point was given to a study if a criterion was met, whereas no points were given when a criterion was not fulfilled or when it was not (sufficiently) described. A methodological quality score (ranging from 0 to 9) was calculated by accumulating all positive items. Studies scoring 0 to 2 points were of low methodological quality, studies with 3 to 5 points were of moderate quality, and studies scoring 6 or above were of high methodological quality. If the study was a non-RCT, the score had to be at least 5 for high methodological quality (owing to the fact that 1 item was regarded to the randomization procedure).

With the aid of the selected quality characteristics, the risk of the 4 most important forms of bias could be examined, which could influence the (internal) validity of a study: selection bias, performance bias, measurement bias, and attrition bias [57].

Strength of Evidence

The strength of evidence was evaluated based on a previously used evidence synthesis method [60,63,64]. Therefore, the effects of interventions with the use of tools of mHealth or wearable activity trackers on PA were rated using an evidence rating system adopted from a study by Liang et al (Figure 2) [62]. As a result, the following 5 levels were defined based on the study design and methodological quality: (1) strong; (2) moderate; (3) limited; (4) inconclusive; and (5) no effect. The studies were stratified based on their intervention tool: in the field of mHealth or wearable activity trackers. Following a review by van Sluijs et al, the overall results were considered as consistent if at least two-thirds of the relevant studies had significant results in the same direction [60].

Results

Included Studies

In total, 864 records were found through a systematic search of 5 databases and other sources that were thematically related reviews and retrieved studies. Finally, 7 trials were identified matching the inclusion criteria (Figure 1), of which 5 used a tool of mHealth [65-69] and the other 2 studies made use of wearable activity trackers to promote PA among children and/or adolescents [70,71]. All 7 studies are described in detail, distinguished by their intervention field: tools of mHealth (Multimedia Appendix 4) and wearable activity trackers (Multimedia Appendix 5). For the description of the intervention characteristics, the study protocols of 2 included studies were additionally consulted [72,73].

The designs of the intervention programs were guided from theoretical frameworks and behavior change techniques (BCTs). Some were referred only to one theoretical model, but most of the studies integrated more than one theoretical model. As basic theoretical model, the *Social Cognitive Theory* by Bandura was most frequently used [65,68-70]. In total, 3 studies integrated

additional BCTs [65,66,68] such as feedback on behavior, self-monitoring behavior, goal setting, and strategies to overcome barriers. Except for 2 studies [66,71], no information was given.

The variety of and the inconsistency in the methods used for data assessment across the studies make it difficult to compare the effects on PA. Accelerometers were used predominantly for objective PA measurement; however, the inclusion criteria varied for the data to be included in the evaluation. Participants' data were generally included in the analyses if valid data for at least 3 days existed. Only the study by Dewar et al requested an additional weekend day [65]. However, wearing time of the monitors varied. In total, 2 studies determined a wearing time of at least 600 min per day [65,66], whereas others demanded ≥480 min per day [67,68]. To evaluate the PA self-efficacy, for example, Direito et al used the *Physical Activity Self-Efficacy Scale* questionnaire [66]; however, Dewar et al designed their own questionnaire that inquired self-efficacy as one item [65].

Publication Bias

This review included only a small number of studies (n=7), so the presentation of the funnel plot was omitted. Therefore, the existence of publication bias is to be assessed as unclear.

Methodological Quality

Multimedia Appendix 6 reports the methodological quality for the 5 selected intervention studies with group comparison [65-69]: 4 were designed as RCTs [65,66,68,69] and 1 was a randomized controlled cross-over trial [67]. In the case of the study by Garde et al [67], it was necessary to consult the study protocol to clarify whether the item (9) was met [74]. In total, 60% of the studies (n=3 studies) were of high methodological quality [65,66,68] and 2 studies of moderate quality (40%) [67,69]. Lubans et al reached the highest score [68]. Their study met all the criteria except for the blinding criterion, which, however, was also not fulfilled by any other study. However, the application of blinding strategies in mHealth is often impossible, impractical, or infeasible, thus making blinding more difficult [75].

To obtain a more differentiated insight into the intervention studies, each methodological criterion was examined on its own (Multimedia Appendix 7). All studies adequately carried out pretest analyses, met the criterion of the timing of measurement, and used valid measurement instruments. In most cases, accelerometers were applied to measure PA levels objectively [65-68], whereby the *ActiGraph* accelerometer was most commonly used [65,66,68]. In 3 out of 7 studies, the assessors of the pretest were not blinded [65-67], and the other 2 studies had no sufficient information about their blinding process [68,69]. The randomization criterion was met by 5 out of 7 studies. Although Zach et al divided their participants randomly, they did not explicitly describe the method used [69], and Garde et al included only 42 participants in their randomization process [67]. A clear randomization process at the level of experimental planning is necessary to ensure that potential confounders are evenly distributed among the comparison groups. Through randomization, a relation between potential confounders and the exposure can be excluded [76]. The criterion of dropout rate was met by 2 out of 7 studies, and the criterion of systematic dropout was met by 3 out of 7 studies. Only the study by Lubans et al [68] included a follow-up measurement realized at a minimum of 3 months after completion of the intervention. Finally, the criterion of the sample size was met by 4 out of 7 studies. Thereby, 2 studies examined samples larger than 250 participants [65,68] and the other 2 had smaller sample sizes but carried out a power calculation [66,69].

Strength of Evidence

In total, 5 studies used a tool of mHealth to promote PA, including 3 high-quality RCTs [65,66,68] and 2 RCTs with moderate methodological quality [67,69] with one of those designed as a cross-over study [67]. The 3 high-quality RCTs with an objectively measured or a self-reported PA level, that is, MVPA, consistently reported no statistically significant effects on their PA outcome. This means that there was evidence of no effect on PA-related outcomes among interventions with tools of mHealth.

Both the intervention studies with wearable activity trackers were designed as before-and-after trials [70,71]. According to the flowchart (Figure 2), only a low-quality CT was available. Therefore, no effect of wearable activity trackers on PA was identified.

To summarize, no evidence of an effect among all interventions with tools of mHealth or wearable activity trackers, or both to promote PA among healthy children and/or adolescents was identified.

Intervention Studies: Tools of Mobile Health

Text Message

Only the study by Dewar et al used short message service (SMS) text messaging as one element in their intervention program [65]. In total, 8 different components were integrated in their study [72]. The overall aim was promoting PA and healthy eating and preventing obesity in female adolescents. The participants (n=357 girls; mean age 13.2 [SD 0.5]) came from economically disadvantaged secondary schools (n=12) and were disengaged in physical education (PE) and/or not currently participating in organized team sports or individual sports. With 12 months, this study had the longest intervention period. The primary outcome (BMI) was reported in a previous publication [77]. In this study, the secondary outcomes (ie, objectively measured PA with *ActiGraph* accelerometers) were examined. According to the study protocol, the text messages were sent to the participants each morning during the 7-day monitoring period to remind wear and improve compliance. However, the study did not report this step.

After 12 months, PA data from 246 girls could be evaluated. There were no significant group-by-time effects for moderate PA (MPA) and vigorous PA (VPA), as well as MVPA. Changes for most of the social-cognitive variables were found in the IG. However, there were no statistically significant effects. Follow-up data were not published. Only baseline results and measures after 12 months were analyzed.

Smartphone App

Direito et al investigated the effects of 2 commercial smartphone apps (*Zombies, Run!* as an immersive app and *Get Running* as a nonimmersive app) [66]. The participants (*n*=51; mean age 15.7 [SD 1.2]) owned an iPod touch or smartphone running at least iOS 6.0 or Android 2.2, respectively. In relation to PA, they were able to perform PAs but were not achieving the PA recommendations of their age group. The apps served to improve PA as well as the ability to run 5 kilometers. Participants received gift cards to a local shopping center for each visit to complete study measures independent of their usage of the app. PA levels were evaluated by self-reporting and measuring using the *ActiGraph* accelerometer. In addition to cardiorespiratory fitness and PA outcomes, the features of the app design were also evaluated with regard to their acceptability and usability.

After 8 weeks, no significant increases for self-reported PA and PA self-efficacy were recorded. In addition, there were no statistically significant effects on physical fitness. The average daily time spent in MVPA had a decrease toward the baseline in the *Zombies, Run!* group and the CG (posttest 33.04 signified 55% to daily recommendations [8] and 30.54 signified 51% to daily recommendations) [8]. The *Get Running* group reported an increase from 21.29 to 23.34 (signified 38.9% to daily recommendations) [8]. This group was overall the weakest in terms of PA-level measurements.

In Canada, Garde et al examined the efficacy of the mobile exergame *MobileKids Monster Manor* (MKMM) in a school-based environment (*n*=42; mean age 11.3 [SD 1.2]) [67]. The most innovative and special feature of MKMM was that the player had to earn in-game playtime by performing PA.

After 4 weeks, they could confirm their hypothesis that children with a higher BMI *z*-score had greater benefit while playing the game. The increase in PA was significantly greater relative to their counterparts with a lower BMI *z*-score ($P<.05$), which was also observed in Garde et al previous community-based study [74]. Furthermore, there was a significant PA difference between the game intervention and control weeks, showing more steps ($P<.001$) and active minutes ($P<.001$) per day during the intervention. PA was recorded using the *Tractivity* activity tracker, which considers an active minute only if it contains at least 20 steps within a window of 7 active min. In addition to PA, the experiences playing the game were evaluated using a survey. In total, 90% of the children thought that MKMM was very effective at promoting PA.

Lubans et al were the only group that evaluated an 18-month follow-up [68]. This study was designed to be culturally appropriate and incorporated mHealth technology with the goal of preventing obesity among adolescents [73]. The participants were disadvantaged boys failing to meet the international PA guidelines, why they were *at-risk* of obesity, whereby the weight status was not an inclusion criterion (*n*=361; mean age 12.7 [SD 0.5]). The study included 7 components (ie, a smartphone app) and involved teachers, parents, and students. After the study endpoint, participants still had access to the smartphone app.

After 18 months from baseline, significant intervention effects for PA were not reported ($P>.05$). The differences at baseline between completers and dropouts for the outcomes at 18 months were not meaningful.

Zach et al included only female high school students in their program, which was carried out during PE lessons, and measured psychological and physiological effects (*n*=154; ages 16 to 18 years) [69]. The reason for the focus on female subjects was that PE in grades 7 to 12 is single sex and not coeducational, and the lack of motivation to participate in PA in leisure time is also more prominent in Israel among female adolescents. In 1 IG, they used the smartphone app, *WhatsApp*, so that the participants could write a short personal report to the class *WhatsApp* group. This report was also shared with the teacher. Overall, they chose the internet and the app as 2 different kinds of technical methods to evaluate which of the two would serve as a better means for also increasing self-efficacy for independent PA.

At the end of the 12-week intervention, the results of the participants' perception of self-efficacy for independent training were inconsistent. Predominantly, no prepost differences or interactions were observed. Zach et al also hypothesized that the best IG would have the greatest significant benefit in physical fitness. But significant differences between groups showed up in all physical fitness measures ($P<.05$).

Intervention Studies: Wearable Activity Trackers

Bronikowski et al used the *Garmin vívofit* activity tracker (model not reported) as an evaluation and intervention tool to examine the effectiveness of different target strategies on PA behaviors among children and adolescents (*n*=193; aged 11 to 17 years) [70]. The *Goal set* group had a daily goal of 10,000 steps for adolescents and 12,000 steps for young adolescents and children. The *Do your best* group did as many steps they could and wanted to do daily. The activity tracker was worn during the whole intervention. All participants could see the number of steps and had an internet account created on the *Garmin Connect* program to follow their progress and weekly trends. The level of PA was determined by means of a PA screening measure, and an MVPA index was thus calculated. The average number of steps was only identified in the posttest. In addition, classmate and teacher support in MVPA during PE lessons was evaluated.

At the end of the 8-week intervention, all adolescents of the 2 groups achieved 10,000 steps, whereas all young adolescent girls could not reach their recommendation of 12,000 steps. In comparison with this, in the group of children, only the girls from the *Do your best* group met the criterion. Generally, in most cases, the daily average was higher in the *Do your best* group. However, the MVPA index decreased in this group. The external support of classmates and teachers was not taken into account in this review as sufficient information about the exact actions in which this support should promote PA was not given.

In addition, Hayes and Van Camp applied an activity tracker to promote PA [71]. They used *Fitbit* (model also not reported) as a tool to increase the PA levels of girls in the third grade in an elementary school (*n*=6; aged 8 years) during unstructured school recess as well as their evaluation tool. During the baseline process, the participants did not receive feedback regarding their number of steps. The criterion for moving out of baseline

was stability or the absence of an upward trend. After baseline data were collected from 7 recess periods, girls were provided with step goals for 7 further recess periods. In addition, they were encouraged to self-monitor their steps against goals. The incremental increase was based on the baseline data. Subsequently, data were collected for a further 7 periods, in which there were no step goals provided. For the final intervention session, 3 goals were given (20%, 30%, and 40% increase). Moreover, a tangible reward (eg, small toys) was provided based on the goal(s) achieved. Results revealed an increase in steps by 47% from baseline, which contributed 18% to the daily step recommendations [78]. The percentages of time spent in MVPA increased from 4% (range 2% to 6%) to 25% (range 10% to 41%), which equate to 5 min of MVPA during recess or a contribution of 8% to the WHO's daily recommendations of at least 60 min MVPA [8]. Without the use of the *Fitbit* activity tracker to self-monitor recess activity, the number of steps and MVPA decreased to initial baseline levels.

Discussion

Principal Findings

In total, 7 intervention studies were identified that reported the use of mHealth tools or wearable activity trackers in healthy children and adolescents with PA-related outcomes. Most of them (5/7, 71%) included mHealth technologies such as mobile apps, games, and SMS text messaging in efforts to promote PA among children and/or adolescents [65-69]. None of the studies had the same mean age of the trial population, which ranged from 11 to 18 years. The study results of the 3 high methodological RCTs consistently reported no statistically significant effects on their PA-related outcomes. However, there was evidence of no effect in relation to the applied scheme. It should be noted that in relation to smartphone apps, 81% were interested in trying various PA-promoting apps in the future [66] and 92% of the children in the study by Garde et al enjoyed the requirement of being active [67]. Games in the form of mobile apps seem to be an attractive tool to promote PA among the youth. However, game design should be appropriate for specific age groups.

Only 2 studies using wearable activity trackers met the criteria to be included. In these studies, the activity trackers were used as the intervention as well as the evaluation tool to measure MVPA [70,71]. Neither of them used an RCT but used a before-and-after study design. On the basis of the applied scheme by Liang et al [62], there was no effect of these devices on PA.

Intervention Approaches

Most of the included studies were set in a school environment. The importance of schools as a setting in PA promotion has already been highlighted [79]. Children and adolescents spend much of their time there, so this environment offers itself to implement such interventions aimed at promoting PA [80].

The studies had different restrictions on the PA of their included participants. Examining the PA status (time spent in VPA, MPA, and MVPA) to select suitable participants is useful. For example, Bronikowski et al reported that their subjects had a reasonably high level of MVPA before their intervention, which is why their effects could have been weakened [70]. Therefore, the approach of some studies focusing on healthy children and adolescents who did not meet their PA recommendations by WHO appears appropriate [66,68]. Moreover, PA intervention programs focusing on children and/or adolescents with a greater BMI *z*-score are necessary. Garde et al could affirm their hypothesis that they have a greater benefit on PA [67]. In total, 2 studies recruited their participants from socioeconomically disadvantaged schools and focused on a specific gender [65,68]. In addition, Hayes and Van Camp and Zach et al included only female participants [69,71]. It is known that especially girls with a low socioeconomic status showed a lack of PA [81]. Therefore, sex-specific interventions should be considered in future research.

Intervention Studies with Tools of Mobile Health and Their Effects

First, all studies with tools of mHealth were designed as RCTs. Second, except for 1 study, where no sufficient information was given [66], the intervention programs were grounded on theoretical frameworks. Third, most of the studies also integrated BCTs [65,66,68]. BCTs are defined as an observable, replicable, and irreducible component of an intervention program designed to change or redirect causal processes that regulate behavior (eg, feedback and self-monitoring) [82]. If an intervention's description of a study named their BCTs, the effectiveness of the intervention can also be associated with these strategies [82]. For example, the apps used by Direito et al included self-regulatory BCTs (ie, prompt specific goal setting, prompt self-monitoring, and provision of feedback on performance) [83]. However, no significant effects were reported.

The limited significant effects observed in intervention studies with tools of mHealth included in this review may be the result of the intervention design and the PA evaluation method that has been elaborated below:

First, the intervention duration of most studies lasted from a minimum of 4 weeks to a maximum of 12 weeks. Nguyen et al reported in their review published in 2016 that behavioral interventions with a duration of ≥6 months had greater success in changing PA levels [84]. Thus, it is possible that a short period of time could be insufficient to change behavior. However, Dewar et al delivered their intervention over 12 months, but did not find significant effects [65]. However, their main aim was to reduce the BMI; thus, it is possible that the components of their program were not sufficient for changing PA positively. There is a need for PA studies with comparable long intervention durations to find out whether significant effects are found.

Second, only 2 studies involved a large number of participants (≥250) [65,68], so most of the results may not be representative for the total population.

Third, researchers did not supervise the intervention programs, with the resulting limitation not guaranteeing the extent to which the intervention measures were implemented [85]. For example, Dewar et al had information that 91% of the IG accessed text

messages, but it was unknown if these were read by the participants [65]. Lubans et al did not have objective usage data to determine participants' continual engagement with the smartphone app [68]. In addition, Direito et al did not closely monitor use of the apps during their intervention [66]. Indeed, they wanted to show the usefulness of the apps in real life [66]; however, important data are missing for evaluation. These examples show how essential it is to verify the intervention components to draw conclusions about their effectiveness.

Fourth, a general problem was the lack of compliance. All studies with the use of an accelerometer reported failing compliance [65-68]. On the basis of insufficient wear time, the inclusion criteria to evaluate PA were not met by the majority of subjects. Only 24.6% met the criteria by Dewar et al at posttest [65]; Lubans et al included only 32% of their participants [68]; and Garde et al received valid data from 28 out of a total of 42 participants (66.7%) [67]. In comparison with this, Direito et al had a low failure of data (loss of 4%) at posttest [66]. However, the reduced number of cases in most of the studies resulted in underpowered analyses, so their findings are not meaningful enough.

Fifth, 3 studies used questionnaires for PA-related outcome measurement [65,66,69]. However, self-reported measures are vulnerable for recall as well as report bias, which is also susceptible for social desirability bias.

Finally, 80% of the included studies used accelerometers to measure PA levels objectively [65-68]. These devices have adequate reliability for PA surveillance. However, there still exist several issues associated with their validity [86]. Accelerometers lack the sensitivity to recognize and record nonambulatory movements, so not all forms of movement are detected [65]. These restrictions could have led to the reported lack of intervention effects.

In conclusion, future research projects in this field are encouraged to develop intervention programs with a longer period of time (≥6 months), including a sufficiently large number of participants (≥250) to receive meaningful results about their efficacy. Moreover, bias should be avoided. For this reason, self-reported measurements should be bypassed. Furthermore, measuring instruments should be checked in advance for their specificity. Owing to the preferable use of accelerometers to determine PA, despite their poor compliance among the youth, further investigations to improve this are indicated.

Future intervention programs should use the intervention mapping (IM) by Kok et al to develop theory- and evidence-based health promotion interventions [87] to be able to assess the impact these factors (eg, election of BCTs, intervention components, and implementation of the program) may have had on the overall findings.

Intervention Studies With Wearable Activity Trackers and Their Effects

Although 1 study was able to observe a clear percentage increase in the number of steps during their intervention phase (increase of 47%), there was a lack of statistical evaluation. Another limitation was that some data for MVPA calculation were lost because of syncing failures, including the last intervention session. In addition, the increase in MVPA during brief 20-min sessions was not consistent and may not currently suffice for clinical significance as well as for a transfer into practical recommendations [71]. Overall, it can be concluded that there were limited intervention effects in the studies with wearable activity trackers as their intervention tool. This result may be attributable to several factors as follows:

First, the intervention duration amounted to a maximum of 8 weeks [70] which might be too short to change behavior [84]. There exists a clear need for studies using longer intervention periods to obtain more meaningful results regarding the effectiveness of these devices in the youth. Second, Hayes and Van Camp had recruited a small number of children; therefore, the study has been underpowered to detect a significant change in PA [71]. Third, in this study, it was not reported if the intervention was grounded on behavioral theories [71] that are essential for intervention effectiveness [88]. Fourth, the selection of the participants should be characterized with regard to their PA level in advance. Bronikowski et al assumed that the missing significant intervention effect might have resulted from MVPA of the participants already present before the beginning of the intervention [70]. One possibility would be to include children and/or adolescents who are healthy but do not meet the recommendations of PA. Fifth, the studies used the data of their wearable activity tracker to measure MVPA. However, it has been noted that, to date, these devices are not validated for assessing PA-related outcomes in the youth [39]. As long as this restriction still exists, the intervention effects should be recorded using validated objective monitors such as accelerometers. It is further important to pay attention to compliance. Finally, the school setting in which the interventions were implemented could influence the study results. Schools are the ideal place to carry out PA promotion interventions. Moreover, there already is sufficient evidence for the increase of PA and fitness in the youth through school-based interventions [89]. However, a process evaluation is essential to control the intervention implementation as well as to examine the range of the program [85]. Otherwise, it is possible that the intervention's effects are influenced by deficits in execution. For example, Bronikowski et al did not monitor their intervention. Therefore, they argued that it is possible that the participants received homework to reach their goals, which could have influenced the study results as well as the fact that all other daily and weekly activities (eg, PE lessons) continued but were not accurately recorded [70].

These notes could be considered as limitations of both the included studies with the use of wearable activity trackers. If intervention studies use these devices as tools with a focus to facilitate behavior change by motivating and supporting, the limitations may be less of an issue. However, if the studies want to evaluate their outcomes by means of these devices, validity and reliability should be established before such use.

Finally, based on the 2 studies, a clear need for RCTs with longer intervention duration (≥6 months) and sufficient participants is also indicated here. In addition, a follow-up is essential to evaluate the sustainability of using wearable activity trackers. Future research should also be grounded on proved theoretical frameworks to identify the effectiveness of wearable

activity trackers for promoting and increasing PA among children and adolescents. In addition, here IM is a helpful planning program framework for development, which integrates, as already mentioned, an evaluation of the intervention program [87].

Self-Monitoring Using Wearable Activity Trackers

The studies also used their wearable activity tracker to self-monitor PA in combination with various intervention approaches. In both studies, the approach of *goal setting* was used [70,71], which is an effective BCT [82]. This technique was often used in PA interventions, because setting specific difficult goals is suitable to enhance PA levels [90]. If the goals are not difficult enough to reach, as in the study by Bronikowski et al, the effect of changing PA levels could not be significant [70]. However, who set the goals varied in both studies. In 1 study, 1 group had fixed aims of daily steps and the other group did as many steps as they could and wanted. The number of steps was apparent for both groups. Regular support or receipt of tailored advice was not reported [70]. The missing masking of the wearable activity tracker could have represented a motivating factor for the group without fixed aims. The blinding of the monitors could be a benefit for the participants to exploit all their abilities and not just achieve their potential goals [70]. In addition, it has been noted that research has shown that allocated goals are equally effective as self-set goals [91]. In contrast, in the intervention program with elementary school children, Hayes and Van Camp set the goals based on the baseline values monitored and then provided rewards in relation to reaching the goals [71]. Rewarding participants for their achievement of their goals is associated with significantly higher PA effect sizes (P<.05) [90].

For future research, it seems to be of great importance to create an incentive for increasing PA in terms of high but achievable goals or also by praising or rewarding participants for achieving their goals or their attempts. Other proved BCTs to increase PA can also be taken into account. Furthermore, the comment by Ridgers et al is appropriate. They demanded to evaluate how children and adolescents engage with wearable activity trackers. Therefore, it can be determined whether the frequency of self-monitoring is mediated by the activity goal [14]. These demands on future research results will show if the approaches will be more effective against the background of this review and provide evidence as to how these technical devices can be successfully integrated into future health promotion interventions to promote PA and to support children and adolescents to reach their recommendations of PA constantly.

Strengths and Limitations

There are some strengths of this review that should be noted. One of these is focusing on healthy children and adolescents. To date, only a small number of reviews had their focus on this population group without any medical diseases or health restrictions. Another strength is the use of an established evidence synthesis method to evaluate the effects of the tools in the field of mHealth or wearable activity trackers on PA-related outcomes. Moreover, comprehensive conditions have to be fulfilled to achieve high methodological quality. All studies with tools of mHealth could be assessed according to this scheme.

However, there are also some limitations. It is appropriate to include only RCTs in a review. They are considered to be studies with high methodological quality so that their results are particularly meaningful [57]. In this review, several study designs were included that have a methodologically lower quality. This is due to the fact that for the review question about trials with tools of wearable activity trackers, studies with a high-quality design were not available at a preliminary examination. For this reason, these study effects need to be interpreted with caution. Furthermore, only PA-related outcomes were considered and analyzed. However, PA was often not the primary outcome of the included studies; therefore, conclusions on the effect on PA are limited.

Conclusions

On the basis of the findings of this study, to date, no clear recommendations can be derived. Some studies made restrictions in relation to participants' PA level so that the populations were not always compared with each other. Moreover, most of the studies based their intervention on several components so that the focus was not only on the mHealth tool or wearable activity tracker. However, as tools in the field of mHealth, mobile games as apps were widely accepted. Future research should focus on developing age-appropriate games to increase PA among children and adolescents. In addition, multicomponent approaches could be more effective in encouraging PA among the youth and should be promoted. A combination of school-based interventions with family or community involvement for social support was applied in some studies and could be an effective strategy.

Overall, the evidence of no effect for intervention studies with tools of mHealth on PA-related outcomes as well as both studies with wearable activity trackers with lower methodological quality shows a clear need for future intervention programs. There is a great lack of studies that seems to exist, especially in the European area. Future studies should be based on IM to develop theory- and evidence-based interventions. By means of this framework, implementation issues are also becoming transparent. Moreover, future studies should aim to strengthen the evidence with a high methodological quality design, an appropriate sample size, a focus on special target groups, follow-up beyond postintervention to assess sustainability, and the use of objective and valid measuring instruments to determine overall activity. In addition, for future transfer of strategies into public health promotion, cost-effectiveness analyses should be carried out.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy including MeSH (Cochrane Central Register of Controlled Trials, PubMed).

[PDF File (Adobe PDF File), 413KB - mhealth_v7i4e8298_app1.pdf]

Multimedia Appendix 2

Search strategy excluding MeSH (Scopus, SPORTDiscus, Web of Science).

[PDF File (Adobe PDF File), 428KB - mhealth_v7i4e8298_app2.pdf]

Multimedia Appendix 3

Criteria for assessment of methodological quality.

[PDF File (Adobe PDF File), 458KB - mhealth_v7i4e8298_app3.pdf]

Multimedia Appendix 4

Intervention characteristics of included studies aimed to promote physical activity among healthy children and/or adolescents with mHealth tools.

[PDF File (Adobe PDF File), 103KB - mhealth_v7i4e8298_app4.pdf]

Multimedia Appendix 5

Intervention characteristics of included studies aimed to promote physical activity among children and/or adolescents with wearable activity trackers.

[PDF File (Adobe PDF File), 87KB - mhealth_v7i4e8298_app5.pdf]

Multimedia Appendix 6

Methodological quality of the selected intervention studies with group comparison.

[PDF File (Adobe PDF File), 171KB - mhealth_v7i4e8298_app6.pdf]

Multimedia Appendix 7

Methodological quality of the selected intervention studies (number of studies and percentages).

[PDF File (Adobe PDF File), 132KB - mhealth_v7i4e8298_app7.pdf]

References

1. World Health Organization. 2014. Global status report on noncommunicable diseases 2014 URL: https://www.who.int/nmh/publications/ncd-status-report-2014/en/ [accessed 2019-04-15] [WebCite Cache ID 77etlFPgW]
2. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
3. Van Kann DH, Jansen MW, de Vries SI, de Vries NK, Kremers SP. Active Living: development and quasi-experimental evaluation of a school-centered physical activity intervention for primary school children. BMC Public Health 2015 Dec 29;15:1315 [FREE Full text] [doi: 10.1186/s12889-015-2633-1] [Medline: 26714755]
4. Manz K, Schlack R, Poethko-Müller C, Mensink G, Finger J, Lampert T, KiGGS Study Group. [Physical activity and electronic media use in children and adolescents: results of the KiGGS study: first follow-up (KiGGS wave 1)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2014 Jul;57(7):840-848. [doi: 10.1007/s00103-014-1986-4] [Medline: 24950833]
5. Katzmarzyk PT, Denstel KD, Beals K, Bolling C, Wright C, Crouter SE, et al. Results from the United States of America's 2016 report card on physical activity for children and youth. J Phys Act Health 2016 Dec;13(11 Suppl 2):S307-S313 [FREE Full text] [doi: 10.1123/jpah.2016-0321] [Medline: 27848726]
6. Barnes JD, Cameron C, Carson V, Chaput JP, Faulkner GE, Janson K, et al. Results from Canada's 2016 ParticipACTION report card on physical activity for children and youth. J Phys Act Health 2016 Dec;13(11 Suppl 2):S110-S116. [doi: 10.1123/jpah.2016-0300] [Medline: 27848727]

7. Wijtzes AI, Verloigne M, Mouton A, Cloes M, De Ridder KA, Cardon G, et al. Results from Belgium's 2016 report card on physical activity for children and youth. J Phys Act Health 2016 Dec;13(11 Suppl 2):S95-S103. [doi: 10.1123/jpah.2016-0306] [Medline: 27848723]
8. World Health Organization. Physical activity and young people: recommended levels of physical activity for children aged 5-17 years URL: http://www.who.int/dietphysicalactivity/factsheet_young_people/en/ [accessed 2019-04-15] [WebCite Cache ID 77eOwwv4y]
9. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
10. World Health Organization. 2015 Jan 20. Noncommunicable diseases URL: https://www.who.int/en/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2019-04-15] [WebCite Cache ID 77eo6oB6b]
11. World Health Organization. 2009. Global health risks: Mortality and burden of disease attributable to selected major risks URL: https://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf [accessed 2019-04-15] [WebCite Cache ID 77etp9YrT]
12. Telama R, Yang X, Leskinen E, Kankaanpää A, Hirvensalo M, Tammelin T, et al. Tracking of physical activity from early childhood through youth into adulthood. Med Sci Sports Exerc 2014;46(5):955-962. [doi: 10.1249/MSS.0000000000000181] [Medline: 24121247]
13. Department of Health. 2011. Start Active, Stay Active: a report on physical activity from the four home countries? Chief Medical Officers URL: https://www.sportengland.org/media/2928/dh_128210.pdf [accessed 2019-04-15] [WebCite Cache ID 77eQ5rAIV]
14. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129 [FREE Full text] [doi: 10.2196/mhealth.6540] [Medline: 27881359]
15. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act 2010 May 11;7:40 [FREE Full text] [doi: 10.1186/1479-5868-7-40] [Medline: 20459784]
16. Altenburg TM, Kist-van Holthe J, Chinapaw MJ. Effectiveness of intervention strategies exclusively targeting reductions in children's sedentary time: a systematic review of the literature. Int J Behav Nutr Phys Act 2016 Jun 09;13:65 [FREE Full text] [doi: 10.1186/s12966-016-0387-5] [Medline: 27276873]
17. Bucksch J, Dreger S. [Sedentary behavior as a risk factor in childhood and adolescence]. Präv Gesundheitsf 2013 Nov 19;9(1):39-46 [FREE Full text] [doi: 10.1007/s11553-013-0413-2]
18. Lau PW, Lau EY, Wong DP, Ransdell L. A systematic review of information and communication technology-based interventions for promoting physical activity behavior change in children and adolescents. J Med Internet Res 2011 Jul 13;13(3):e48 [FREE Full text] [doi: 10.2196/jmir.1533] [Medline: 21749967]
19. Carlin A, Murphy MH, Gallagher AM. Current influences and approaches to promote future physical activity in 11-13 year olds: a focus group study. BMC Public Health 2015 Dec 21;15:1270 [FREE Full text] [doi: 10.1186/s12889-015-2601-9] [Medline: 26689932]
20. Fatehi F, Wootton R. Telemedicine, telehealth or e-health? A bibliometric analysis of the trends in the use of these terms. J Telemed Telecare 2012 Dec;18(8):460-464. [doi: 10.1258/jtt.2012.gth108] [Medline: 23209265]
21. World Health Organization. 2011. mHealth: New horizons for health through mobile technologies URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2019-04-15] [WebCite Cache ID 77eR12ine]
22. Bitkom. Pressekonferenz - Studie ?Kinder und Jugend 3 URL: https://tinyurl.com/y4psbf7n [accessed 2019-04-15] [WebCite Cache ID 77erDExNg]
23. Maslakovic M. Gadgets & Wearables. 2016. Top fitness trackers for kids URL: http://gadgetsandwearables.com/2016/04/11/fitness-trackers-kids/ [accessed 2019-04-15] [WebCite Cache ID 77eW68A5h]
24. Moser M. Vodafone. 2016. Digital parenting guide: Digital helpers for a safe everyday life with a child] URL: https://tinyurl.com/y6tnj3xy [accessed 2019-04-16] [WebCite Cache ID 77fiVYoWa]
25. Kahn JG, Yang JS, Kahn JS. 'Mobile' health needs and opportunities in developing countries. Health Aff (Millwood) 2010 Feb;29(2):252-258. [doi: 10.1377/hlthaff.2009.0965] [Medline: 20348069]
26. Fritz T, Huang EM, Murphy GC, Zimmermann T. Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness, in Presented at: CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2014. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2014 Presented at: Conference on Human Factors in Computing Systems; April 26-May 01, 2014; Toronto, Ontario, Canada p. 487-496. [doi: 10.1145/2556288.2557383]
27. Balow A. vívofit. [The first fitness tracker for children by Garmin] URL: https://tinyurl.com/yxanz28k [accessed 2019-04-15] [WebCite Cache ID 77epK7zwZ]
28. CleanKids. 2015. [Fitness tracker for children? useful or not?] URL: https://tinyurl.com/y3obm3zx [accessed 2019-04-15] [WebCite Cache ID 77etsj9IF]

29. Jee H. Review of researches on smartphone applications for physical activity promotion in healthy adults. J Exerc Rehabil 2017 Feb;13(1):3-11 [FREE Full text] [doi: 10.12965/jer.1732928.464] [Medline: 28349027]
30. Ly H. The impact of utilizing mobile phones to promote physical activity among post-secondary students: a scoping review. Mhealth 2016;2:47 [FREE Full text] [doi: 10.21037/mhealth.2016.12.03] [Medline: 28293617]
31. Flores Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015 Nov 10;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
32. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]
33. Chaplais E, Naughton G, Thivel D, Courteix D, Greene D. Smartphone interventions for weight treatment and behavioral change in pediatric obesity: a systematic review. Telemed J E Health 2015 Oct;21(10):822-830. [doi: 10.1089/tmj.2014.0197] [Medline: 26290954]
34. Chen JL, Wilkosz ME. Efficacy of technology-based interventions for obesity prevention in adolescents: a systematic review. Adolesc Health Med Ther 2014;5:159-170 [FREE Full text] [doi: 10.2147/AHMT.S39969] [Medline: 25177158]
35. Rehman H, Kamal AK, Sayani S, Morris PB, Merchant AT, Virani SS. Using mobile health (mHealth) technology in the management of diabetes mellitus, physical inactivity, and smoking. Curr Atheroscler Rep 2017 Apr;19(4):16. [doi: 10.1007/s11883-017-0650-5] [Medline: 28243807]
36. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do mobile phone applications improve glycemic control (HbA1c) in the self-management of diabetes? A systematic review, meta-analysis, and GRADE of 14 randomized trials. Diabetes Care 2016 Nov;39(11):2089-2095. [doi: 10.2337/dc16-0346] [Medline: 27926892]
37. Ershad Sarabi R, Sadoughi F, Jamshidi Orak R, Bahaadinbeigy K. The effectiveness of mobile phone text messaging in improving medication adherence for patients with chronic diseases: a systematic review. Iran Red Crescent Med J 2016 May;18(5):e25183 [FREE Full text] [doi: 10.5812/ircmj.25183] [Medline: 27437126]
38. Gandapur Y, Kianoush S, Kelli HM, Misra S, Urrea B, Blaha MJ, et al. The role of mHealth for improving medication adherence in patients with cardiovascular disease: a systematic review. Eur Heart J Qual Care Clin Outcomes 2016 Oct 01;2(4):237-244 [FREE Full text] [doi: 10.1093/ehjqcco/qcw018] [Medline: 29474713]
39. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
40. An HS, Jones GC, Kang SK, Welk GJ, Lee JM. How valid are wearable physical activity trackers for measuring steps? Eur J Sport Sci 2017 Apr;17(3):360-368. [doi: 10.1080/17461391.2016.1255261] [Medline: 27912681]
41. Schaefer S, Ching CC, Breen H, German JB. Wearing, thinking, and moving: testing the feasibility of fitness tracking with urban youth. Am J Health Educ 2016 Jan 08;47(1):8-16 [FREE Full text] [doi: 10.1080/19325037.2015.1111174]
42. Schaefer SE, Van Loan M, German JB. A feasibility study of wearable activity monitors for pre-adolescent school-age children. Prev Chronic Dis 2014 May 22;11:E85 [FREE Full text] [doi: 10.5888/pcd11.130262] [Medline: 24854236]
43. Wang JB, Cadmus-Bertram LA, Natarajan L, White MM, Madanat H, Nichols JF, et al. Wearable sensor/device (Fitbit One) and SMS text-messaging prompts to increase physical activity in overweight and obese adults: a randomized controlled trial. Telemed J E Health 2015 Oct;21(10):782-792 [FREE Full text] [doi: 10.1089/tmj.2014.0176] [Medline: 26431257]
44. Lewis ZH, Lyons EJ, Jarvis JM, Baillargeon J. Using an electronic activity monitor system as an intervention modality: a systematic review. BMC Public Health 2015 Jun 24;15:585 [FREE Full text] [doi: 10.1186/s12889-015-1947-3] [Medline: 26104189]
45. Schoffman DE, Turner-McGrievy G, Jones SJ, Wilcox S. Mobile apps for pediatric obesity prevention and treatment, healthy eating, and physical activity promotion: just fun and games? Transl Behav Med 2013 Sep;3(3):320-325 [FREE Full text] [doi: 10.1007/s13142-013-0206-3] [Medline: 24073184]
46. Wearing JR, Nollen N, Befort C, Davis AM, Agemy CK. iPhone app adherence to expert-recommended guidelines for pediatric obesity prevention. Child Obes 2014 Apr;10(2):132-144 [FREE Full text] [doi: 10.1089/chi.2013.0084] [Medline: 24655230]
47. Mameli C, Brunetti D, Colombo V, Bedogni G, Schneider L, Penagini F, et al. Combined use of a wristband and a smartphone to reduce body weight in obese children: randomized controlled trial. Pediatr Obes 2018 Dec;13(2):81-87. [doi: 10.1111/ijpo.12201] [Medline: 27900849]
48. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
49. Weihrauch-Blüher S, Koormann S, Brauchmann J, Wiegand S. [Electronic media in obesity prevention in childhood and adolescence]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2016 Nov;59(11):1452-1464. [doi: 10.1007/s00103-016-2455-z] [Medline: 27757512]

50. Mojica CM, Parra-Medina D, Yin Z, Akopian D, Esparza LA. Assessing media access and use among Latina adolescents to inform development of a physical activity promotion intervention incorporating text messaging. Health Promot Pract 2014 Jul;15(4):548-555 [FREE Full text] [doi: 10.1177/1524839913514441] [Medline: 24357863]
51. Hoy MB. Personal activity trackers and the quantified self. Med Ref Serv Q 2016;35(1):94-100. [doi: 10.1080/02763869.2016.1117300] [Medline: 26794199]
52. Pal S, Cheng C, Egger G, Binns C, Donovan R. Using pedometers to increase physical activity in overweight and obese women: a pilot study. BMC Public Health 2009 Aug 25;9:309 [FREE Full text] [doi: 10.1186/1471-2458-9-309] [Medline: 19703317]
53. Marshall AL. Should all steps count when using a pedometer as a measure of physical activity in older adults? J Phys Act Health 2007 Jul;4(3):305-314. [Medline: 17846459]
54. Van Hoye K, Boen F, Lefevre J. The impact of different degrees of feedback on physical activity levels: a 4-week intervention study. Int J Environ Res Public Health 2015 Jun 09;12(6):6561-6581 [FREE Full text] [doi: 10.3390/ijerph120606561] [Medline: 26067990]
55. Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. Cochrane Database Syst Rev 2011 Aug 10(8):MR000026. [doi: 10.1002/14651858.MR000026.pub2] [Medline: 21833989]
56. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
57. Sterne JA, Sutton AJ, Ioannidis JA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. Br Med J 2011 Jul 22;343:d4002. [doi: 10.1136/bmj.d4002] [Medline: 21784880]
58. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009 Oct;62(10):e1-34 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.06.006] [Medline: 19631507]
59. Higgins JP, Green S. Cochrane. 2011. Cochrane Handbook for Systematic Reviews of Interventions. 5.1.0 updated March 2011, The Cochrane Collaboration URL: http://handbook.cochrane.org [accessed 2019-04-15] [WebCite Cache ID 77eg1Egji]
60. van Sluijs EM, McMinn AM, Griffin SJ. Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. Br J Sports Med 2008 Aug;42(8):653-657. [Medline: 18685076]
61. Demetriou Y, Sudeck G, Thiel A, Höner O. The effects of school-based physical activity interventions on students' health-related fitness knowledge: a systematic review. Educational Research Review 2015 Oct;16:19-40. [doi: 10.1016/j.edurev.2015.07.002]
62. Liang Y, Lau PW. Effects of active videogames on physical activity and related outcomes among healthy children: a systematic review. Games Health J 2014 Jun;3(3):122-144. [doi: 10.1089/g4h.2013.0070] [Medline: 26196172]
63. van Sluijs EM, van Poppel MN, van Mechelen W. Stage-based lifestyle interventions in primary care: are they effective? Am J Prev Med 2004 May;26(4):330-343. [doi: 10.1016/j.amepre.2003.12.010] [Medline: 15110061]
64. Engbers LH, van Poppel MN, Chin A Paw MJ, van Mechelen W. Worksite health promotion programs with environmental changes: a systematic review. Am J Prev Med 2005 Jul;29(1):61-70. [doi: 10.1016/j.amepre.2005.03.001] [Medline: 15958254]
65. Dewar DL, Morgan PJ, Plotnikoff RC, Okely AD, Batterham M, Lubans DR. Exploring changes in physical activity, sedentary behaviors and hypothesized mediators in the NEAT girls group randomized controlled trial. J Sci Med Sport 2014 Jan;17(1):39-46. [doi: 10.1016/j.jsams.2013.02.003] [Medline: 23506657]
66. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for IMproving FITness and Increasing Physical Activity Among Young People: the AIMFIT pragmatic randomized controlled trial. J Med Internet Res 2015 Aug 27;17(8):e210 [FREE Full text] [doi: 10.2196/jmir.4568] [Medline: 26316499]
67. Garde A, Umedaly A, Abulnaga SM, Junker A, Chanoine JP, Johnson M, et al. Evaluation of a novel mobile exergame in a school-based environment. Cyberpsychol Behav Soc Netw 2016 Mar;19(3):186-192. [doi: 10.1089/cyber.2015.0281] [Medline: 26882222]
68. Lubans DR, Smith JJ, Plotnikoff RC, Dally KA, Okely AD, Salmon J, et al. Assessing the sustained impact of a school-based obesity prevention program for adolescent boys: the ATLAS cluster randomized controlled trial. Int J Behav Nutr Phys Act 2016 Aug 20;13:92 [FREE Full text] [doi: 10.1186/s12966-016-0420-8] [Medline: 27542825]
69. Zach S, Raviv T, Meckel Y. Using information communication technologies (ICTs) for motivating female adolescents to exercise/run in their leisure time. Computers in Human Behavior 2016 Jul;60:593-601 [FREE Full text] [doi: 10.1016/j.chb.2016.02.096]
70. Bronikowski M, Bronikowska M, Glapa A. Do they need goals or support? A report from a goal-setting intervention using physical activity monitors in youth. Int J Environ Res Public Health 2016 Dec 13;13(9) [FREE Full text] [doi: 10.3390/ijerph13090914] [Medline: 27649219]
71. Hayes LB, Van Camp CM. Increasing physical activity of children during school recess. J Appl Behav Anal 2015 Sep;48(3):690-695. [doi: 10.1002/jaba.222] [Medline: 26119136]
72. Lubans DR, Morgan PJ, Dewar D, Collins CE, Plotnikoff RC, Okely AD, et al. The Nutrition and Enjoyable Activity for Teen Girls (NEAT girls) randomized controlled trial for adolescent girls from disadvantaged secondary schools: rationale,

study protocol, and baseline results. BMC Public Health 2010 Oct 28;10:652 [FREE Full text] [doi: 10.1186/1471-2458-10-652] [Medline: 21029467]
73. Smith JJ, Morgan PJ, Plotnikoff RC, Dally KA, Salmon J, Okely AD, et al. Rationale and study protocol for the 'active teen leaders avoiding screen-time' (ATLAS) group randomized controlled trial: an obesity prevention intervention for adolescent boys from schools in low-income communities. Contemp Clin Trials 2014 Jan;37(1):106-119. [doi: 10.1016/j.cct.2013.11.008] [Medline: 24291151]
74. Garde A, Umedaly A, Abulnaga SM, Robertson L, Junker A, Chanoine JP, et al. Assessment of a Mobile Game ("MobileKids Monster Manor") to Promote Physical Activity Among Children. Games Health J 2015 Apr;4(2):149-158. [doi: 10.1089/g4h.2014.0095] [Medline: 26181809]
75. Page SJ, Persch AC. Recruitment, retention, and blinding in clinical trials. Am J Occup Ther 2013;67(2):154-161 [FREE Full text] [doi: 10.5014/ajot.2013.006197] [Medline: 23433269]
76. Beaglehole R, Bonita R, Kjellstrom T. World Health Organization. Geneva; 1993. [Introduction into Epidemiology] URL: https://apps.who.int/iris/bitstream/handle/10665/43541/9783456845357_ger_Contents.pdf?sequence=2&isAllowed=y
77. Lubans DR, Morgan PJ, Okely AD, Dewar D, Collins CE, Batterham M, et al. Preventing obesity among adolescent girls: one-year outcomes of the nutrition and enjoyable activity for teen girls (NEAT Girls) cluster randomized controlled trial. Arch Pediatr Adolesc Med 2012 Sep 01;166(9):821-827. [doi: 10.1001/archpediatrics.2012.41] [Medline: 22566517]
78. Tudor-Locke C, Craig CL, Beets MW, Belton S, Cardon GM, Duncan S, et al. How many steps/day are enough? for children and adolescents. Int J Behav Nutr Phys Act 2011 Jul 28;8:78 [FREE Full text] [doi: 10.1186/1479-5868-8-78] [Medline: 21798014]
79. Lonsdale C, Rosenkranz RR, Peralta LR, Bennie A, Fahey P, Lubans DR. A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical education lessons. Prev Med 2013 Feb;56(2):152-161. [doi: 10.1016/j.ypmed.2012.12.004] [Medline: 23246641]
80. Guerra PH, Nobre MR, Silveira JA, Taddei JA. The effect of school-based physical activity interventions on body mass index: a meta-analysis of randomized trials. Clinics 2013 Sep 26;68(9):1263-1273 [FREE Full text] [doi: 10.6061/clinics/2013(09)14] [Medline: 24141844]
81. Lampert T, Mensink GB, Romahn N, Woll A. [Physical activity among children and adolescents in Germany. Results of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2007;50(5-6):634-642. [doi: 10.1007/s00103-007-0224-8] [Medline: 17514447]
82. 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]
83. Direito A, Jiang Y, Whittaker R, Maddison R. Smartphone apps to improve fitness and increase physical activity among young people: protocol of the Apps for IMproving FITness (AIMFIT) randomized controlled trial. BMC Public Health 2015 Jul 11;15:635 [FREE Full text] [doi: 10.1186/s12889-015-1968-y] [Medline: 26159834]
84. Nguyen S, Häcker AL, Henderson M, Barnett T, Mathieu ME, Pagani L, et al. Physical activity programs with post-intervention follow-up in children: a comprehensive review according to categories of intervention. Int J Environ Res Public Health 2016 Dec 30;13(7) [FREE Full text] [doi: 10.3390/ijerph13070664] [Medline: 27376315]
85. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Dec 26;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
86. Pedišić Ž, Bauman A. Accelerometer-based measures in physical activity surveillance: current practices and issues. Br J Sports Med 2015 Feb;49(4):219-223. [doi: 10.1136/bjsports-2013-093407] [Medline: 25370153]
87. Kok G, Bartholomew LK, Parcel GS, Gottlieb NH, Fernández ME. Finding theory- and evidence-based alternatives to fear appeals: intervention mapping. Int J Psychol 2014 Apr;49(2):98-107 [FREE Full text] [doi: 10.1002/ijop.12001] [Medline: 24811880]
88. Salmon J, Booth ML, Phongsavan P, Murphy N, Timperio A. Promoting physical activity participation among children and adolescents. Epidemiol Rev 2007;29:144-159. [doi: 10.1093/epirev/mxm010] [Medline: 17556765]
89. Kriemler S, Meyer U, Martin E, van Sluijs EM, Andersen LB, Martin BW. Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update. Br J Sports Med 2011 Sep;45(11):923-930 [FREE Full text] [doi: 10.1136/bjsports-2011-090186] [Medline: 21836176]
90. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? Health Educ Res 2011 Apr;26(2):308-322. [doi: 10.1093/her/cyr005] [Medline: 21321008]
91. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717. [Medline: 12237980]

Abbreviations

BCT: behavior change technique

BMI: body mass index
CENTRAL: Cochrane Central Register of Controlled Trials
CG: control group
CT: controlled trial
IG: intervention group
IM: intervention mapping
MeSH: Medical Subject Headings
mHealth: mobile health
MKMM: MobileKids Monster Manor
MPA: moderate physical activity
MVPA: moderate- to vigorous-intensity physical activity
PA: physical activity
PE: physical education
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SMS: short message service
VPA: vigorous physical activity
WHO: World Health Organization

Edited by G Eysenbach; submitted 27.06.17; peer-reviewed by J Prescott, M Lang, N Ridgers, G Kok, P Wark; comments to author 21.08.17; revised version received 18.06.18; accepted 07.02.19; published 30.04.19.

Please cite as:
Böhm B, Karwiese SD, Böhm H, Oberhoffer R
Effects of Mobile Health Including Wearable Activity Trackers to Increase Physical Activity Outcomes Among Healthy Children and Adolescents: Systematic Review
JMIR Mhealth Uhealth 2019;7(4):e8298
URL: http://mhealth.jmir.org/2019/4/e8298/
doi:10.2196/mhealth.8298
PMID:31038460

Original Paper

Relative Validity of a Method Based on a Smartphone App (Electronic 12-Hour Dietary Recall) to Estimate Habitual Dietary Intake in Adults

Luis María Béjar[1], PhD; María Dolores García-Perea[2], PhD; Óscar Adrián Reyes[3], BMed; Esther Vázquez-Limón[1], PhD

[1]Department of Preventive Medicine and Public Health, School of Medicine, University of Seville, Seville, Spain

[2]Virgen Macarena University Hospital, Seville, Spain

[3]Mutua Balear, Seville, Spain

Corresponding Author:
Luis María Béjar, PhD
Department of Preventive Medicine and Public Health
School of Medicine
University of Seville
Institute of Anatomy, 3rd Floor
Sánchez-Pizjuán Avenue
Seville, 41009
Spain
Phone: 34 954551771
Fax: 34 954556481
Email: lmbprado@us.es

Abstract

Background: Accurate dietary assessment is key to understanding nutrition-related outcomes and for estimating the dietary change in nutrition-based interventions. When researching the habitual consumption of selected food groups, it is essential to be aware of factors that could possibly affect reporting accuracy.

Objective: This study aimed to evaluate the *relative validity* of the current-day dietary recall, a method based on a smartphone app called electronic 12-hour dietary recall (e-12HR), to categorize individuals according to habitual intake, in the whole sample of adults and in different strata thereof.

Methods: University students and employees over 18 years recorded the consumption of 10 selected groups of food using e-12HR during 28 consecutive days. During this period, they also completed 4 dietary records. Once the period was finished, the subjects then completed a food frequency questionnaire (FFQ) and a usability-rating questionnaire for e-12HR. The food group intakes estimated by the e-12HR app, the dietary records, and the FFQ were categorized into sextiles: *less than once a week*, *once or twice a week*, *3-4 times a week*, *5-6 times a week*, *once or twice a day*, and *3 or more times a day*. The 10 selected groups with e-12HR were compared with 4 dietary records and an FFQ reference method, in the whole sample and in different strata thereof: age (years): <25 and ≥25; gender: females and males; occupation: students and employees; smoking: no and yes; physical activity (minutes/week): ≥150 and <150; and body mass index (kg/m^2): <25 and ≥25. The association between the different methods was assessed using Spearman correlation coefficient (SCC). Cross-classification and kappa statistic were used as a measure of agreement between the different methods.

Results: In total, 203 participants completed the study (56.7% [115/203] women, and 43.3% [88/203] men). For all food groups and all participants, the mean SCC for e-12HR versus FFQ was 0.67 (≥0.62 for all strata). On average, 50.7% of participants were classified into the same category (≥47.0% for all strata) and 90.2% within the nearest category (≥88.6% for all strata). Mean weighted kappa was 0.49 (≥0.44 for all strata). For e-12HR versus RDs, mean SCC was 0.65 (≥0.57 for all strata). On average, 50.0% of participants were classified into the same category (≥47.0% for all strata) and 88.2% within the nearest category (≥86.1% for all strata). Mean weighted kappa was 0.50 (≥0.44 for all strata).

Conclusions: The results indicate that e-12HR generated categories of dietary intake highly comparable with the 2 reference methods in the whole sample and in different strata thereof. The inclusion of photographs to facilitate estimation of the servings

consumed generated correlation/agreement data between e-12HR and the FFQ that were similar to a previous study using an older version of the app, which did not include photographs.

(JMIR Mhealth Uhealth 2019;7(4):e11531) doi:10.2196/11531

KEYWORDS

epidemiologic methods; diet records; mobile apps; nutrition assessment

Introduction

Background

Habitual intake (or average long-term consumption) is an essential part of epidemiological investigations and intervention studies [1-3]. Many of these studies do not require the characterization of all foods and beverages consumed (hereafter referred to as *food*) [4], as it can represent an unnecessary workload for study participants and an avoidable waste of the scarce resources available for research [5]. The characterization of foods may mean assessing whether survey items can be reduced to binaries (was a food eaten or not?) or requiring an accurate weight [6]. Categorizing individuals according to categories of habitual consumption of specific food groups might be used for evaluating the relationship between relative ranking and disease [2,7-13] and for evaluating the effectiveness of personalized methods that are implemented to promote changes in dietary patterns [2,4,8-11] with regard to the selected food groups.

Dietary records (DRs) and 24-hour recalls (short-term methods), and food frequency questionnaires (FFQs; a long-term method) are the 3 main assessment methods that are traditionally used to assess dietary intake [14-16]. The strengths and weaknesses of these instruments are well documented [2,7,14,17-20].

In large-scale epidemiological and intervention studies, where detailed dietary assessment is not feasible [9], FFQs have been the most accessible and commonly utilized dietary assessment tool [1,7,10,11,15,16,21]. FFQs are retrospective methods that require respondents to report the frequency of consumption of a predefined list of food groups over an extended period of time (weeks or months) [14,22]. FFQs are practical and easy to administer; they do not affect food intake patterns and can assess habitual dietary patterns with a single administration [10]. One inherent limitation to most FFQs is that they are paper-based. As a result, on the one hand, errors such as skipped questions or multiple marks are common, whereas on the other hand, they do not allow precise estimation of food portion size [8], and finally, there is the necessary posterior manual introduction of data for statistical analysis, which increases research costs and time consumption considerably [19,23]. FFQs in digital format (mobile phone apps or Web-based) offer straightforward solutions to these limitations, incorporating complex skip patterns and a broad and varying number of portion-size options for extensive food groups. In addition, FFQs administered electronically do not require posterior manual introduction of the collected data [14,23-27]. However, all FFQs (paper and digital format) depend on the long-term memory of the interviewed subject, and they do not take day-to-day intrapersonal variation into account during the period of the study [2,7,14,17-20]. For these reasons, developing new methods that overcome the limitations of FFQs to assess the habitual intake of selected food groups in large-scale epidemiological and intervention studies is well motivated.

The aim of this study was to evaluate the *relative validity* of the current-day dietary recall (current-day recall), a method which is based on a smartphone app called electronic 12-hour dietary recall (e-12HR). Moreover, 4 estimated DRs and a semiquantitative FFQ were used as reference methods to verify comparability of the data with regard to 10 selected food groups among the whole sample and across different strata (sociodemographic characteristics, lifestyle factors, and weight category).

Previous Research

When researching the habitual consumption of selected food groups, it is essential to be aware of factors that could possibly affect reporting accuracy: gender, age, ethnic group, education level, occupation, employment status, socioeconomic status, diagnosed diseases, a sedentary lifestyle, slimming regimens, smoking, alcohol consumption, weight category, psychological factors, etc [28-31].

This study is an extension of the study previously published in *JMIR mHealth and uHealth* titled *Electronic 12-Hour Dietary Recall (e-12HR): Comparison of a Mobile Phone App for Dietary Intake Assessment With a Food Frequency Questionnaire and Four Dietary Records* [32]. This study, with regard to the previous one, compares e-12HR against an FFQ and 4 DRs in a whole sample of adults (students and employees of the Schools of Medicine or Pharmacy, University of Seville), and in different strata thereof (sociodemographic characteristics, lifestyle factors and weight category), and not only a sample of the university students. It is true that the research team uses, consciously, exactly the same protocol and the same statistical analysis with the idea of making the results comparable. However, in the study at hand, the sample is increased to include new population groups. This has allowed the research team to analyze, on the one hand, if the study objective has been achieved (determining the *relative validity* of current-day recall), which can be extrapolated to a wider amount of the target population, and on the other hand, the influence of certain factors (age, gender, occupation, smoking status, physical activity status, and weight category).

Methods

Procedure

The investigation protocol has been published previously elsewhere [32]. In brief, the study was carried out in 2 centers: the Schools of Medicine and Pharmacy at the University of Seville (Andalusia, Spain, South of Europe). Different events

were organized to present the project to the students and employees from both faculties. Participant recruitment took place from January 2017 to December 2017. The participants were incorporated in the study progressively during the entire recruitment period in such a way that every day of the week and every season of the year would be represented [33].

Inclusion criteria were as follows: (1) older than 18 years of age, (2) a student or employee of the Schools of Medicine or Pharmacy (University of Seville), and (3) possesses a smartphone with internet access (3G/4G/Wi-Fi) and an Android operating system.

All procedures on human beings were approved by the Research Ethics Committee at the University of Seville.

In the first interview, the participants started by providing informed consent; then, they were assigned a unique alphanumeric code to preserve their anonymity in accordance with current Spanish legislation [34], and they performed the following activities:

1. Each participant filled out an initial questionnaire (on paper), which included the date of the interview and self-reported date of birth, gender, occupation, weight and height measurements, as well as smoking and physical activity status. Body mass index (BMI; kg/m^2) was estimated from self-reported body weight and height [13,14,35].
2. Each participant downloaded the e-12HR app for their personal smartphone, and a member of the research team personally explained how to use the app with a practical demonstration before written instructions were given to the participants [12,36] to be consulted later if necessary.
3. The same research team member personally gave each participant detailed instructions on how to complete the 4 estimated DRs and how to estimate serving sizes consumed. In addition, an explanatory pamphlet was also given to the participants [10,11].

Written instructions ("how to use the app," "how to complete the four estimated DRs," and "how to estimate serving sizes consumed") are subject to copyright and thus are not included in the manuscript.

In the second interview, at the end of the e-12HR app data collection period and at the convenience of each participant, the participant was required to fill out a semiquantitative FFQ. A research team member explained to each participant the process for completing the FFQ. Finally, each participant filled out a usability rating questionnaire [14,28,37] for e-12HR app (Figure 1), which comprised 5 questions about the completion of e-12HR (Multimedia Appendix 1).

The Electronic 12-Hour Dietary Recall App

The e-12HR app was developed to record daily consumption of a list of 10 food groups: fruit, vegetables, legumes, chicken/turkey, fish, red meat, soft drinks, sweets, prepared foods, and beer (Multimedia Appendix 2). Other food groups such as dairy and derivatives, eggs, nuts, potatoes, pasta, rice, or bread have not been included. In any case, the food groups included can be modified to meet the needs of each study [32]. The list could not be too long to minimize the workload of the participants as well as the research costs [21]. These food groups were selected as they are indicators of health/disease and are considered protective factors (fruit, vegetables, legumes, or fish) or risk factors (soft drinks, commercial baked foods, and precooked meals) for chronic illnesses [1,10,38]. They also provide consumption patterns that range from almost every day for every inhabitant of the population to infrequently for the majority [1].

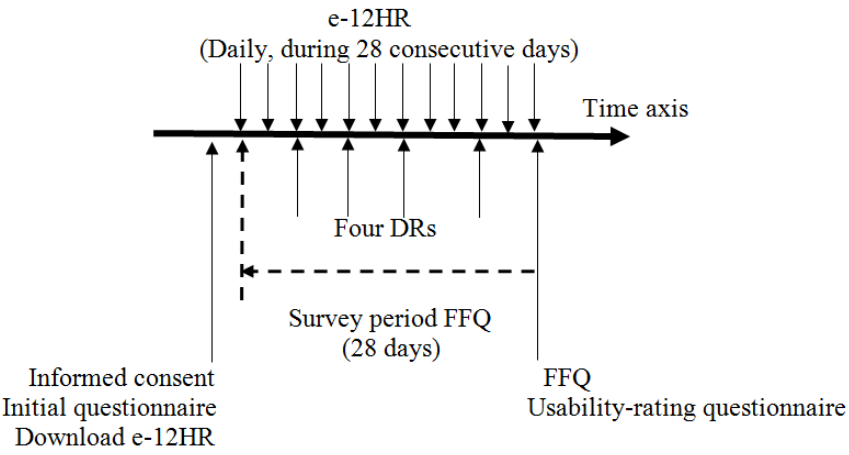
When the e-12HR app was used the first time, the participants were required to introduce their personally assigned alphanumeric code and the email of the researcher who would receive the data from the app. Participants were instructed to use the app after consuming the last food of the day [12,36]. For each food group, the participant would choose the most appropriate image (or images) from a series of color photographs with 2 to 4 possible options, shown simultaneously [12,36], that illustrated the different serving sizes to assist with selecting the number of standard servings consumed [7,10,12,28,36]. To further assist with estimating serving sizes, each photograph was accompanied by an explanatory text and 3 objects of known/predictable size [39,40] (fiducial markers): a commonly used pencil, pen, and a marker. For example, on the screen of the app, the following would appear: *How many servings of soft drinks have you had today?*, with the *Rations* button and *Next* button. Supposing that the participant had, throughout the day, consumed 2 cans of soft drinks, 1 normal size and another larger size, they would proceed as follows: (1) tap the *Rations* button-a new window opens with different photographs of soft drinks, an *Accept* button, and a *Cancel* button; (2) tap once on the photo corresponding to the normal size; (3) scroll down on the screen; (4) tap once on the photo corresponding to the large size; (5) tap the *Accept* button-the app returns to the previous window; and (6) tap the *Next* button to access the next food group and proceed as before-if an error occurs, the participant can tap the *Cancel* button instead of *Accept*, starting the process over again (Multimedia Appendix 3).

After completing the daily questionnaire with the e-12HR app, the information is automatically saved and sent, via 3G/4G/Wi-Fi, to the e-mail address of the research administrator. Once the questionnaire is completed and sent, the participant cannot change their responses or access the app until the following day.

The consumption record of the selected food groups on the app was performed for 28 consecutive days. The time interval selected is similar to other comparison/validation studies [3,10,13,14,35].

The questionnaire and the size of the rations used in the e-12HR app are based on a semiquantitative FFQ previously validated for the population of Spain [41].

Figure 1. Assessment process using the electronic 12-hour dietary recall (e-12HR) app, 4 dietary records (DRs), food frequency questionnaire (FFQ), and usability rating questionnaire for the e-12HR app.



Dietary Records

During the 28-day period that e-12HR was in use for each participant, 4 estimated DRs (on paper) were scheduled on randomly assigned, nonconsecutive days [9,13]: 3 days during the weekdays and 1 day during the weekend [9-11,13]. The choice between 3 and 7 DRs is normally considered sufficient to evaluate food group intake [42]. Four estimated DRs were chosen instead of weighed DRs for logistical reasons [9,10].

Each participant, during the first interview, received an explanation of how to use the estimated DRs and how to estimate the serving size consumed, through the use of a pamphlet with a series of 2 to 4 color photographs [7,11,12,36] (1 series for each food group). To assist with estimating serving sizes, each photograph was accompanied by an explanatory text and 3 reference objects of known/predictable size [39,40] (fiducial markers). The explanatory text and the fiducial markers were the same as for the e-12HR app.

The DRs used were based on a DR previously validated for another European country (Denmark) [11,43], but structured according to the typical Spanish diet (breakfast, lunch, an afternoon snack, and dinner), and precodified including the same 10 food groups selected for e-12HR. The precoded DR includes 10 rows (1 for each of the food groups selected by the study) and 3 columns for morning, afternoon and evening, and night (Multimedia Appendix 4). This was done to minimize the burden on the participants. The serving sizes were based on a semiquantitative FFQ previously validated for the Spanish population [41].

Participants were told that they must record the consumption data on a separate page for each day [29] and immediately after consuming the food [11,29].

Food Frequency Questionnaire

The FFQ was a structured, semiquantitative FFQ (on paper) that included the same 10 food groups selected for the e-12HR app and the DRs. A research team member provided participants with an explanatory pamphlet to estimate what was considered a standard serving for each food group. This pamphlet contained a photograph of a standard serving for each food group along with an explanatory text and 3 reference objects of a known/predictable size [39,40] (fiducial markers). The explanatory text and the fiducial markers were the same as for the e-12HR app and the DRs (for a standard serving). The time period considered by the FFQ corresponded to the 28 days of the app. All the participants completed the FFQ within the first week of finishing the e-12HR app with the exception of 4 participants, who completed the FFQ 8 to 14 days later.

The semiquantitative FFQ as well as the standard serving sizes were based on a semiquantitative FFQ previously validated for the Spanish population [41].

Data Conversion

Using e-12HR, each participant recorded the number of standard serving sizes consumed daily for each food group throughout the 28-day study period. With the 4 estimated DRs, each participant collected the number of standard serving sizes consumed daily for each food group on 4 different days throughout the 28-day monitoring period. On the semiquantitative FFQ, each participant selected the number of standard serving sizes habitually consumed for each food group throughout the 28-day monitoring period (Figure 1).

For each participant, the data from the e-12HR app, the 4 DRs, and the FFQ had to be expressed in the same categories of habitual consumption to make comparisons (6 categories: *less than once a week, once or twice a week, 3-4 times a week, 5-6 times a week, once or twice a day,* and *3 or more times a day*). On the FFQ, these different options for habitual consumption were already available for the participants to choose from, and as such, the FFQ data were not modified. With regard to the e-12HR app, the data needed to be transformed. As an example, 1 participant registered an average daily consumption of 0.76 standard servings of vegetables over 28 days using the app. This average consumption represents 5.32 standard servings per week (0.76×7=5.32), which would be classified as *5 to 6 times a week* [32,44,45]. As for the 4 DRs, the information they contained

also needed to be converted [9]. As an example, 1 participant recorded consuming 0, 0.5, and 1 standard pieces of red meat on the DRs during the weekdays and 0.5 standard pieces of red meat on the DR completed at the weekend. This represents an average daily consumption during weekdays of: (0 standard pieces+0.5 standard pieces+1 standard piece)/3 weekdays=0.5 standard pieces per weekday. For weekly consumption, the conversion was as follows: (0.5×5 weekdays)+(0.5×2 weekend days)=2.5+1=3.5 standard pieces, which would then be classified as *3 to 4 times a week*.

To make comparisons, the 3 tools registered the consumption of the same food groups, used the same standard servings as a reference, and the intake record corresponded to the same time period, to avoid possible variations in individual diets during different periods [13,22,28,46].

Statistical Analysis

The food group intakes estimated by the e-12HR app, the FFQ, and the DRs were categorized into sextiles. For each food group, the consumption category assigned by e-12HR is compared with the category assigned by each of the different reference methods (FFQ and 4 DRs). The association between dietary intake methods (the current-day recall vs the FFQ and vs the 4 DRs) was assessed using Spearman correlation coefficients (SCC) [4]. Cross-classification analysis and kappa statistic index [4] were used as a measure of agreement between the current-day recall and the FFQ/DRs. The proportion of subjects categorized in the same sextile by the different methods (labeled *exact agreement*), in the same or adjacent sextile (labeled exact *agreement + adjacent*), and in opposite sextiles (labeled *extreme disagreement*) was calculated. Kappa statistic index was weighted to take into account the degree of disagreement between the instruments, assigning partial credit to scores using the Stata prerecorded weights [47].

SCC can have a value between -1 and +1; according to Cohen cut-offs, r=±0.5 is considered strong, r=±0.30 is moderate, and r=±0.10 is weak [48]. Weighted kappa statistic index can oscillate between 1 and +1: values of weighted kappa statistic index over 0.80 indicate very good agreement, between 0.80 and 0.61 indicate good agreement, between 0.60 and 0.41 indicate moderate agreement, between 0.40 and 0.21 indicate fair agreement, and <0.20 indicate poor agreement [49]. The comparison criteria considered in this study were as follows: SCC≥0.5 [4,13]; cross-classification percentage in the *exact agreement category* ≥35.0% [13], in the *exact agreement+adjacent category* ≥75.0% [13], and in the *extreme disagreement category* ≤8.0% [14]; and a weighted kappa statistic index ≥0.41 [4].

All statistical tests were 2 sided, and a significance level was considered at *P* value <.05. All data were analyzed using the statistical software STATA version MP 13.1 (Stata Corp LP, College Station, Texas, USA) [47].

It is important to note that the cross-classification analysis and weighted kappa depend on the number of categories used [4]. For example, imagine 2 participants in the study, participants *A* and *B*. Participant *A* presents an average consumption of a specific food group of 3.2 standard servings per week; participant *B* presents an average consumption of the same food group of 5.4 standard servings per week. If the categories considered in the study were 3 categories (Category 1: less than 3 times a week; Category 2: 3-6 times a week; and Category 3: once or more times a day), both participants (*A* and *B*) would be included in category 2. However, if the categories considered in the study are 6 categories (Category 1: less than once a week; Category 2: once or twice a week; Category 3: 3-4 times a week; Category 4: 5-6 times a week; Category 5: once or twice a day; and Category 6: 3 or more times a day), then participant *A* would be included in category 3, whereas participant *B* would be included in category 4.

Results

Overview

Of the 217 participants who signed the informed consent, 14 did not complete the study. The results of these individuals were not included in later statistical analysis. Information on the number of days completed with e-12HR can be found in Table 1. To highlight, 58.1% (118/203) of the participants completed the task every day (28 days of monitoring), 10.3% (21/203) completed the task 27 days, and 11.8% (24/203) completed the app 26 days.

The average age of the participants was 32 years. Moreover, 59.6% (121/203) were ≥25 years old; 56.7% (115/203) were women. In addition, 57.1% (116/203) were employees and 42.9% (87/203) were students. A majority (83.7% [170/203]) of the participants were nonsmokers. Two-thirds (66.5% [135/203]) of the respondents performed 150 min or more of moderate-intensity physical activities per week [50]. The mean BMI was 24.2 kg/m^2, with 4.9% (10/203) of the participants in the underweight range (BMI<18.5), 61.1% (124/203) in the healthy weight range (BMI: 18.5-24.9), 25.1% (51/203) being overweight (BMI: 25.0-29.9), and 8.9% (18/203) obese (BMI>30.0) [51] (Table 1). No statistically significant differences in the variables studied were found among the participants who completed the study and those who did not.

Table 1. Characteristics of study participants.

Characteristics	Statistics, n (%)	Mean (SD)	95% CI
Participants who completed the study	203 (100)	—[a]	—
Number of days completed the app			
28 days	118 (58.1)	—	—
27 days	21 (10.3)	—	—
26 days	24 (11.8)	—	—
25 days	12 (5.9)	—	—
24 days	11 (5.4)	—	—
23 days	6 (3.0)	—	—
22 days	7 (3.4)	—	—
21 days	4 (2.0)	—	—
Age (years)	—	32.0 (11.4)	—
<25	82 (40.4)	—	33.6-47.2
≥25	121 (59.6)	—	52.8-66.4
Gender			
Females	115 (56.7)	—	49.8-63.5
Males	88 (43.3)	—	36.5-50.2
Occupation			
Students	87 (42.9)	—	36.0-49.7
Employees	116 (57.1)	—	50.3-64.0
Smoking status			
No	170 (83.7)	—	78.6-88.9
Yes	33 (16.3)	—	11.1-21.4
Physical activity status (minutes/week)			
≥150	135 (66.5)	—	60.0-73.1
<150	68 (33.5)	—	26.9-40.0
Body mass index (kg/m^2>)	—	24.2 (4.1)	—
<25	134 (66.0)	—	59.4-72.6
≥25	69 (34.0)	—	27.4-40.6

[a]—: not applicable.

The Electronic 12-Hour Dietary Recall App Versus the Food Frequency Questionnaire

For all the food groups, for all participants, and for all strata, the average SCC was 0.67 (by strata, from 0.62 [smokers] to 0.70 [student]; Tables 2 and 3). Cross-classification analysis showed that the average percentage of individuals classified in the *exact agreement* category was 50.7% (by strata, from 47.0% [males] to 53.6% [females]); *exact agreement+adjacent* was 90.2% (by strata, from 88.6% [BMI≥25] to 91.8% [students]); and no participants (0%) were classified in the *extreme disagreement* category (Tables 4 and 5; see Multimedia Appendix 5 for full details). The average weighted kappa was .49 (by strata, from 0.44 [smokers] to 0.51 [students]; Tables 6 and 7).

Table 2. Spearman correlation coefficients derived from the electronic 12-hour dietary recall app versus the food frequency questionnaire for categories of food group consumption and strata: age (years): <25 and ≥25; gender: females and males; and occupation: students and employees. Comparison of the electronic 12-hour dietary recall app versus food frequency questionnaire. Spearman correlation coefficients (95% CI).

Food groups	All[a]	Age (years)		Gender		Occupation	
		<25	≥25	Females	Males	Students	Employees
Fruit	0.80 (0.75-0.84)	0.82 (0.73-0.88)	0.79 (0.71-0.85)	0.76 (0.67-0.83)	0.85 (0.78-0.90)	0.84 (0.76-0.89)	0.77 (0.68-0.83)
Vegetables	0.70 (0.62-0.76)	0.77 (0.66-0.85)	0.64 (0.52-0.74)	0.70 (0.60-0.78)	0.69 (0.56-0.78)	0.80 (0.71-0.86)	0.61 (0.48-0.71)
Legumes	0.48 (0.37-0.58)	0.47 (0.29-0.63)	0.51 (0.37-0.63)	0.51 (0.36-0.63)	0.46 (0.27-0.61)	0.50 (0.32-0.64)	0.49 (0.34-0.62)
Chicken/Turkey	0.58 (0.48-0.67)	0.49 (0.31-0.64)	0.63 (0.51-0.73)	0.58 (0.45-0.69)	0.60 (0.45-0.72)	0.53 (0.36-0.66)	0.62 (0.49-0.72)
Fish	0.53 (0.42-0.62)	0.62 (0.47-0.74)	0.48 (0.33-0.60)	0.52 (0.37-0.64)	0.56 (0.40-0.69)	0.65 (0.51-0.76)	0.44 (0.28-0.58)
Red meat	0.63 (0.53-0.70)	0.71 (0.58-0.80)	0.58 (0.45-0.69)	0.65 (0.53-0.75)	0.54 (0.37-0.67)	0.69 (0.56-0.79)	0.58 (0.44-0.69)
Soft drinks	0.83 (0.78-0.87)	0.79 (0.69-0.86)	0.85 (0.79-0.89)	0.82 (0.75-0.87)	0.84 (0.77-0.89)	0.81 (0.73-0.87)	0.85 (0.78-0.89)
Sweets	0.72 (0.64-0.78)	0.72 (0.59-0.81)	0.71 (0.61-0.79)	0.71 (0.61-0.79)	0.74 (0.63-0.82)	0.71 (0.59-0.80)	0.71 (0.60-0.79)
Prepared foods	0.59 (0.49-0.67)	0.57 (0.40-0.70)	0.59 (0.46-0.69)	0.57 (0.44-0.69)	0.61 (0.46-0.73)	0.61 (0.46-0.73)	0.55 (0.41-0.66)
Beer	0.88 (0.85-0.91)	0.86 (0.78-0.90)	0.88 (0.84-0.92)	0.87 (0.82-0.91)	0.88 (0.81-0.92)	0.86 (0.80-0.91)	0.89 (0.84-0.92)
Average	0.67 (—[b])	0.68 (—)	0.67 (—)	0.67 (—)	0.68 (—)	0.70 (—)	0.65 (—)

[a]*P*<.001 for all data.

[b]—: not applicable.

Table 3. Spearman correlation coefficients derived from the electronic 12-hour dietary recall app versus the food frequency questionnaire for categories of food group consumption and strata: smoking: no and yes; physical activity (minutes/week): ≥150 and <150; and body mass index (kg/m^2): <25 and ≥25. Comparison of the electronic 12-hour dietary recall app versus food frequency questionnaire. Spearman correlation coefficients (95% CI).

Food groups	All[a]	Smoking		Physical activity (minutes/week)		Body mass index (kg/m^2)	
		No	Yes	≥150	<150	<25	≥25
Fruit	0.80 (0.75-0.84)	0.80 (0.74-0.85)	0.81 (0.64-0.90)	0.80 (0.72-0.85)	0.81 (0.71-0.88)	0.78 (0.70-0.84)	0.85 (0.77-0.91)
Vegetables	0.70 (0.62-0.76)	0.73 (0.65-0.79)	0.53[b] (0.22-0.74)	0.70 (0.60-0.77)	0.64 (0.48-0.76)	0.69 (0.59-0.77)	0.75 (0.62-0.84)
Legumes	0.48 (0.37-0.58)	0.49 (0.37-0.60)	0.41[b] (0.07-0.66)	0.46 (0.32-0.59)	0.52 (0.33-0.68)	0.47 (0.33-0.59)	0.51 (0.31-0.66)
Chicken/Turkey	0.58 (0.48-0.67)	0.63 (0.52-0.71)	0.38[b] (0.04-0.64)	0.57 (0.44-0.67)	0.61 (0.44-0.74)	0.61 (0.49-0.71)	0.54 (0.34-0.69)
Fish	0.53 (0.42-0.62)	0.52 (0.40-0.62)	0.59 (0.31-0.78)	0.52 (0.39-0.64)	0.51 (0.31-0.67)	0.58 (0.45-0.68)	0.42 (0.21-0.60)
Red meat	0.63 (0.53-0.70)	0.62 (0.52-0.71)	0.64 (0.38-0.81)	0.64 (0.53-0.73)	0.60 (0.42-0.73)	0.62 (0.50-0.71)	0.64 (0.47-0.76)
Soft drinks	0.83 (0.78-0.87)	0.83 (0.77-0.87)	0.85 (0.72-0.92)	0.84 (0.79-0.89)	0.80 (0.70-0.87)	0.82 (0.76-0.87)	0.83 (0.73-0.89)
Sweets	0.72 (0.64-0.78)	0.72 (0.64-0.79)	0.67 (0.42-0.82)	0.71 (0.62-0.79)	0.75 (0.62-0.84)	0.71 (0.61-0.78)	0.74 (0.61-0.83)
Prepared foods	0.59 (0.49-0.67)	0.62 (0.52-0.71)	0.44[b] (0.12-0.68)	0.58 (0.45-0.68)	0.59 (0.41-0.72)	0.58 (0.45-0.68)	0.61 (0.43-0.74)
Beer	0.88 (0.85-0.91)	0.87 (0.83-0.91)	0.88 (0.77-0.94)	0.91 (0.88-0.94)	0.82 (0.73-0.89)	0.91 (0.87-0.93)	0.85 (0.76-0.90)
Average	0.67 (—[c])	0.68 (—)	0.62 (—)	0.67 (—)	0.67 (—)	0.68 (—)	0.67 (—)

[a]*P*<.001 for all data, except:

[b]*P*<.05.

[c]—: not applicable.

Table 4. Cross-classification analysis derived from the electronic 12-hour dietary recall app versus the food frequency questionnaire for categories of food group consumption and strata: age (years): <25 and ≥25; gender: females and males; and occupation: students and employees. Comparison of the electronic 12-hour dietary recall app versus food frequency questionnaire.

Agreement	All[a]	Age (years)		Gender		Occupation	
		<25	≥25	Females	Males	Students	Employees
Exact agreement[b] (%)	50.7	50.9	50.7	53.6	47.0	51.5	50.2
Exact agreement + adjacent[c] (%)	90.2	90.7	90.0	90.9	89.5	91.8	89.2
Extreme disagreement[d] (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[a]Data presented are mean agreement for 10 different food groups: fruit, vegetables, legumes, chicken/turkey, fish, red meat, soft drinks, sweets, prepared foods, and beer.

[b]Exact agreement: cases cross-classified into the same category.

[c]Exact agreement + adjacent: cases cross-classified into the same or adjacent category.

[d]Extreme disagreement: cases cross-classified into extreme categories.

Table 5. Cross-classification analysis derived from the electronic 12-hour dietary recall app versus the food frequency questionnaire for categories of food group consumption and strata: smoking: no and yes; physical activity (minutes/week): ≥150 and <150; and body mass index (kg/m^2): <25 and ≥25. Comparison of the electronic 12-hour dietary recall app versus the food frequency questionnaire.

Agreement	All[a]	Smoking		Physical activity (minutes/week)		Body mass index (kg/m^2)	
		No	Yes	≥150	<150	<25	≥25
Exact agreement[b] (%)	50.7	50.9	49.7	51.4	49.4	51.5	49.6
Exact agreement + adjacent[c] (%)	90.2	90.4	90.0	90.4	90.0	91.3	88.6
Extreme disagreement[d] (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[a]Data presented are mean agreement for 10 different food groups: fruit, vegetables, legumes, chicken/turkey, fish, red meat, soft drinks, sweets, prepared foods, and beer.

[b]Exact agreement: cases cross-classified into the same category.

[c]Exact agreement + adjacent: cases cross-classified into the same or adjacent category.

[d]Extreme disagreement: cases cross-classified into extreme categories.

Table 6. Weighted kappa derived from the electronic 12-hour dietary recall app versus the food frequency questionnaire for categories of food group consumption and strata: age (years): <25 and ≥25; gender: females and males; and occupation: students and employees. Comparison of the electronic 12-hour dietary recall app versus the food frequency questionnaire.

Food groups	All[a]	Age (years)		Gender		Occupation	
		<25	≥25	Females	Males	Students	Employees
Fruit	0.65	0.64	0.64	0.63	0.67	0.66	0.63
Vegetables	0.54	0.59	0.48	0.59	0.48	0.61	0.46
Legumes	0.37	0.35	0.38	0.38	0.34	0.38	0.37
Chicken/Turkey	0.40	0.32	0.44	0.34	0.46	0.34	0.44
Fish	0.30	0.37	0.26	0.29	0.32	0.41	0.22
Red meat	0.46	0.55	0.38	0.48	0.37	0.54	0.39
Soft drinks	0.59	0.52	0.62	0.60	0.58	0.59	0.60
Sweets	0.49	0.48	0.48	0.51	0.45	0.48	0.47
Prepared foods	0.40	0.38	0.40	0.40	0.40	0.41	0.36
Beer	0.68	0.67	0.66	0.64	0.67	0.66	0.67
Average	0.49	0.49	0.48	0.49	0.48	0.51	0.46

[a]P<.001 for all data.

The Electronic 12-Hour Dietary Recall App Versus the 4 Dietary Records

For all the food groups, for all participants, and for all strata, the average SCC was 0.65 (by strata, from 0.57 [smokers] to 0.67 [males]; Tables 8 and 9). Cross-classification analysis showed that the average percentage of individuals classified in the *exact agreement* category was 50.0% (by strata, from 47.0% [males] to 52.3% [females]); *exact agreement+adjacent* was 88.2% (by strata, from 86.1% [males] to 89.8% [females]); and no participants (0%) were classified in the *extreme disagreement* category (Tables 10 and 11; see Multimedia Appendix 5 for full details). The average weighted kappa was .50 (by strata, from 0.44 [smokers] to 0.50 [≥25 years, males, employees, non-smokers, ≥150 min/week, and BMI<25]; Tables 12 and 13).

Table 7. Weighted kappa derived from the electronic 12-hour dietary recall app versus the food frequency questionnaire for categories of food group consumption and strata: smoking: no and yes; physical activity (minutes/week): ≥150 and <150; and body mass index (kg/m^2): <25 and ≥25. Comparison of the electronic 12-hour dietary recall app versus the food frequency questionnaire.

Food groups	All[a]	Smoking		Physical activity (minutes/week)		Body mass index (kg/m^2)	
		No	Yes	≥150	<150	<25	≥25
Fruit	0.65	0.65	0.65	0.63	0.67	0.63	0.68
Vegetables	0.54	0.56	0.43[b]	0.56	0.48	0.54	0.55
Legumes	0.37	0.37	0.33[b]	0.37	0.35	0.35	0.40
Chicken/Turkey	0.40	0.43	0.26[b]	0.39	0.41	0.45	0.31
Fish	0.30	0.30	0.34[b]	0.30	0.30	0.34	0.23[b]
Red meat	0.46	0.46	0.45	0.44	0.49	0.46	0.43
Soft drinks	0.59	0.59	0.60	0.62	0.55	0.60	0.55
Sweets	0.49	0.49	0.44	0.51	0.45	0.48	0.49
Prepared foods	0.40	0.44	0.22[b]	0.41	0.37	0.39	0.44
Beer	0.68	0.67	0.66	0.71	0.61	0.70	0.63
Average	0.49	0.50	0.44	0.49	0.47	0.49	0.47

[a]*P*<.001 for all data, except:

[b]*P*<.05.

Table 8. Spearman correlation coefficients derived from the electronic 12-hour dietary recall app versus the 4 dietary records for categories of food group consumption and strata: age (years): <25 and ≥25; gender: females and males; and occupation: students and employees. Comparison of the electronic 12-hour dietary recall app versus dietary records. Spearman correlation coefficients (95% CI).

Food groups	All[a]	Age (years)		Gender		Occupation	
		<25	≥25	Females	Males	Students	Employees
Fruit	0.78 (0.73-0.83)	0.82 (0.74-0.88)	0.74 (0.65-0.81)	0.80 (0.72-0.86)	0.76 (0.65-0.83)	0.83 (0.75-0.88)	0.74 (0.64-0.81)
Vegetables	0.72 (0.65-0.78)	0.74 (0.63-0.83)	0.68 (0.56-0.76)	0.70 (0.60-0.78)	0.75 (0.64-0.83)	0.72 (0.60-0.81)	0.69 (0.58-0.77)
Legumes	0.40 (0.28-0.51)	0.40 (0.20-0.56)	0.43 (0.27-0.56)	0.34 (0.17-0.50)	0.47 (0.29-0.62)	0.46 (0.27-0.61)	0.39 (0.22-0.53)
Chicken/Turkey	0.63 (0.54-0.71)	0.59 (0.42-0.71)	0.67 (0.56-0.76)	0.63 (0.51-0.73)	0.60 (0.44-0.72)	0.60 (0.45-0.72)	0.66 (0.54-0.75)
Fish	0.52 (0.41-0.61)	0.53 (0.35-0.67)	0.49 (0.34-0.62)	0.50 (0.35-0.63)	0.55 (0.38-0.68)	0.47 (0.28-0.62)	0.51 (0.36-0.63)
Red meat	0.52 (0.42-0.62)	0.50 (0.32-0.65)	0.54 (0.40-0.66)	0.45 (0.29-0.59)	0.55 (0.38-0.68)	0.50 (0.32-0.64)	0.54(0.39-0.65)
Soft drinks	0.77 (0.71-0.82)	0.70 (0.57-0.79)	0.82 (0.75-0.87)	0.73 (0.64-0.81)	0.82 (0.74-0.88)	0.72 (0.60-0.81)	0.81 (0.74-0.86)
Sweets	0.72 (0.64-0.78)	0.78 (0.68-0.85)	0.68 (0.57-0.77)	0.71 (0.61-0.79)	0.74 (0.63-0.82)	0.74 (0.63-0.82)	0.70 (0.60-0.79)
Prepared foods	0.63 (0.53-0.70)	0.57 (0.40-0.70)	0.66 (0.55-0.75)	0.62 (0.49-0.72)	0.67 (0.53-0.77)	0.60 (0.45-0.72)	0.63 (0.51-0.73)
Beer	0.81 (0.75-0.85)	0.71 (0.59-0.80)	0.84 (0.77-0.88)	0.83 (0.76-0.88)	0.77 (0.67-0.84)	0.70 (0.58-0.80)	0.85 (0.78-0.89)
Average	0.65 (—[b])	0.63 (—)	0.65 (—)	0.63 (—)	0.67 (—)	0.63 (—)	0.65 (—)

[a]*P*<.001 for all data.

[b]—: not applicable.

Table 9. Spearman correlation coefficients derived from the electronic 12-hour dietary recall app versus the 4 dietary records for categories of food group consumption and strata: smoking: no and yes; physical activity (minutes/week): ≥150 and <150; and body mass index (kg/m^2): <25 and ≥25. Comparison of the electronic 12-hour dietary recall app versus dietary records. Spearman correlation coefficients (95% CI).

Food groups	All[a]	Smoking		Physical activity (minutes/week)		Body mass index (kg/m^2)	
		No	Yes	≥150	<150	<25	≥25
Fruit	0.78 (0.73-0.83)	0.79 (0.72-0.84)	0.72 (0.50-0.85)	0.76 (0.68-0.83)	0.82 (0.72-0.88)	0.75 (0.67-0.82)	0.82 (0.73-0.89)
Vegetables	0.72 (0.65-0.78)	0.74 (0.66-0.80)	0.61 (0.34-0.79)	0.72 (0.63-0.79)	0.69 (0.54-0.80)	0.70 (0.60-0.77)	0.78 (0.66-0.86)
Legumes	0.40 (0.28-0.51)	0.42 (0.29-0.53)	0.29[c] (0.00-0.57)	0.40 (0.25-0.54)	0.39[b] (0.17-0.58)	0.38 (0.23-0.52)	0.43 (0.21-0.60)
Chicken/Turkey	0.63 (0.54-0.71)	0.64 (0.54-0.72)	0.54[b] (0.25-0.75)	0.60 (0.48-0.70)	0.70 (0.55-0.80)	0.63 (0.52-0.72)	0.63 (0.46-0.76)
Fish	0.52 (0.41-0.61)	0.51 (0.38-0.61)	0.57[b] (0.28-0.76)	0.54 (0.41-0.65)	0.46 (0.25-0.63)	0.63 (0.51-0.72)	0.33[b] (0.10-0.53)
Red meat	0.52 (0.42-0.62)	0.55 (0.44-0.65)	0.31[d] (0.00-0.59)	0.59 (0.47-0.69)	0.39[b] (0.17-0.58)	0.52 (0.39-0.64)	0.57 (0.38-0.71)
Soft drinks	0.77 (0.71-0.82)	0.77 (0.71-0.83)	0.75 (0.54-0.87)	0.79 (0.72-0.85)	0.73 (0.59-0.82)	0.77 (0.69-0.83)	0.75 (0.62-0.84)
Sweets	0.72 (0.64-0.78)	0.72 (0.64-0.79)	0.54[b] (0.24-0.74)	0.72 (0.63-0.79)	0.72 (0.59-0.82)	0.73 (0.64-0.80)	0.70 (0.56-0.80)
Prepared foods	0.63 (0.53-0.70)	0.63 (0.53-0.71)	0.64 (0.38-0.81)	0.61 (0.49-0.70)	0.67 (0.51-0.78)	0.59 (0.47-0.69)	0.68 (0.53-0.79)
Beer	0.81 (0.75-0.85)	0.81 (0.74-0.85)	0.78 (0.59-0.89)	0.80 (0.73-0.85)	0.82 (0.73-0.89)	0.81 (0.74-0.86)	0.82 (0.72-0.88)
Average	0.65 (—[e])	0.66 (—)	0.57 (—)	0.65 (—)	0.64 (—)	0.65 (—)	0.65 (—)

[a]*P*<.001 for all data, except:

[b]*P*<.05.

[c]*P*=.106.

[d]*P*=.083.

[e]—: not applicable.

Table 10. Cross-classification analysis derived from the electronic 12-hour dietary recall app versus the 4 dietary records for categories of food group consumption and strata: age (years): <25 and ≥25; gender: females and males; and occupation: students and employees. Comparison of the electronic 12-hour dietary recall app versus dietary records.

Agreement	All[a]	Age (years)		Gender		Occupation	
		<25	≥25	Females	Males	Students	Employees
Exact agreement[b] (%)	50.0	47.7	51.6	52.3	47.0	47.1	52.2
Exact agreement + adjacent[c] (%)	88.2	89.3	87.5	89.8	86.1	89.2	87.5
Extreme disagreement[d] (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[a]Data presented are mean agreement for 10 different food groups: fruit, vegetables, legumes, chicken/turkey, fish, red meat, soft drinks, sweets, prepared foods, and beer.

[b]Exact agreement: cases cross-classified into the same category.

[c]Exact agreement + adjacent: cases cross-classified into the same or adjacent category.

[d]Extreme disagreement: cases cross-classified into extreme categories.

Table 11. Cross-classification analysis derived from the electronic 12-hour dietary recall app versus the 4 dietary records for categories of food group consumption and strata: smoking: no and yes; physical activity (minutes/week): ≥150 and <150; and body mass index (kg/m^2): <25 and ≥25. Comparison of the electronic 12-hour dietary recall app versus dietary records.

Agreement	All[a]	Smoking		Physical activity (minutes/week)		Body mass index (kg/m^2)	
		No	Yes	≥150	<150	<25	≥25
Exact agreement[b] (%)	50.0	50.5	47.3	51.0	48.1	51.0	48.0
Exact agreement + adjacent[c] (%)	88.2	88.1	89.1	88.1	88.4	89.0	86.8
Extreme disagreement[d] (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[a]Data presented are mean agreement for 10 different food groups: fruit, vegetables, legumes, chicken/turkey, fish, red meat, soft drinks, sweets, prepared foods, and beer.

[b]Exact agreement: cases cross-classified into the same category.

[c]Exact agreement + adjacent: cases cross-classified into the same or adjacent category.

[d]Extreme disagreement: cases cross-classified into extreme categories.

Table 12. Weighted kappa derived from the electronic 12-hour dietary recall app versus the 4 dietary records for categories of food group consumption and strata: age (years): <25 and ≥25; gender: females and males; and occupation: students and employees. Comparison of the electronic 12-hour dietary recall app versus dietary records.

Food groups	All[a]	Age (years)		Gender		Occupation	
		<25	≥25	Females	Males	Students	Employees
Fruit	0.67	0.73	0.61	0.69	0.64	0.72	0.61
Vegetables	0.56	0.54	0.53	0.56	0.55	0.51	0.55
Legumes	0.29	0.28[b]	0.30	0.21[b]	0.38	0.34	0.26
Chicken/Turkey	0.43	0.36	0.47	0.42	0.41	0.36	0.47
Fish	0.31	0.29	0.29	0.29	0.33	0.26	0.31
Red meat	0.34	0.34	0.35	0.31	0.32	0.34	0.34
Soft drinks	0.64	0.58	0.67	0.64	0.63	0.60	0.66
Sweets	0.54	0.52	0.56	0.55	0.53	0.49	0.58
Prepared foods	0.47	0.41	0.50	0.43	0.52	0.44	0.47
Beer	0.72	0.64	0.74	0.77	0.65	0.63	0.76
Average	0.50	0.47	0.50	0.49	0.50	0.47	0.50

[a]$P<.001$ for all data, except:

[b]$P<.05$.

Usability Rating Questionnaire for the Electronic 12-Hour Dietary Recall App

The responses of the participants to the usability-rating questionnaire are shown in Tables 14 and 15.

Table 13. Weighted kappa derived from the electronic 12-hour dietary recall app versus the 4 dietary records for categories of food group consumption and strata: smoking: no and yes; physical activity (minutes/week): ≥150 and <150; and body mass index (kg/m^2): <25 and ≥25. Comparison of the electronic 12-hour dietary recall app versus dietary records.

Food groups	All[a]	Smoking		Physical activity (minutes/week)		Body mass index (kg/m^2)	
		No	Yes	≥150	<150	<25	≥25
Fruit	0.67	0.68	0.60	0.63	0.74	0.66	0.68
Vegetables	0.56	0.58	0.46	0.59	0.48	0.53	0.61
Legumes	0.29	0.31	0.18[c]	0.32	0.22[b]	0.26	0.34
Chicken/Turkey	0.43	0.43	0.42[b]	0.42	0.43	0.45	0.39
Fish	0.31	0.30	0.34[b]	0.32	0.25[b]	0.39	0.16[b]
Red meat	0.34	0.36	0.19[b]	0.39	0.24[b]	0.35	0.33
Soft drinks	0.64	0.63	0.70	0.64	0.64	0.66	0.58
Sweets	0.54	0.54	0.50	0.54	0.54	0.55	0.54
Prepared foods	0.47	0.48	0.42	0.47	0.47	0.43	0.55
Beer	0.72	0.74	0.63	0.72	0.73	0.71	0.74
Average	0.50	0.50	0.44	0.50	0.47	0.50	0.49

[a]$P<.001$ for all data, except:
[b]$P<.05$
[c]$P=.0624$.

Table 14. Responses of the study participants to the usability rating questionnaire for the electronic 12-hour dietary recall app (part 1).

Options	Questions, n (%)			
	Easy to complete	Too time consuming	Interesting to complete	I would be willing to complete again
Strongly agree	139 (68.5)	3 (1.5)	48 (23.6)	57 (28.1)
Agree	62 (30.5)	4 (2.0)	111 (54.7)	104 (51.2)
Neither agree nor disagree	2 (1.0)	4 (2.0)	39 (19.2)	39 (19.2)
Disagree	0 (0.0)	82 (40.4)	5 (2.5)	3 (1.5)
Strongly disagree	0 (0.0)	110 (54.2)	0 (0.0)	0 (0.0)

Table 15. Responses of the study participants to the usability rating questionnaire for the electronic 12-hour dietary recall app (part 2).

Options	Question, n (%)
	Time to complete the app
<1 min/day	23 (11.3)
Approximately 1 min/day	54 (26.6)
Approximately 2 min/day	63 (31.0)
Approximately 3 min/day	41 (20.2)
Approximately 4 min/day	16 (7.9)
5 min/day or more	6 (3.0)

Discussion

Overview

The current-day recall has been designed to categorize participants according to habitual intake of selected food groups. Notwithstanding, this method is not intended to determine the total amount of foods consumed by an individual nor the exact quantity consumed for specific food groups or nutrients. This method is basically a modified 24-hour recall focused on a series of 10 food groups and completed at the end of every day during 28 consecutive days [32,44,45]. In this study, the current-day recall, based on the e-12HR app, has been compared with 2 different reference models, one long term (FFQ) and the other

short-term (4 DRs), in the whole sample of adults and in different strata thereof (sociodemographic characteristics, lifestyle factors, and weight category).

Even though 2 different reference methods were used for e-12HR, the high degree of association and agreement between the data collected when comparing the different methods does not indicate that the current-day recall is exact, as there is no true measurement of dietary intake [2,8,32,52].

Principal Findings: The Electronic 12-Hour Recall App Versus the Food Frequency Questionnaire and the 4 Dietary Records

For each of the 10 food groups considered in this study, a comparison was made using e-12HR versus FFQ as well as e-12HR versus DRs. In both comparisons, 5 criteria were considered to compare the different methods: SCC; cross-classification percentage in the exact agreement category, in the exact agreement+adjacent category, and in the extreme disagreement category; and weighted kappa. Apart from this, and in the comparisons, the complete sample and the 12 individual strata were compared. This generated 130 statistical indicators for each of the food groups. For example, for fruit, a statistical indicator was obtained for each of the comparison criteria (5 comparison criteria), for the complete sample and the different strata (13 strata), for e-12HR versus FFQ comparison (5×13=65 indicators), and finally for e-12HR versus DRs (5×13=65 indicators). The 130 statistical indicators obtained for each food group fulfilled the comparison criteria (see the *Statistical Analysis* section) for fruit, vegetables, soft drinks, sweets, and beer. For the rest of the food groups, of the 130 statistical indicators obtained for each, the following cases did not fulfill the comparison criteria: legumes, 36.1% (47/130); chicken/turkey, 13.1% (17/130); fish, 30.0% (39/130); red meat, 16.1% (21/130); and prepared foods, 7.7% (10/130).

Regarding the SCC, in all of these cases, the agreement between methods was moderate (r=±0.30), except in e-12HR versus RDs, for legumes, and yes smoking strata (0.29). Regarding the cross-classification percentage in the exact agreement category, in all cases the percentage of agreement between the methods was at least 31%, except for the e-12HR app versus FFQ for chicken/turkey and yes smoking strata (27.3%), as well as e-12HR versus DRs for fish and ≥25 years strata (24.6%). Regarding the weighted kappa, in all cases the agreement between the methods was fair (weighted kappa statistic index between 0.40 and 0.21), except in e-12HR versus RDs, for legumes and the yes smoking strata (0.18); for fish, ≥25 kg/m^2 strata (0.16); and for red meat, yes smoking strata (0.19; see Multimedia Appendix 6).

Evaluating the true validity of a method requires measuring, with a high degree of accuracy, the habitual diet of free-living individuals during a prolonged period, which is not feasible [4]. As a result, the researchers of this study have evaluated the relative validity of e-12HR by comparing it with 2 alternative methods of dietary assessment (FFQ and DRs), with their own limitations (there is no perfect measure of dietary intake, which implies that validation studies are not possible) [2,3,6,16-19]. Thus, validation studies never compare an operational method with absolute truth. To do so, the lesser degree of agreement between e-12HR and the reference methods for some food groups (especially legumes and fish) does not imply that e-12HR is a bad categorization method for habitual dietary intake for these food groups. The current-day recall is a method that depends only on short-term memory (e-12HR app is completed at the end of each day); it takes day-to-day intrapersonal variation into account during the period of the study (the app is completed daily). At the same time, the FFQ compiles information at the end of the study period, DR only on 4 of the 28 days of the study period. With regard to the FFQs, we must take into account the fact that the recollection of past consumption of foods can be influenced by more recent food consumption [2,6,17,18]. Regarding the DRs, short-term methods are generally unrepresentative of habitual intake if only one or a few days are assessed [2]. The different characteristics of e-12HR, the FFQ, and DRs can contribute to assigning different categories of habitual consumption depending on the method, especially for those food groups that are consumed infrequently, such as, legumes, and fish. In any case, the research team will develop future studies to explore the reasons for the disagreement between the methods for these 5 food groups.

The majority of the published research reports associations between the methods, measured by correlation coefficients, although agreement is the most appropriate comparison for validation studies [8]. As previously mentioned in the *Statistical Analysis* subsection, the different categorization of individuals according to the number of categories considered would affect the cross-classification analysis and weighted kappa. With regard to the cross-classification analysis, the dependence on the number of categories considered is reduced as a result of the comparison considered by Forster et al [13] and Fallaize et al [14]; these studies used 4 categories. However, the weighted kappa suffered, especially due to the comparison criterion considered by Masson et al [4] for weighted kappa being defined for 3 categories instead of the 6 considered here. The 6 original categories could have been reorganized into 3 [4,8,44], 4 [13,14], or 5 [30,53,54], as other authors have done. However, this research team preferred to maintain 6 categories for the statistical analysis [32,45] as a greater number of categories of habitual consumption provides compact information on the ability of the methods to assign individuals according to the distribution of dietary intake [30]. In any case, the values observed indicate high correlation and good agreement between the e-12HR app and the 2 reference methods, in the whole sample and in all strata considered: age group (<25 years old and ≥25 years old), gender (female and male), occupation (student and employee), smoking status (no and yes), physical activity status (≥150 min/week and <150 min/week), and BMI (<25 kg/m^2 and ≥25 kg/m^2).

e-12HR presents interesting characteristics for both participants and investigators. For participants, the app is easy, brief, and interesting to complete (according to the usability-rating questionnaire). For investigators, with the e-12HR app, data collection is performed digitally, eliminating the need for investigators to later introduce the data manually; it is a self-reporting tool, not requiring interviewers; and overall

research costs are greatly reduced. Notwithstanding, current-day recall presents some weaknesses when determining the category of habitual consumption; although the method only depends on short-term memory, it still depends on the memory of the participant (as e-12HR is not completed immediately after each meal, rather at the end of each day), and the number of different options for servings consumed is limited (with color photographs that represent 2 to 4 possible options). Regarding the use of photographs, when comparing e-12HR with the FFQ, the values obtained are similar to those from a previous study that used only 1 reference method (a semiquantitative FFQ) and the older version of the app (which did not use photographs to facilitate estimation of the servings consumed) [45]. As such, the research team would like to mention that the introduction of photographs in the newer version of the app has not translated into better correlation or agreement data between these 2 methods.

The 3 methods for determining diet refer to the same tracking period to avoid possible variations in the intake of different foods over time [13,22,28,46]. This is especially likely among the university students who make up the sample of this study. Reasons being that dietary intake is variable from day to day, sporadic changes in food intake are common (skipping meals, snacking, school events interfering with meal times), and dining out is more frequent than in the general population [44]. All these reasons could have led to an underestimation of the correlation and agreement between the different methods that were compared. In contrast, using the same period of time could overestimate the correlation and agreement between the different methods compared. There are no bibliographic references from other authors that allow us to evaluate this overestimation of such a new method as current-day recall. In the comparison of e-12HR versus the FFQ, the app was completed daily over 28 consecutive days, and the FFQ was completed after the end of period of app use. It is unlikely that the participant would be able to remember the information collected in the app during the 28-day period and that this reminiscence facilitates completing the FFQ, and overestimating the correlation and agreement between both methods. In the comparison of e-12HR versus DRs, the app is completed during 28 days, and on 4 of them, a DR is completed. On the days on which the participants complete both methods, remembering the answers to the DR will favor completion of the app; however, this only occurs during 4 of the 28 days of the study period, and as such, overestimation of the correlation and agreement of both methods is unlikely to be significant.

Nutrient Intake

It must be reiterated that current-day recall was not designed to determine the exact quantity of specific nutrients consumed. Good agreement between e-12HR and the reference methods (FFQ and DRs), with regard to a group of specific foods, does not imply good agreement between the nutrients that the food group provides an individual. This is due to the fact that specific nutrients may come from different food groups. For example, of the food groups considered in this study, legumes, chicken/turkey, fish, and red meat are all rich in proteins. Although, in general, we have observed good agreement between e-12HR and the reference methods considered by these food groups, this does not imply that e-12HR has the ability to determine the exact quantity of proteins consumed by an individual. This is because of other food groups also being rich in this nutrient (such as nuts, dairy products, or pasta), which were not considered in this study.

Format Used in Questionnaire

Full details on the format used in the questionnaires are available elsewhere [32,44,45]. In short, the e-12HR app is digital, and the FFQ and DRs are completed on paper. Paper formats are typically associated with errors such as unanswered questions, questions with multiple responses [7] (FFQ), and not registering the quantity consumed for some of the different food groups selected (DR) [55]. Despite the potential advantages of utilizing FFQs and DRs in digital format, in the end, it was decided to use paper formats in this study. The research team took into account that, on the one hand, evidence shows that data collected from smartphone apps and Web-based FFQs and DRs are comparable with data from paper formats [12-14,16,22,29,30,35,36,38,40,56], whereas on the other hand, due to the characteristics of this study, the potential disadvantages of developing FFQs and DRs in digital format could surpass the possible benefits. In fact, in this study, the paper-based FFQ and the DRs are very short and simple (they only contain 10 food groups), and the sample population is made up of students and employees at the Schools of Medicine and Pharmacy at the University of Seville, which is easily accessible for the research team. The simplicity of the paper-based FFQ and DRs minimized possible errors, the amount of paper used, problems with storage space, and costs associated with data conversion. These costs were minimal when compared with the potential costs of developing a Web-based or smartphone-based FFQ and DRs. Easy access to the sample made it possible to complete the paper FFQ in person, without the need for researchers or participants to travel or pay mailing costs [32,44,45].

Usability Rating Questionnaire for the Electronic 12-Hour Dietary Recall App

The majority of participants in this study reported that the e-12HR app was easy, brief, and interesting to complete; that they would be willing to complete the e-12HR app again; and that the task took 2 min or less per day to complete (see Tables 14 and 15). According to this latest piece of information from the study participants, the research team considered that the time necessary to complete the app is, normally, 2 min per day or less.

Sample Size

The sample size was established with the sample size software nQuery Advisor Version 7.0 (Statistical Solutions Ltd., One International Place, 100 Oliver Street, Boston MA, USA) [57]. For the SCC, alpha=.05, a value for the null hypothesis (lack of relation)=0.0, a value for the alternative hypothesis=0.5, and power of 90%. The sample size obtained was n=40.

The sample size reached (and amply surpassed) what was indicated in all of the strata except for one: *yes smoking* status with n=33 (see Table 1).

Limitations

Limitations of this study included the fact that the sample used was extremely educated, which is a convenient sample (there is no random selection) and not representative of the population on the national level. In addition, as this is a convenience sample, made up of colleagues, students, and employees, the participants might have responded more favorably to the questions posed by the usability rating questionnaire for e-12HR. The small number of individuals in some of the subgroups is another limitation of the study, for example, smokers (n=33). Another limitation derives from the need to have a smartphone with an Android operating system. Access to these technologies is not universal and could exclude those students or employees with less purchasing power [44].

This method, as it was not designed to collect data on the exact quantity of specific nutrients consumed, does not allow for an analysis of the possible association between nutrients and chronic illnesses, rather only between categories of habitual consumption for food groups and risk of chronic illnesses.

Another limitation is that the soft drinks category does not differentiate between sugary drinks and artificially sugary drinks.

Ideally, validation studies should include the use of nutritional biomarkers, but currently, there are few biomarkers for specific foods [10,52,58,59] and they cannot measure habitual intake [52].

Conclusions

For the whole sample of adults and for all strata thereof, the high correlation and good agreement between the e-12HR app and both reference methods (the FFQ and the 4 DRs), utilizing various procedures of statistical analysis, indicate the *relative validity* of the current-day recall for ranking the habitual intake of selected food groups.

For e-12HR versus FFQ, the inclusion of photographs to facilitate estimation of the servings consumed has not provided better correlation or agreement data between the methods, as the data obtained were similar to that of a previous study using an older version of the app without photographs.

The *relative validity* of current-day recall and the interesting features of e-12HR for users (the app is easy, brief, and interesting to complete [according to the usability rating questionnaire], and has photographs to assist with estimating servings consumed) as well as investigators (data collection is performed digitally, eliminating the need for investigators to later introduce the data manually; it is a self-reporting tool, not requiring interviewers; and overall research costs are greatly reduced), indicate that this method could be considered as a useful alternative to FFQs. This method (FFQ) is the most commonly implemented instrument in large-scale epidemiological and intervention studies, which do not require determining the complete diet nor the exact quantity consumed of a specific food group to analyze possible associations with risks for chronic diseases and for evaluating the effects of interventions.

Acknowledgments

The authors would like to thank the participants in this study.

Authors' Contributions

LMB performed the conception and design of the study, developed the app, analyzed and interpreted data, and wrote the paper. OAR, MDGP, and EVL were involved in data collection and interpretation of the data and contributed in drafting the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Usability rating questionnaire for e-12HR.

[PDF File (Adobe PDF File), 6KB - mhealth_v7i4e11531_app1.pdf]

Multimedia Appendix 2

Questionnaire used in the e-12HR app.

[PDF File (Adobe PDF File), 83KB - mhealth_v7i4e11531_app2.pdf]

Multimedia Appendix 3

Screen capture of e-12HR.

[PDF File (Adobe PDF File), 247KB - mhealth_v7i4e11531_app3.pdf]

Multimedia Appendix 4

Precoded dietary record.

[PDF File (Adobe PDF File), 26KB - mhealth_v7i4e11531_app4.pdf]

Multimedia Appendix 5

Cross-classification analysis derived from the e-12HR app versus the food frequency questionnaire, and the e-12HR app versus the 4 dietary records.

[PDF File (Adobe PDF File), 292KB - mhealth_v7i4e11531_app5.pdf]

Multimedia Appendix 6

Cases that did not fulfill the comparison criteria for categories of food group consumption derived from the e-12HR app versus the food frequency questionnaire, and from the e-12HR app versus the 4 dietary records.

[PDF File (Adobe PDF File), 211KB - mhealth_v7i4e11531_app6.pdf]

References

1. Carroll RJ, Midthune D, Subar AF, Shumakovich M, Freedman LS, Thompson FE, et al. Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. Am J Epidemiol 2012 Feb 15;175(4):340-347 [FREE Full text] [doi: 10.1093/aje/kwr317] [Medline: 22273536]
2. Willett W. Nutritional Epidemiology, Third Edition. New York, MD: Oxford University Press; 2013.
3. Stumbo PJ. New technology in dietary assessment: a review of digital methods in improving food record accuracy. Proc Nutr Soc 2013 Feb;72(1):70-76. [doi: 10.1017/S0029665112002911] [Medline: 23336561]
4. Masson LF, McNeill G, Tomany JO, Simpson JA, Peace HS, Wei L, et al. Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. Public Health Nutr 2003 May;6(3):313-321. [doi: 10.1079/PHN2002429] [Medline: 12740081]
5. Giabbanelli PJ, Adams J. Identifying small groups of foods that can predict achievement of key dietary recommendations: data mining of the UK National Diet and Nutrition Survey, 2008-12. Public Health Nutr 2016 Dec;19(9):1543-1551 [FREE Full text] [doi: 10.1017/S1368980016000185] [Medline: 26879185]
6. Rosso N, Giabbanelli P. Accurately inferring compliance to five major food guidelines through simplified surveys: applying data mining to the UK National Diet and Nutrition Survey. JMIR Public Health Surveill 2018 May 30;4(2):e56 [FREE Full text] [doi: 10.2196/publichealth.9536] [Medline: 29848474]
7. Rutishauser IH. Dietary intake measurements. Public Health Nutr 2005 Oct;8(7A):1100-1107. [doi: 10.1079/PHN2005798] [Medline: 16277822]
8. Kristal AR, Kolar AS, Fisher JL, Plascak JJ, Stumbo PJ, Weiss R, et al. Evaluation of web-based, self-administered, graphical food frequency questionnaire. J Acad Nutr Diet 2014 Apr;114(4):613-621 [FREE Full text] [doi: 10.1016/j.jand.2013.11.017] [Medline: 24462267]
9. Cleghorn CL, Harrison RA, Ransley JK, Wilkinson S, Thomas J, Cade JE. Can a dietary quality score derived from a short-form FFQ assess dietary quality in UK adult population surveys? Public Health Nutr 2016 Nov;19(16):2915-2923. [doi: 10.1017/S1368980016001099] [Medline: 27181696]
10. Saeedi P, Skeaff SA, Wong JE, Skidmore PM. Reproducibility and relative validity of a short food frequency questionnaire in 9–10 year-old children. Nutrients 2016 May 7;8(5) [FREE Full text] [doi: 10.3390/nu8050271] [Medline: 27164137]
11. Buch-Andersen T, Pérez-Cueto FJ, Toft U. Relative validity and reproducibility of a parent-administered semi-quantitative FFQ for assessing food intake in Danish children aged 3-9 years. Public Health Nutr 2016 May;19(7):1184-1194. [doi: 10.1017/S136898001500275X] [Medline: 26434501]
12. Knudsen VK, Hatch EE, Cueto H, Tucker KL, Wise L, Christensen T, et al. Relative validity of a semi-quantitative, web-based FFQ used in the 'Snart Forældre' cohort - a Danish study of diet and fertility. Public Health Nutr 2016 Apr;19(6):1027-1034. [doi: 10.1017/S1368980015002189] [Medline: 26235206]
13. Medin AC, Carlsen MH, Andersen LF. Associations between reported intakes of carotenoid-rich foods and concentrations of carotenoids in plasma: a validation study of a web-based food recall for children and adolescents. Public Health Nutr 2016 Dec;19(18):3265-3275. [doi: 10.1017/S1368980016001622] [Medline: 27346293]
14. Forster H, Fallaize R, Gallagher C, O'Donovan CB, Woolhead C, Walsh MC, et al. Online dietary intake estimation: the Food4Me food frequency questionnaire. J Med Internet Res 2014;16(6):e150 [FREE Full text] [doi: 10.2196/jmir.3105] [Medline: 24911957]
15. Fallaize R, Forster H, Macready AL, Walsh MC, Mathers JC, Brennan L, et al. Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. J Med Internet Res 2014;16(8):e190 [FREE Full text] [doi: 10.2196/jmir.3355] [Medline: 25113936]

16. Riobó SP, Sierra PR, Soldo RJ, Gómez-Candela C, García LP, Serra-Majem L. Special considerations for nutritional studies in elderly. Nutr Hosp 2015 Feb 26;31(Suppl 3):84-90 [FREE Full text] [doi: 10.3305/nh.2015.31.sup3.8756] [Medline: 25719776]
17. Tucker KL, Smith CE, Lai C, Ordovas JM. Quantifying diet for nutrigenomic studies. Annu Rev Nutr 2013;33:349-371 [FREE Full text] [doi: 10.1146/annurev-nutr-072610-145203] [Medline: 23642200]
18. Gibson R. Principles of Nutritional Assessment, Second Edition. New York, Md: Oxford University Press; 2005.
19. Martín-Moreno JM, Gorgojo L. [Assessment of dietary intake at the population level through individual questionnaires: methodological shadows and lights]. Rev Esp Salud Publica 2007;81(5):507-518 [FREE Full text] [Medline: 18274354]
20. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TI, Energy Balance Measurement Working Group. Energy balance measurement: when something is not better than nothing. Int J Obes (Lond) 2014 Nov 13. [doi: 10.1038/ijo.2014.199] [Medline: 25394308]
21. Kim DW, Oh S, Kwon S, Kim J. Comparison of validity of food group intake by food frequency questionnaire between pre- and post- adjustment estimates derived from 2-day 24-hour recalls in combination with the probability of consumption. Asian Pac J Cancer Prev 2012;13(6):2655-2661 [FREE Full text] [Medline: 22938437]
22. Gerdessen JC, Souverein OW, van 't Veer P, de Vries JH. Optimising the selection of food items for FFQs using Mixed Integer Linear Programming. Public Health Nutr 2015 Jan;18(1):68-74. [doi: 10.1017/S1368980013003479] [Medline: 24476899]
23. Rangan AM, O'Connor S, Giannelli V, Yap ML, Tang LM, Roy R, et al. Electronic Dietary Intake Assessment (e-DIA): comparison of a mobile phone digital entry app for dietary data collection with 24-hour dietary recalls. JMIR Mhealth Uhealth 2015 Oct 27;3(4):e98 [FREE Full text] [doi: 10.2196/mhealth.4613] [Medline: 26508282]
24. Falomir Z, Arregui M, Madueño F, Corella D, Coltell O. Automation of food questionnaires in medical studies: a state-of-the-art review and future prospects. Comput Biol Med 2012 Oct;42(10):964-974. [doi: 10.1016/j.compbiomed.2012.07.008] [Medline: 22898338]
25. Christensen SE, Möller E, Bonn SE, Ploner A, Wright A, Sjölander A, et al. Two new meal- and web-based interactive food frequency questionnaires: validation of energy and macronutrient intake. J Med Internet Res 2013 Jun;15(6):e109 [FREE Full text] [doi: 10.2196/jmir.2458] [Medline: 23739995]
26. Swierk M, Williams PG, Wilcox J, Russell KG, Meyer BJ. Validation of an Australian electronic food frequency questionnaire to measure polyunsaturated fatty acid intake. Nutrition 2011 Jun;27(6):641-646. [doi: 10.1016/j.nut.2010.06.011] [Medline: 20869207]
27. Labonté M, Cyr A, Baril-Gravel L, Royer M, Lamarche B. Validity and reproducibility of a web-based, self-administered food frequency questionnaire. Eur J Clin Nutr 2012 Feb;66(2):166-173. [doi: 10.1038/ejcn.2011.163] [Medline: 21934698]
28. Svensson A, Larsson C. A mobile phone app for dietary intake assessment in adolescents: an evaluation study. JMIR Mhealth Uhealth 2015;3(4):e93 [FREE Full text] [doi: 10.2196/mhealth.4804] [Medline: 26534783]
29. Saloheimo T, González SA, Erkkola M, Milauskas DM, Meisel JD, Champagne CM, et al. The reliability and validity of a short food frequency questionnaire among 9-11-year olds: a multinational study on three middle-income and high-income countries. Int J Obes Suppl 2015 Dec;5(Suppl 2):S22-S28 [FREE Full text] [doi: 10.1038/ijosup.2015.15] [Medline: 27152180]
30. Tabacchi G, Filippi AR, Breda J, Censi L, Amodio E, Napoli G, et al. Comparative validity of the ASSO-Food Frequency Questionnaire for the web-based assessment of food and nutrients intake in adolescents. Food Nutr Res 2015;59:26216 [FREE Full text] [doi: 10.3402/fnr.v59.26216] [Medline: 25882537]
31. Alcantara I, Haardörfer R, Gazmararian JA, Hartman TJ, Greene B, Kegler MC. Relative validation of fruit and vegetable intake and fat intake among overweight and obese African-American women. Public Health Nutr 2015 Aug;18(11):1932-1940. [doi: 10.1017/S1368980014002547] [Medline: 25412867]
32. Béjar LM, Reyes Ó, García-Perea MD. Electronic 12-Hour Dietary Recall (e-12HR): comparison of a mobile phone app for dietary intake assessment with a food frequency questionnaire and four dietary records. JMIR Mhealth Uhealth 2018 Jun 15;6(6):e10409 [FREE Full text] [doi: 10.2196/10409] [Medline: 29907555]
33. Souverein OW, Dekkers AL, Geelen A, Haubrock J, de Vries JH, Ocké MC, et al. Comparing four methods to estimate usual intake distributions. Eur J Clin Nutr 2011 Jul;65(Suppl 1):S92-101. [doi: 10.1038/ejcn.2011.93] [Medline: 21731012]
34. Head of State. [State Agency Official State Gazette]. 1999. Law 15/1999, of Protection of Personal Data URL: https://www.boe.es/boe/dias/1999/12/14/pdfs/A43088-43099.pdf [accessed 2019-03-08] [WebCite Cache ID 76j8bnc6i]
35. Marshall SJ, Livingstone KM, Celis-Morales C, Forster H, Fallaize R, O'Donovan CB, Food4Me Study. Reproducibility of the online Food4Me food-frequency questionnaire for estimating dietary intakes across Europe. J Nutr 2016 Dec;146(5):1068-1075. [doi: 10.3945/jn.115.225078] [Medline: 27052541]
36. Medin AC, Astrup H, Kåsin BM, Andersen LF. Evaluation of a web-based food record for children using direct unobtrusive lunch observations: a validation study. J Med Internet Res 2015 Dec 7;17(12):e273 [FREE Full text] [doi: 10.2196/jmir.5031] [Medline: 26680744]
37. Hongu N, Pope BT, Bilgiç P, Orr BJ, Suzuki A, Kim AS, et al. Usability of a smartphone food picture app for assisting 24-hour dietary recall: a pilot study. Nutr Res Pract 2015 Apr;9(2):207-212 [FREE Full text] [doi: 10.4162/nrp.2015.9.2.207] [Medline: 25861429]

38. Henriksson H, Bonn SE, Bergström A, Bälter K, Bälter O, Delisle C, et al. A new mobile phone-based tool for assessing energy and certain food intakes in young children: a validation study. JMIR Mhealth Uhealth 2015 Apr 24;3(2):e38 [FREE Full text] [doi: 10.2196/mhealth.3670] [Medline: 25910494]
39. Daugherty BL, Schap TE, Ettienne-Gittens R, Zhu FM, Bosch M, Delp EJ, et al. Novel technologies for assessing dietary intake: evaluating the usability of a mobile telephone food record among adults and adolescents. J Med Internet Res 2012;14(2):e58 [FREE Full text] [doi: 10.2196/jmir.1967] [Medline: 22504018]
40. Casperson SL, Sieling J, Moon J, Johnson L, Roemmich JN, Whigham L. A mobile phone food record app to digitally capture dietary intake for adolescents in a free-living environment: usability study. JMIR Mhealth Uhealth 2015;3(1):e30 [FREE Full text] [doi: 10.2196/mhealth.3324] [Medline: 25775506]
41. Rodríguez IT, Ballart JF, Pastor GC, Jordà EB, Val VA. [Validation of a short questionnaire on frequency of dietary intake: reproducibility and validity]. Nutr Hosp 2008;23(3):242-252 [FREE Full text] [Medline: 18560701]
42. Vereecken CA, Maes L. A Belgian study on the reliability and relative validity of the Health Behaviour in School-Aged Children food-frequency questionnaire. Public Health Nutr 2003 Sep;6(6):581-588. [doi: 10.1079/PHN2003466] [Medline: 14690039]
43. Knudsen VK, Gille M, Nielsen TH, Christensen T, Fagt S, Biltoft-Jensen A. Relative validity of the pre-coded food diary used in the Danish National Survey of Diet and Physical Activity. Public Health Nutr 2011 Dec;14(12):2110-2116. [doi: 10.1017/S1368980011001650] [Medline: 21806861]
44. Bejar LM, Sharp BN, García-Perea MD. The e-EPIDEMIOLOGY mobile phone app for dietary intake assessment: comparison with a food frequency questionnaire. JMIR Res Protoc 2016 Nov 2;5(4):e208 [FREE Full text] [doi: 10.2196/resprot.5782] [Medline: 27806922]
45. Béjar LM. First evaluation steps of a new method for dietary intake estimation regarding a list of key food groups in adults and in different sociodemographic and health-related behaviour strata. Public Health Nutr 2017 Oct;20(15):2660-2669. [doi: 10.1017/S1368980017001641] [Medline: 28789723]
46. Watanabe M, Yamaoka K, Yokotsuka M, Adachi M, Tango T. Validity and reproducibility of the FFQ (FFQW82) for dietary assessment in female adolescents. Public Health Nutr 2011 Feb;14(2):297-305. [doi: 10.1017/S1368980010001618] [Medline: 20537215]
47. StataCorp L. Stata Statistical Software: Release 13 URL: https://www.stata.com/ [accessed 2018-07-07] [WebCite Cache ID 6xuS3EeeH]
48. Cohen J. Statistical Power Analysis for the Behavioral Sciences, Second Edition. New York: Lawrence Erlbaum Associates; 1988.
49. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005 May;37(5):360-363 [FREE Full text] [Medline: 15883903]
50. World Health Organization. 2001. Global Recommendations on Physical Activity for Health URL: http://www.who.int/dietphysicalactivity/physical-activity-recommendations-18-64years.pdf?ua=1 [accessed 2018-07-07] [WebCite Cache ID 70mNUz1FG]
51. World Health Organization. 1998. Programme of nutrition, family and reproductive health. Obesity. Preventing and managing the global epidemic URL: http://whqlibdoc.who.int/hq/1998/WHO_NUT_NCD_98.1_(p1-158).pdf [accessed 2018-07-07] [WebCite Cache ID 70mNailiT]
52. England CY, Thompson JL, Jago R, Cooper AR, Andrews RC. Development of a brief, reliable and valid diet assessment tool for impaired glucose tolerance and diabetes: the UK Diabetes and Diet Questionnaire. Public Health Nutr 2017 Feb;20(2):191-199 [FREE Full text] [doi: 10.1017/S1368980016002275] [Medline: 27609314]
53. Tabacchi G, Filippi AR, Amodio E, Jemni M, Bianco A, Firenze A, et al. A meta-analysis of the validity of FFQ targeted to adolescents. Public Health Nutr 2016 May;19(7):1168-1183. [doi: 10.1017/S1368980015002505] [Medline: 26354204]
54. Deschamps V, de Lauzon-Guillain B, Lafay L, Borys J, Charles MA, Romon M. Reproducibility and relative validity of a food-frequency questionnaire among French adults and adolescents. Eur J Clin Nutr 2009 Feb;63(2):282-291 [FREE Full text] [doi: 10.1038/sj.ejcn.1602914] [Medline: 17882132]
55. Vereecken CA, Rossi S, Giacchi MV, Maes L. Comparison of a short food-frequency questionnaire and derived indices with a seven-day diet record in Belgian and Italian children. Int J Public Health 2008;53(6):297-305. [doi: 10.1007/s00038-008-7101-6] [Medline: 19112592]
56. Matthys C, Pynaert I, De Keyzer W, De Henauw S. Validity and reproducibility of an adolescent web-based food frequency questionnaire. J Am Diet Assoc 2007 Apr;107(4):605-610. [doi: 10.1016/j.jada.2007.01.005] [Medline: 17383266]
57. nQuery Advisor Version 7 URL: https://www.statsols.com/nquery [accessed 2019-03-08] [WebCite Cache ID 76j9zpNvR]
58. Bassett JK, English DR, Fahey MT, Forbes AB, Gurrin LC, Simpson JA, et al. Validity and calibration of the FFQ used in the Melbourne Collaborative Cohort Study. Public Health Nutr 2016 Sep;19(13):2357-2368. [doi: 10.1017/S1368980016000690] [Medline: 27075344]
59. Hedrick VE, Dietrich AM, Estabrooks PA, Savla J, Serrano E, Davy BM. Dietary biomarkers: advances, limitations and future directions. Nutr J 2012 Dec 14;11:109 [FREE Full text] [doi: 10.1186/1475-2891-11-109] [Medline: 23237668]

Abbreviations

BMI: body mass index
DR: dietary record
e-12HR: electronic 12-hour dietary recall
FFQ: food frequency questionnaire
SCC: Spearman correlation coefficient

Edited by G Eysenbach; submitted 09.07.18; peer-reviewed by P Giabbanelli, JJ Mira, C Cleghorn; comments to author 08.10.18; revised version received 30.11.18; accepted 24.02.19; published 11.04.19.

Please cite as:
Béjar LM, García-Perea MD, Reyes ÓA, Vázquez-Limón E
Relative Validity of a Method Based on a Smartphone App (Electronic 12-Hour Dietary Recall) to Estimate Habitual Dietary Intake in Adults
JMIR Mhealth Uhealth 2019;7(4):e11531
URL: https://mhealth.jmir.org/2019/4/e11531/
doi:10.2196/11531
PMID:30973343

Original Paper

Improving Linkage to and Retention in Care in Newly Diagnosed HIV-Positive Patients Using Smartphones in South Africa: Randomized Controlled Trial

Willem Daniel Francois Venter[1], FCP(SA); Alex Fischer[1], BSc; Samanta Tresha Lalla-Edward[1], PhD; Jesse Coleman[1], PhD; Vincent Lau Chan[1†], MPH; Zara Shubber[2], PhD; Mothepane Phatsoane[1], BMedSci; Marelize Gorgens[2], MA; Lynsey Stewart-Isherwood[3], MSc; Sergio Carmona[3], PhD; Nicole Fraser-Hurt[2], PhD

[1]Wits Reproductive Health and HIV Institute, University of Witwatersrand, Johannesburg, South Africa

[2]World Bank, World Bank Group, Washington, DC, United States

[3]National Health Laboratory Service, Johannesburg, South Africa

[†]deceased

Corresponding Author:
Alex Fischer, BSc
Wits Reproductive Health and HIV Institute
University of Witwatersrand
Hillbrow Health Precinct
22 Esselen St
Johannesburg, 2001
South Africa
Phone: 27 011 358 5500
Email: afischer@wrhi.ac.za

Abstract

Background: South Africa provides free antiretroviral therapy for almost 5 million people living with HIV, but only 71% of the eligible people are on treatment, representing a shortfall in the care cascade, especially among men and youth. Many developing countries have expanded access to smartphones; success in health apps raises the possibility of improving this cascade.

Objective: SmartLink is a health app for Android smartphones providing HIV-related laboratory results, information, support, and appointment reminders to engage and link patients to care. This study aimed to evaluate the ability of SmartLink to improve linkage to care for HIV-positive smartphone owners.

Methods: This study was a multisite randomized controlled trial in Johannesburg. The intervention arm received the app (along with referral to a treatment site) and the control arm received the standard of care (referral alone). Linkage to care was confirmed by an HIV-related blood test reported on the National Health Laboratory Service database between 2 weeks and 8 months after initiation.

Results: A total of 345 participants were recruited into the study; 64.9% (224/345) of the participants were female and 44.1% (152/345) were aged less than 30 years. In addition, 46.7% (161/345) were employed full time, 95.9% (331/345) had at least secondary school education, and 35.9% (124/345) were from Zimbabwe. Linkage to care between 2 weeks and 8 months was 48.6% (88/181) in the intervention arm versus 45.1% (74/164) in the control (P=.52) and increased to 64.1% (116/181) and 61.0% (100/164) (P=.55), respectively, after the initial 8-month period. Moreover, youth aged 18 to 30-years showed a statistically significant 20% increase in linkage to care for the intervention group.

Conclusions: Youth aged less than 30 years have been historically difficult to reach with traditional interventions, and the SmartLink app provides a proof of concept that this population reacts to mobile health interventions that engage patients in HIV care.

Trial Registration: ClinicalTrials.gov NCT02756949; https://clinicaltrials.gov/ct2/show/NCT02756949 (Archived by WebCite at http://www.webcitation.org/6z1GTJCNW)

(JMIR Mhealth Uhealth 2019;7(4):e12652) doi:10.2196/12652

KEYWORDS

cell phones; HIV; app; Africa; linkage to care; patient information

Introduction

Background

South Africa has the largest antiretroviral therapy (ART) program in the world, which provides free ART to approximately 4.4 million people living with HIV [1], and since its introduction in 2004, AIDS-related deaths and new HIV infections have been reduced by 58% and 46%, respectively [1]. The country strategy has been created in line with international guidelines and updated with the emergence of new bodies of evidence and global initiatives [2-4].

In 2015, the *90-90-90* initiative was introduced by the Joint United Nations Programme on HIV/AIDS and the World Health Organization as a way to further decrease new infections among the population, recognizing the large impact of ART on infectiousness, while also optimizing individual health. The initiative maximizes the effect of ART coverage by emphasizing that 90% of HIV-positive people should know their status, 90% of those eligible for ART should be initiated on ART, and 90% of those on ART should achieve and maintain viral suppression [5].

South Africa has accomplished moderate success with HIV testing and viral suppression, achieving 85% and 86% success rates, respectively; however, only 71% of the people eligible for ART are on treatment [6]. It is well documented that patients, especially young people aged less than 30 years and men, are being lost to follow-up along the entire HIV care cascade, but the most significant attrition is found during the stage from HIV diagnosis to the start of treatment [7-9]. Improving this deficit is needed to ensure that patients are initiated on ART early as patients lost during linkage to care often return as late presenters when they become seriously ill. Late presenters may also continue spreading the virus, further increasing the risk of infection and threatening the 90-90-90 targets [10].

In September 2016, South Africa adopted the *treatall* approach for ART treatment by dropping the CD4 thresholds for ART initiation completely [11], yet patients could still expect several clinic visits before initiating ART [12]. These visits consist of initial HIV testing, followed by determination of treatment eligibility, adherence counseling, and education, as well as baseline blood tests and a physical examination before receiving the antiretrovirals [12]. Each of these visits represents a risk to the continuum of care of the newly diagnosed HIV cases and simplifying this process has been hypothesized as a way to decrease patient drop-off. Various interventions such as home-based testing and treatment and same-day initiation of ART have been tested to address this attrition, but there remains a gap [10,12-14].

The emergence of mobile health (mHealth) in developing countries has enabled some successful interventions across the continuum of HIV care, especially on the promotion of treatment adherence. With 90% of the world's population living in areas with mobile phone coverage and two-thirds of these people able to access data on their devices, mHealth provides an efficient method to engage the population [15]. Short message service (SMS) text messages and mobile apps have been used with moderate success in developing countries to improve ART adherence and appointment attendance [16-19]. South Africa has also experienced success with mHealth interventions, including the MomConnect program, which provides antenatal support through SMS and a help desk to almost 2 million pregnant mothers across the country [20].

The majority of mHealth interventions still focus on SMS text messaging, but by 2020, smartphone penetration in South Africa is expected to exceed 50% of the population [21]. The smartphones allow for data-based messaging, which should be considered for population scaling, as these platforms are much cheaper than SMS text messaging. Research surrounding linkage to care and the piloting, feasibility, and effectiveness of mHealth apps is needed to ensure that these interventions remain current as the population transition from basic phones to smartphones [15,19,22].

SmartLink is an mHealth app designed to provide HIV-positive smartphone owners with their laboratory results securely and rapidly, coupled with supportive information as well as prompts to link to care. Methods and information on the app development, including the challenges and limitations of the study, have been previously published [23] and will not be discussed in detail here.

Objectives

This study presents the evaluation of SmartLink to improve linkage to care for newly diagnosed HIV-positive smartphone owners through a randomized controlled trial. Of particular interest is the linkage to care of men and youth aged less than 30 years, as these populations have been historically hard to reach with traditional interventions [7-9]. Virological suppression was also evaluated as a secondary outcome.

Methods

Trial Design

The study was designed as a multisite randomized controlled trial where newly diagnosed HIV-positive participants were approached upon having a positive HIV test and were then screened for trial eligibility. Eligible and consenting trial candidates were randomized 1:1 into either the intervention or the control arm of the study. Participants in the intervention arm were then aided with the installation and setup of SmartLink.

Setting

The inner city of Johannesburg is one of South Africa's most densely populated areas, with an estimated population of 1 million people; numerous socioeconomic challenges such as overcrowding, unemployment, crime, poverty, substance abuse, and sex work; and a high HIV prevalence [24]. The area has a well-established HIV testing and ART program, with some

health care facilities providing ART to over 20,000 patients. However, the transient nature of the community makes it difficult to measure actual testing, linkage, and retention rates at the population level [25]. Participants were recruited at 5 public HIV testing sites (1 community health center, 3 clinics, and 1 tertiary hospital) from October 2015 to June 2016 and then followed up until February 2017.

Participants

Trained field workers at the 5 testing sites approached newly diagnosed HIV-positive people for trial participation after they had blood drawn for CD4 count measuring. Trial candidates were prescreened. Participants were considered for the trial if they were a resident in the area, aged 18 years and above, not pregnant, and could read English or Zulu (2 commonly understood languages in the area) [23]. Individuals were then screened for app compatibility; ineligible participants were excluded from the study if they had no active subscriber identity module card in their phone, no Android smartphone, or no data on their phone. It was discovered that the app could not be installed if the participant had insufficient RAM on their phone or if their Android version was too old (pre-version 4.2), so these parameters were also added to the exclusion criteria. Eligible participants who passed screening were then recruited into the study and randomized 1:1 into the intervention arm or the control arm using a pregenerated randomization table.

Intervention

Study staff assisted participants from the intervention arm with the installation of the SmartLink app, which was done with an Android install file and Wi-Fi dongle to allow installation at no data cost to the participants.

The app, available in English or Zulu, was designed to engage participants in their own care by directly providing them with 2 laboratory results; appointment reminders; and information about the laboratory tests, ART adherence, and HIV in general (Multimedia Appendix 1). The 2 laboratory results were CD4 count and viral load, and they were communicated in simple language. These values were also expressed visually on a color-coded scale that showed *normal* values and were accompanied by a short explanation of the results and guidance as to what action, if any, should be taken.

Participants randomized into the control arm received the standard of care, where participants received counseling and were referred to their local ART initiation site to collect their laboratory results and initiate appropriate treatment as needed. All participants, regardless of the study arm, were instructed to attend their local clinic for a follow-up within a few weeks of trial commencement and not to wait for the results on their phone.

App Security

The SmartLink logo, app icon, and landing page made no reference to HIV, AIDS, or health care to ensure that a participant's HIV or other health status would not be accidentally disclosed when viewing the app name or icon on a participant's phone. Furthermore, to protect confidential medical information from being available to other people, app security was modeled after local banking apps. This ensured security and privacy by employing a username, password, and a personal identification number to gain access to personal health data.

Outcomes

To capture HIV-related laboratory monitoring (our proxy for linkage to HIV care), evidence of an HIV-related laboratory test result between 2 weeks and 8 months of participant recruitment was sought. Test results were available on the National Health Laboratory Service (NHLS) database, which covers all local public facilities (but not initiation by private general practitioners or workplaces, although these provide very limited access in terms of absolute numbers), and included CD4, viral load, or creatinine clearance. Clinic visits were tracked after the initial 8 months until the completion of the follow-up in February 2017 to see if any lag to linkage to care was present in either trial arm. The viral load results were also analyzed to determine if virological suppression was achieved as a secondary outcome.

Due to these abovementioned independent databases as well as analytical data from the app developers, consolidation of these data was required. The investigators implemented a method to keep track of trial participants and their laboratory results by creating a centralized universal study dataset. To ensure intervention fidelity, this dataset was continuously monitored and evaluated by researchers to identify any potential variances [23].

Data Analysis

On the basis of the market research conducted in early 2015 at the study sites and a primary outcome measured as a second HIV-related laboratory test between 2 weeks and 8 months, a sample size of at least 1000 participants for each study arm was anticipated to measure a 20% difference in linkage to care between the intervention and control arms of each study subgroup such as young men. This was calculated based on a significance of .05, a power of 80%, and an estimated loss to follow-up of 27% (Hillbrow Community health Centre data).

Descriptive statistics were used to summarize baseline characteristics, presented as categorical data with frequency (percentage). All outcomes were compared between the intervention and control arms by linkage to care with the Pearson Chi-square test for significance. All data analyses were performed with Stata version 12.1 (StataCorp LP, College Station, TX).

The SmartLink protocol was approved by the University of Witwatersrand's Medical Human Research Ethics Committee (Certificate: M150606), the City of Johannesburg, and Gauteng's Department of Health at the provincial level and was registered in ClinicalTrials.gov (NCT02756949).

Results

Participant Flow

The participant flow diagram is shown in Figure 1. Of the 4537 individuals approached about the study, only 90 people (2.0%) declined to participate; however, a total of 4094 people (90.2%)

were found to be ineligible during the prescreening and screening. The data from 8 participants in the control arm were also removed from analysis because of the erroneous sending of SMS reminders for their 6-month clinic appointment. Once removed, 164 participants (3.6%) remained in the control arm, and 181 participants (4.0%) remained in the intervention arm. A complete breakdown of enrollment based on inclusion and exclusion criteria has been reported [23].

Figure 1. SmartLink participant flow diagram. SMS: short message service.

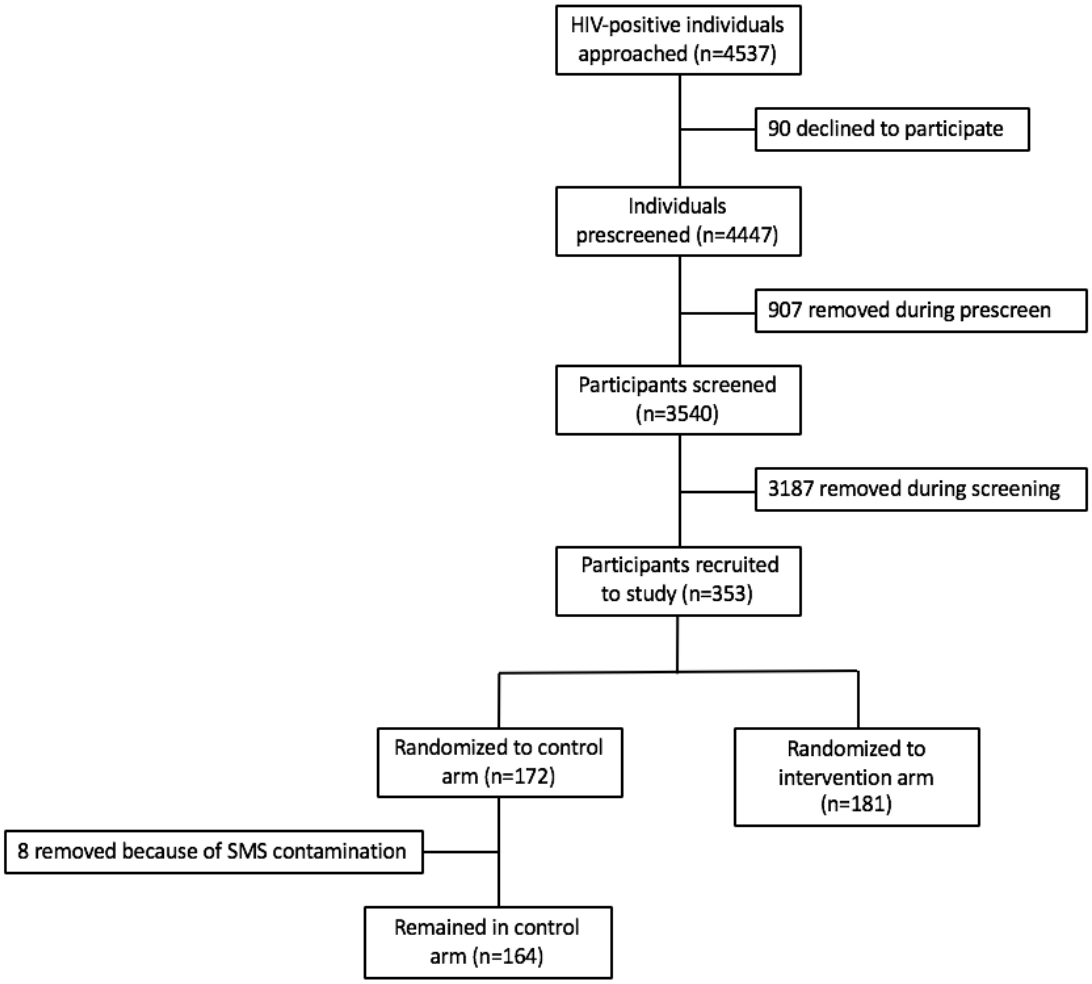


Table 1. Baseline characteristics of SmartLink trial participants.

Characteristic	Control (n=164), n (%)[a]	Intervention (n=181), n (%)	Total (N=345),n (%)
Sex			
Male	61 (37.2)	60 (33.1)	121 (35.1)
Female	103 (62.8)	121 (66.9)	224 (64.9)
Age (years)			
18-30	69 (42.1)	83 (45.9)	152 (44.1)
31+	95 (57.9)	98 (54.1)	193 (55.9)
Country of birth			
South Africa	95 (57.9)	103 (56.9)	198 (57.4)
Zimbabwe	61 (37.2)	63 (34.1)	124 (35.9)
Other	8 (5.9)	15 (8.3)	23 (6.67)
Education			
Primary only	6 (3.7)	8 (4.4)	14 (4.1)
Some secondary school	44 (26.9)	51 (28.2)	95 (27.5)
Completed secondary school	85 (51.8)	98 (54.1)	183 (53.0)
Attended/completed tertiary	29 (17.7)	24 (13.3)	53 (15.4)
Employment status			
Employed full time	79 (48.2)	82 (45.3)	161 (46.7)
Employed part time	22 (13.4)	37 (20.4)	59 (17.1)
Unemployed	40 (24.4)	49 (27.1)	89 (25.8)
Self-employed	16 (9.8)	10 (5.5)	26 (7.5)
Student	7 (4.4)	3 (1.7)	10 (2.9)

[a]Total may not add to 100% because of decimal rounding.

Baseline Characteristics

There were no significant differences in the baseline characteristics between the intervention and control arms (Table 1). Overall, only one-third of the participants were male (35.1%) and nearly half (44.1%) were youth aged less than 30 years. Almost half of the participants were employed full time (46.7%) and the majority had at least attended secondary school (95.9%). In addition, 57.4% of the participants were South African and just over one-third (35.9%) were from Zimbabwe. These baseline characteristics reflect the demographics of inner-city Johannesburg, where many migrants from Zimbabwe have settled and become part of the local population. These migrants are often well educated and possibly more likely to be employed than South Africans living in the inner city.

Although the 2 trial arms were well balanced in terms of participants' characteristics, there was, however, an important bias as to which demographic groups entered the trial because of the smartphone eligibility criteria [23]. Those with less education, those earning less, and those unemployed or underemployed are less likely to have smartphones to be eligible for the study [23].

Primary Outcome: Linkage to Care

This study called for a sample size of 2000 total participants; however, because of several challenges and limitations outlined, recruitment numbers were much lower than anticipated at 345 [23].

Total Cohort

Linkage to HIV care between 2 weeks and 8 months was just under 50% (45.1% control vs 48.6% intervention; *P*=.516) and increased to just over 60% (61.0% control vs 64.1% intervention; *P*=.551) after the initial 8-month period (Table 2).

Males

The male population, which was of specific interest, showed a slightly higher (but not statistically significant) linkage to care with the app between 2 weeks and 8 months (47.5% control vs 55.0% intervention; *P*=.412), but after 8 months, both these values were similar, approximately 66% (67.2% control vs 66.7% intervention; *P*=.949).

Youth Aged Between 18 and 30 Years

Despite the small sample size, a statistically significant difference was seen with youth aged between 18 and 30 years. Linkage to care between 2 weeks and 8 months was approximately 20% higher for youth with the app (31.9% control vs 53.0% intervention; *P*=.009), and this remained true after 8 months as well (50.7% control vs 69.9% intervention; *P*=.016; Table 2).

Table 2. Linkage to care.

Group	Linked to care for 2 weeks to 8 months					Ever linked to care				
	NLC[a], n (%)	LC[b], n (%)	Total, n	Pearson χ^2	*P* value	NLC, n (%)	LC, n (%)	Total, n	Pearson χ^2	*P* value
Total cohort										
Control	90 (54.9)	74 (45.1)	164	0.4	.52	64 (39.0)	100 (61.0)	164	0.4	.55
Intervention	93 (51.4)	88 (48.6)	181	0.4	.52	65 (35.9)	116 (64.1)	181	0.4	.55
Total	183 (53.0)	162 (47.0)	345	0.4	.52	129 (37.4)	216 (62.6)	345	0.4	.55
Males										
Control	32 (52.5)	29 (47.5)	61	0.7	.41	20 (32.8)	41 (67.2)	61	0.0	.95
Intervention	27 (45.0)	33 (55.0)	60	0.7	.41	20 (33.3)	40 (66.7)	60	0.0	.95
Total	59 (48.4)	63 (51.6)	121	0.7	.41	40 (33.1)	81 (66.9)	121	0.0	.95
Females										
Control	58 (56.3)	45 (43.7)	103	0.1	.79	44 (42.7)	59 (57.3)	103	0.7	.40
Intervention	66 (54.6)	55 (45.5)	121	0.1	.79	45 (37.2)	76 (62.8)	121	0.7	.40
Total	124 (55.4)	100 (44.6)	224	0.1	.79	89 (39.7)	135 (60.3)	224	0.7	.40
Youth aged between 18 and 30 years										
Control	47 (68.1)	22 (31.9)	69	6.8	.01	34 (49.3)	35 (50.7)	69	5.8	.02
Intervention	39 (47.0)	44 (53.0)	83	6.8	.01	25 (30.1)	58 (69.9)	83	5.8	.02
Total	86 (56.6)	66 (43.4)	152	6.8	.01	59 (38.8)	93 (61.2)	152	5.8	.02
Aged over 30 years										
Control	43 (45.3)	52 (54.7)	95	1.9	.17	30 (31.6)	65 (68.4)	95	1.8	.18
Intervention	54 (55.1)	44 (44.9)	98	1.9	.17	40 (40.8)	58 (59.2)	98	1.8	.18
Total	97 (50.3)	96 (49.7)	193	1.9	.17	70 (36.3)	123 (63.7)	193	1.8	.18

[a]NLC: not linked to care.

[b]LC: linked to care.

Table 3. Viral load suppression.

Study group	Virally suppressed				
	Yes, n (%)	No, n (%)	Total, n (%)	Pearson χ^2	*P* value
Intervention	28 (63.6)	16 (36.4)	44 (100.0)	0.2	.66
Control	23 (59.0)	16 (41.0)	39 (100.0)	0.2	.66
Total	51 (61.5)	32 (38.6)	83 (100.0)	0.2	.66

Secondary Outcome: Viral Load Suppression

For participants who had viral load tests in the NHLS database, virological suppression was assessed as an outcome. As recruitment numbers were much lower than anticipated, participant results were also low, and no statistically significant results were reached; however, these values are presented for completeness (Table 3). As of February 2017, a total of 83 participants had viral load tests that could be used for analysis, 39 out of 164 (23.8%) from the control arm and 44 out of 181 (24.3%) from the intervention arm. With viral load suppression defined as less than 400 copies/mL, 59.0% of the control arm and 63.6% of the intervention arm experienced suppression; however, the *P* value of .663 negated any significance.

Discussion

Principal Findings

Although this was the first evaluation using a smartphone-enabled app to support linkage to HIV care in Africa, as far as we are aware, the study outcomes were limited due to being underpowered as a result of complications and limitations surrounding app compatibility. As a proof of concept, the SmartLink app worked as anticipated; however, the smartphone specifications required for installation excluded over 90% of candidates who volunteered to participate in the study. This is unfortunately a common trend in mHealth studies, where many interventions show generally positive results; however, they are often inconclusive or are not substantial

enough when extrapolated out to a broader population or scaled up [16,26].

Although this study demonstrated that app-linked information and prompting can lead to increased linkage to care, the specific technology was not evaluated. The SmartLink app provided patients with laboratory results, information, and appointment reminders, but the relative efficacy of these specific components could not be explored. Despite the challenges in trial enrolment, one population of interest, youth aged 18 to 30 years, showed a statistically significant benefit of the app. This subpopulation experienced a 20% increase in linkage to care for the app group, and this is encouraging as HIV patients in this age group have historically been very difficult to engage with traditional interventions [8]. In South Africa, this population is 16% more likely to own a smartphone and 19% more likely to access the internet with their phone than their parents [27]. The high smartphone ownership coupled with our evidence of increased linkage in care strongly suggests that mHealth apps for engagement in care should be considered for this age group.

This demographic will become more and more familiar with technology, reinforcing the need to create a strong body of evidence surrounding these mHealth interventions. For children aged 9 to 17 years, 80% have access to internet on a smartphone and 84% own their own device. This generation is growing up with the internet, social media, and apps and already possesses the same mobile skills set as their parents, with children even surpassing them with knowledge about creating media and installing apps [28]. Future studies should focus on tailoring mHealth interventions toward youth, while also providing an opportunity to standardize counseling and support communications from health care providers [29].

Conclusions

This proof-of-concept study has demonstrated that SmartLink can significantly increase linkage to care for youth aged 18 to 30 years; however, further evaluation with larger samples is required to recommend such an intervention for programmatic rollout. This research is of timely importance as demand for entry-level smartphones (sub-US $100) in developing countries had led to over 400 million smartphone units being sold in the first quarter of 2018 alone [30]. As smartphone penetration increases and prices decrease, new innovations such as using quick response code technology coupled with patient-held smartcards can allow for information to be transferred without internet access or data [31]. During this shift, mHealth apps should also be considered for incorporation into multifaceted interventions as bundling apps with SMS text messaging, phone calls, or in-person communications could be a way to optimally engage patients while app familiarity and technology continue to improve [32].

Limitations

Secondary outcomes, such as ART initiation rates, feasibility, satisfaction, and participants' knowledge, could not be evaluated because of the limitations, as outlined by Venter et al in 2018 [23]. Analytics on app use by participants also could not be evaluated because of complications in data collection between the devices and the back-end analytics software. Essentially, data exchange between the relevant systems could not be achieved during the trial, limiting the scope of log-in analytics to counts of app openings only. Finally, we acknowledge the limitations of the trial in terms of generalizability, as already mentioned. The eligibility criteria lead to a selected patient group. For instance, a relatively high proportion of Zimbabwean patients and more educated patients were better able to qualify for the trial.

Acknowledgments

The authors would like to acknowledge all the patients and health care workers who participated in the study, National Health Laboratory Service (NHLS) that facilitated in the data collection, and the World Bank for their technical and financial assistance.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the SmartLink app.

[PNG File, 916KB - mhealth_v7i4e12652_app1.png]

Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 551KB - mhealth_v7i4e12652_fig.pdf]

References

1. UNAIDS. Geneva; 2018. URL: http://www.unaids.org/sites/default/files/media_asset/unaids-data-2018_en.pdf [accessed 2018-10-22] [WebCite Cache ID 73LxdKdmv]

2. National Department of Health (2018). Annual Performance Plan 2018/19-2020/21. Pretoria, South Africa. ISBD: 9780621461749.
3. World Health Organization. Geneva; 2015 Nov 25. Guideline on When to Start Antiretroviral Therapy and on Pre-Exposure Prophylaxis for HIV URL: https://www.who.int/sorry/ [accessed 2019-02-18] [WebCite Cache ID 76IBOdg89]
4. World Health Organization. Congo; 2016. WHO Country Cooperation Strategy 2016-2020 URL: https://apps.who.int/iris/bitstream/handle/10665/255007/ccs_zaf_2016_2020.pdf?sequence=1&isAllowed=y
5. AIDS. UNAIDS. Geneva; 2017. Global AIDS Update URL: http://www.unaids.org/en/resources/documents/2017/%202017_data_book [accessed 2019-02-18] [WebCite Cache ID 76IBqoFe0]
6. The Human Sciences Research Council. 2018. The Fifth South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2017: HIV Impact Assessment Summary Report URL: http://www.hsrc.ac.za/uploads/pageContent/9234/SABSSMV_Impact_Assessment_Summary_ZA_ADS_cleared_PDFA4.pdf [accessed 2019-02-18] [WebCite Cache ID 76IEzX9V0]
7. NACOSA - Networking HIV & AIDS Community of Southern Africa NPC. 2016. Adherence Guidelines for HIV, TB and NCDs URL: https://www.nacosa.org.za/wp-content/uploads/2016/11/Integrated-Adherence-Guidelines-NDOH.pdf [accessed 2018-07-08] [WebCite Cache ID 6rmrZXOMi]
8. Maheu-Giroux M, Tanser F, Boily M, Pillay D, Joseph SA, Bärnighausen T. Determinants of time from HIV infection to linkage-to-care in rural KwaZulu-Natal, South Africa. AIDS 2017 Dec 24;31(7):1017-1024 [FREE Full text] [doi: 10.1097/QAD.0000000000001435] [Medline: 28252526]
9. Takarinda KC, Harries AD, Shiraishi RW, Mutasa-Apollo T, Abdul-Quader A, Mugurungi O. Gender-related differences in outcomes and attrition on antiretroviral treatment among an HIV-infected patient cohort in Zimbabwe: 2007-2010. Int J Infect Dis 2015 Jan;30:98-105 [FREE Full text] [doi: 10.1016/j.ijid.2014.11.009] [Medline: 25462184]
10. Rosen S, Fox MP. Retention in HIV care between testing and treatment in sub-Saharan Africa: a systematic review. PLoS Med 2011 Jul;8(7):e1001056 [FREE Full text] [doi: 10.1371/journal.pmed.1001056] [Medline: 21811403]
11. Motsoaledi A. Bhekisisa Centre for Health Journalism. #AIDS2016: Health Minister Aaron Motsoaledi admits that 'key leaders were in denial' URL: https://bhekisisa.org/article/2016-07-15-aids2016-health-minister-aaron-motsoaledi-admits-that-key-leaders-were-in-denial/ [accessed 2018-10-22] [WebCite Cache ID 73M7Xp6zp]
12. Long LC, Maskew M, Brennan AT, Mongwenyana C, Nyoni C, Malete G, et al. Initiating antiretroviral therapy for HIV at a patient's first clinic visit: a cost-effectiveness analysis of the rapid initiation of treatment randomized controlled trial. AIDS 2017 Jul 17;31(11):1611-1619 [FREE Full text] [doi: 10.1097/QAD.0000000000001528] [Medline: 28463879]
13. Iwuji CC, Orne-Gliemann J, Larmarange J, Okesola N, Tanser F, Thiebaut R, ANRS 12249 TasP trial group. Uptake of home-based HIV testing, linkage to care, and community attitudes about ART in rural KwaZulu-Natal, South Africa: descriptive results from the first phase of of the ANRS 12249 TasP cluster-randomised trial. PLoS Med 2016 Aug;13(8):e1002107 [FREE Full text] [doi: 10.1371/journal.pmed.1002107] [Medline: 27504637]
14. Larson BA, Schnippel K, Brennan A, Long L, Xulu T, Maotoe T, et al. Same-day CD4 testing to improve uptake of HIV care and treatment in South Africa: point-of-care is not enough. AIDS Res Treat 2013;2013:941493 [FREE Full text] [doi: 10.1155/2013/941493] [Medline: 23956850]
15. Stephani V, Opoku D, Quentin W. A systematic review of randomized controlled trials of mHealth interventions against non-communicable diseases in developing countries. BMC Public Health 2016 Dec 15;16(1):572-510 [FREE Full text] [doi: 10.1186/s12889-016-3226-3] [Medline: 27417513]
16. Mbuagbaw L, Mursleen S, Lytvyn L, Smieja M, Dolovich L, Thabane L. Mobile phone text messaging interventions for HIV and other chronic diseases: an overview of systematic reviews and framework for evidence transfer. BMC Health Serv Res 2015 Jan 22;15:33 [FREE Full text] [doi: 10.1186/s12913-014-0654-6] [Medline: 25609559]
17. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014 Feb;9(2):e88166 [FREE Full text] [doi: 10.1371/journal.pone.0088166] [Medline: 24505411]
18. Kanters S, Park JJ, Chan K, Socias ME, Ford N, Forrest JI, et al. Interventions to improve adherence to antiretroviral therapy: a systematic review and network meta-analysis. Lancet HIV 2017 Dec;4(1):e31-e40. [doi: 10.1016/S2352-3018(16)30206-5] [Medline: 27863996]
19. Perera AI, Thomas MG, Moore JO, Faasse K, Petrie KJ. Effect of a smartphone application incorporating personalized health-related imagery on adherence to antiretroviral therapy: a randomized clinical trial. AIDS Patient Care STDS 2014 Nov;28(11):579-586 [FREE Full text] [doi: 10.1089/apc.2014.0156] [Medline: 25290556]
20. Seebregts C, Dane P, Parsons AN, Fogwill T, Rogers D, Bekker M, et al. Designing for scale: optimising the health information system architecture for mobile maternal health messaging in South Africa (MomConnect). BMJ Glob Health 2018 Apr;3(Suppl 2):e000563 [FREE Full text] [doi: 10.1136/bmjgh-2017-000563] [Medline: 29713506]
21. GSMA Intelligence. 2017. The Mobile Economy Sub-Saharan Africa 2018 URL: https://www.gsmaintelligence.com/research/?file=809c442550e5487f3b1d025fdc70e23b&download

22. Fox M, Rosen S, Geldsetzer P, Bärnighausen T, Negussie E, Beanland R. Interventions to improve the rate or timing of initiation of antiretroviral therapy for HIV in sub-Saharan Africa: meta-analyses of effectiveness. J Int AIDS Soc 2016;19(1):20888 [FREE Full text] [doi: 10.7448/IAS.19.1.20888] [Medline: 27507249]
23. Venter W, Coleman J, Chan VL, Shubber Z, Phatsoane M, Gorgens M, et al. Improving linkage to HIV care through mobile phone apps: randomized controlled trial. JMIR Mhealth Uhealth 2018 Jul 17;6(7):e155 [FREE Full text] [doi: 10.2196/mhealth.8376] [Medline: 30021706]
24. Tshuma N, Muloongo K, Setswe G, Chimoyi L, Sarfo B, Burger D, et al. Potential barriers to rapid testing for human immunodeficiency virus among a commuter population in Johannesburg, South Africa. HIV AIDS (Auckl) 2015;7:11-19 [FREE Full text] [doi: 10.2147/HIV.S71920] [Medline: 25565898]
25. Venter WF, Coleman J, Lau CV, Phatsoane M. World Bank Group. 2017. Smart Linkage to Care : Evaluation Report URL: http://documents.worldbank.org/curated/en/374611542662090016/pdf/WP-132214-SmartLtCEvaluationReportUpdateNovFINAL.pdf [accessed 2019-02-19] [WebCite Cache ID 76INX0Ipd]
26. Aranda-Jan C, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014 Feb 21;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
27. Poushter J. Pew Research Center.: Pew Global; 2016 Feb 22. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies URL: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ [accessed 2018-10-22] [WebCite Cache ID 73MDoUknh]
28. LSE Research Online. 2016 Sep. Global Kids Online South Africa: barriers, opportunities and risks. A glimpse into South African children's internet use and online activities URL: http://eprints.lse.ac.uk/71267/2/GKO_Country-Report_South-Africa_CJCP_upload.pdf [accessed 2019-02-19] [WebCite Cache ID 76IOCZ6gF]
29. Section 27. How we can improve HIV testing and counselling URL: https://section27.org.za/2012/04/4568/ [accessed 2018-10-22] [WebCite Cache ID 73MEGfPQe]
30. Costello K, Hippold S. Gartner. 2018. Gartner Says Worldwide Sales of Smartphones Returned to Growth in First Quarter of 2018 URL: https://www.gartner.com/newsroom/id/3876865 [accessed 2018-10-22] [WebCite Cache ID 73MEkAOAz]
31. Gbadamosi SO, Eze C, Olawepo JO, Iwelunmor J, Sarpong DF, Ogidi AG, et al. A patient-held smartcard with a unique identifier and an mHealth platform to improve the availability of prenatal test results in rural Nigeria: demonstration study. J Med Internet Res 2018 Jan 15;20(1):e18 [FREE Full text] [doi: 10.2196/jmir.8716] [Medline: 29335234]
32. Mugo PM, Wahome EW, Gichuru EN, Mwashigadi GM, Thiong'o AN, Prins HA, et al. Effect of text message, phone call, and in-person appointment reminders on uptake of repeat HIV testing among outpatients screened for acute HIV infection in Kenya: a randomized controlled trial. PLoS One 2016 Apr 14;11(4):e0153612 [FREE Full text] [doi: 10.1371/journal.pone.0153612] [Medline: 27077745]

Abbreviations

ART: antiretroviral therapy
mHealth: mobile health
NHLS: National Health Laboratory Service
SMS: short message service

Edited by G Eysenbach; submitted 30.10.18; peer-reviewed by K Chigudu, S Gbadamosi, E Green; comments to author 06.12.18; revised version received 21.12.18; accepted 03.01.19; published 02.04.19.

Please cite as:
Venter WDF, Fischer A, Lalla-Edward ST, Coleman J, Lau Chan V, Shubber Z, Phatsoane M, Gorgens M, Stewart-Isherwood L, Carmona S, Fraser-Hurt N
Improving Linkage to and Retention in Care in Newly Diagnosed HIV-Positive Patients Using Smartphones in South Africa: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(4):e12652
URL: https://mhealth.jmir.org/2019/4/e12652/
doi:10.2196/12652
PMID:30938681

Original Paper

Mobile Health for Tuberculosis Management in South India: Is Video-Based Directly Observed Treatment an Acceptable Alternative?

Anil A Kumar[1], MBBS; Ayesha De Costa[2], MBBS, MD, PhD; Arundathi Das[3], MBBS, MD; GA Srinivasa[3], MBBS, DPH; George D'Souza[4], MBBS, MD, DNB; Rashmi Rodrigues[2,5,6], MBBS, MD, PhD

[1]St John's Research Institute, St John's National Academy of Health Sciences, Bangalore, India

[2]Public Health Sciences, Karolinska Institutet, Stockholm, Sweden

[3]Rajiv Gandhi University of Health Sciences, Bangalore, India

[4]Department of Chest Medicine, St John's Medical College, St John's National Academy of Health Sciences, Bangalore, India

[5]Department of Community Health, St John's Medical College, St John's National Academy of Health Sciences, Bangalore, India

[6]The Wellcome Trust/DBT India Alliance, Hyderabad, India

Corresponding Author:
Rashmi Rodrigues, MBBS, MD, PhD
Department of Community Health
St John's Medical College
St John's National Academy of Health Sciences
Koramangala 4th Block
Bangalore, 560034
India
Phone: 91 9845389538
Email: rashmijr@gmail.com

Abstract

Background: With the availability of low-cost mobile devices and the ease of internet access, mobile health (mHealth) is digitally revolutionizing the health sector even in resource-constrained settings. It is however necessary to assess end-user perceptions before deploying potential interventions.

Objective: This study aimed to assess the mobile phone usage patterns and the acceptability of mobile phone support during care and treatment in patients with tuberculosis (TB) in South India.

Methods: This exploratory study was conducted at an urban private tertiary care teaching hospital and nearby public primary-level health care facilities in Bangalore, South India. We recruited 185 patients with TB through consecutive sampling. Subsequent to written informed consent, participants responded to an interviewer-administered pretested questionnaire. The questionnaire included questions on demographics, phone usage patterns, and the benefits of using of mobile phone technology to improve health outcomes and treatment adherence. Frequency, mean, median, and SD or interquartile range were used to describe the data. Bivariate associations were assessed between demographics, clinical details, phone usage, and mHealth communication preferences using the chi-square test and odds ratios. Associations with a P value ≤.20 were included in a logistic regression model. A P value of <.05 was considered significant.

Results: Of the 185 participants, 151 (81.6%) used a mobile phone, and half of them owned a smartphone. The primary use of the mobile phone was to communicate over voice calls (147/151, 97.4%). The short message service (SMS) text messaging feature was used by only 66/151 (43.7%) mobile phone users. A total of 87 of the 151 mobile phone users (57.6%) knew how to use the camera. Only 41/151 (27.2%) mobile phone users had used their mobile phones to communicate with their health care providers. Although receiving medication reminders via mobile phones was acceptable to all participants, 2 participants considered repeated reminders as an intrusion of their privacy. A majority of the participants (137/185, 74.1%) preferred health communications via voice calls. Of the total participants, 123/185 (66.5%) requested reminders to be sent only at specific times during the day, 22/185 (11.9%) suggested reminders should synchronize with their prescribed medication schedule, whereas 40/185 (21.6%) did not have any time preferences. English literacy was associated with a preference for SMS in comparison with voice calls. Most participants (142/185, 76.8%) preferred video-based directly observed treatment when compared with in-person directly observed treatment.

Conclusions: Although mobile phones for supporting health and treatment adherence were acceptable to patients with TB, mHealth interventions should consider language, mode of communication, and preferred timing for communication to improve uptake.

(JMIR Mhealth Uhealth 2019;7(4):e11687) doi:10.2196/11687

KEYWORDS

medical informatics; tuberculosis; mHealth; adherence; mobile phone; reminder; SMS; voice call; DOT; vDOT; video DOT

Introduction

World tuberculosis (TB) surveillance estimates that 10 million people are either diagnosed or relapse with TB every year [1]. With a case fatality rate of 16%, TB is one of the most frequent causes for death from a single infectious agent, second only to HIV/AIDS [1]. Furthermore, the emergence of HIV infection in 1983, led to a resurgence in TB, making TB-HIV coinfection a threat of greater significance. India contributes 27% to the world's burden of TB, the highest among the 10 high TB burden countries globally [1,2].

To address the burden of TB globally, the World Health Organization introduced the directly observed treatment, short-course (DOTS) strategy in 1992 [3]. DOTS comprises the following 5 elements: (1) political will for TB control, (2) case finding through quality diagnostics, (3) regular supply of antitubercular treatment (ATT), (4) short-course chemotherapy, which is the directly observed treatment (DOT) and, (5) a reliable TB information system [3]. DOTS was initiated in India in 1993, in a phased manner, through the Revised National Tuberculosis Control Program (RNTCP) [4]. The RNTCP provides ATT at no cost to patients with TB. However, as it implements DOTS, patients are required to visit a health care provider and swallow their medication under observation [4]. The alternative is for patients with TB to avail treatment through the private health care sector, at a cost. It is estimated that twice the number of patients with TB are treated in the private health care sector when compared with the public sector [5].

Given that TB requires 6 months of treatment with up to 4 antitubercular drugs, ensuring treatment success is a challenge, both from the patients' and health care providers' perspective. The need to stay motivated throughout the treatment along with a high pill burden, medication side effects, poverty, stigma, and discrimination, serve as barriers to treatment adherence [6-8]. In addition, forgetfulness and HIV coinfection can influence adherence negatively [7,9]. Along with these patient-associated barriers, health care system–related factors, such as ATT stockouts and unfavorable attitudes of health care providers toward patients with TB, also play a role [8,10-13].

Most Indian literature on ATT nonadherence reflects the proportions of patients lost to follow-up once initiated on ATT. The global TB report 2018 indicates that only 69% of Indian patients initiated on ATT are treated successfully; the rest either fail the treatment or are lost to follow-up or succumb to the illness [14]. Literature indicates that loss to follow-up with ATT in India ranges from 6% to 44% [15-18], whereas proportions of patients interrupting treatment for more than 1 month range from 14% to 50% [7,19].

Nonadherence to ATT has led to the emergence of drug-resistant strains of TB, which are resistant to either a single drug (monoresistance) or several drugs (multidrug resistance, MDR) or are extremely drug-resistant TB [20]. MDR TB treatment is associated with a higher financial burden, longer duration of treatment, and lower treatment success rates [21]. Given this, ensuring early diagnosis and treatment of TB minimizes the pill burden, making treatment regimens shorter, cheaper, and easier to comply with.

India has approximately 1.18 billion wireless subscribers, including mobile phone users [22]. Such a high wireless user base makes the use of mobile phones for health care delivery inevitable. Mobile phone reminders, such as voice calls and short message service (SMS) text messaging, to improve adherence to ATT have shown mixed results [23,24]. However, mobile reminders are known to improve clinic attendance [25]. Furthermore, mobile health (mHealth) [26] interventions have led to better retention of patients with TB when compared with historical cohorts [27]. A study from Lesotho, Africa, indicates that 92% of HIV/TB patients found SMS reminders for medications acceptable [28]. However, a randomized controlled trial from Pakistan found that SMS did not significantly improve treatment outcomes compared with a control group [29]. *Photovoice*, an app that used video recordings from patients cured of TB to promote ATT adherence and outcomes in Pune, India, showed better outcomes in patients exposed to the intervention [30]. In addition, mobile video-based directly observed treatment (vDOT), an alternative to conventional in-person DOT [31], holds promise, given the high mobile phone penetration and wireless users, especially in the Indian context [32].

Many mobile phone apps that are in use for TB are health care provider–centric and aid in either data collection or referral of patients with TB [33,34]. Given that few mobile apps for management of TB exist, mHealth for TB and its treatment is underexplored [34]. Furthermore, there is a paucity of information in the Indian context regarding the use of mobile phone interventions for TB. We therefore decided to explore the acceptability of mobile phone apps and the type of apps patients with TB would prefer. Such information is expected to support the development of patient-centric mobile phone apps for TB in the Indian context.

Methods

Study Site

The participants for this cross-sectional, exploratory study were recruited through consecutive sampling from both private and public sector health care facilities in Bangalore, Karnataka,

India. The private facility was St John's Medical College Hospital (SJMCH), Bangalore. This is a 1250 bedded, nonprofit, tertiary care teaching hospital that caters to patients largely from the South Indian states of Karnataka, Andhra Pradesh, and Tamil Nadu. The public health facilities involved were urban health centers in the vicinity of SJMCH in Bangalore. These public health facilities implement national health programs and provide health care at no cost to all those in need.

Diagnosis and Treatment of Tuberculosis at St John's Medical College Hospital (Private Health Care Facility)

Patients suspected of clinically active TB are subjected to microbiological and/or radiological tests for confirmation. Newly diagnosed patients with TB are started on a Category I ATT comprising isoniazid (H), rifampicin (R), pyrazinamide, and ethambutol for 6 months. This includes an initial 2 months of an intensive phase of treatment with all the 4 drugs, followed by 4 months of a continuation phase of treatment with 2 drugs (H and R). Patients with relapse or drug resistance receive more complex ATT regimens involving injectable medications. ATT at the hospital is available either through (1) the DOT center, a public-private initiative that enables no-cost treatment through the RNTCP or (2) for a cost through the hospital's pharmacy. Patients follow up with physicians every month routinely for health appraisals that include clinical examination, monitoring adverse effects, or prescription refills.

Diagnosis and Treatment of Tuberculosis at Public Facilities

Diagnostic protocols at the public health care facilities are similar to those at SJMCH. However, all patients diagnosed with TB and treated at public health care facilities receive treatment only through the RNTCP. All patients within the RNTCP are expected to receive DOT.

Participants and Data Collection

Between February 2016 and December 2017, 185 patients with TB aged between 18 and 60 years, receiving treatment at the study sites, were enrolled in the study. Of the participants enrolled, 159/185 (85.9%) received ATT at SJMCH, whereas 26/185 (14.1%) received ATT at public health care facilities. Both newly diagnosed patients with TB and those already receiving ATT were included in the study. Patients who were seriously ill or those who did not understand the purpose of the study were excluded.

Subsequent to written informed consent, trained research assistants administered a questionnaire in the local language to the study participants. The questionnaire obtained basic sociodemographic information from the participants along with information regarding (1) the basic functionality of their mobile phones, (2) acceptability of delivering adherence support via mobile phones, and (3) the type of mobile phone intervention acceptable to them, such as SMS, voice call, interactive voice response system, or vDOT.

Data Analysis

Data were analyzed using IBM-SPSS version 24. Frequencies, means, medians, SDs, and interquartile ranges were used to describe the data. The outcome variables studied were (1) the preference for voice calls compared with SMS reminders and (2) the preference for in-person DOT compared with vDOT. Some categorical variables that had multiple categories were converted into binary variables. Bivariate associations were assessed between demographics, clinical details, phone usage, and the preference-based outcome variables using the chi-square test and odds ratio (OR). Unadjusted logistic regression was used to derive OR for variables with more than 2 categories. Bivariate associations with a P value ≤.20 were included in an adjusted logistic regression model. Associations with P values <.05 were considered significant.

Ethics Statement

Ethical clearance for the study was obtained from the Institutional Ethics Committee, St John's Medical College, Bangalore, India. Written informed consent was obtained from all participants after providing them with study-related information, either verbally or in writing, before administering the questionnaire.

Results

Overview

A total of 185 patients with TB participated in the study. The mean age of the participants was 35.25 (SD 11.59) years. Of the participants, 114/185 (61.6%) were males, and 121/185 (65.4%) resided in an urban area. There were 44/185 (23.8%) participants on in-person DOT, 45/185 (24.3%) on self-administered treatment, and 96/185 (51.9%) participants for whom treatment was yet to be initiated. The demographic characteristics of the patients are presented in Table 1.

Clinically, 98/185 (53.0%) patients had pulmonary TB, and 159/185 (85.9%) were newly diagnosed patients with TB on category I ATT.

Ownership of Mobile Phones

Of the 185 participants, 151/185 (81.6%) used a mobile phone. Among these 144/151 (95.4%) owned the phone, and 85/151 (56.3%) had used mobile phones for 6 years or more. Of those who owned mobile phones, 65/144 (45.1%) owned smartphones, and the rest (79/144, 54.9%) owned basic phones. The major reasons cited for not owning a phone included not needing a phone (19/34, 56%), inability to use a mobile phone (8/34, 24%), and financial constraints (3/34, 9%). Only 7/144 (4.9%) mobile phone users reported using a phone shared with other family members. Men were 4 times as likely as women to own mobile phones (unadjusted OR 3.816; 95% CI 1.747-8.338). Significant factors associated with phone ownership included education and a monthly income of 500 INR or more (Table 2).

Table 1. Demographic profile of study participants (N=185).

Variables	Total (n=185)	Female (n=71)	Male (n=114)	*P* value
Age (years)				
Median (IQR[a])	32 (26-45)	30 (24-45.5)	33 (27-44)	—[b]
≥32, n (%)	97 (52.4)	37 (52)	58 (50.9)	Referent[c]
<32, n (%)	88 (47.6)	34 (48)	56 (49.1)	.20
Marital status, n (%)				
Married	123 (66.5)	45 (63)	78 (68.4)	Referent
Single	62 (33.5)	26 (37)	36 (31.6)	.48
Residence, n (%)				
Rural	64 (34.6)	20 (28)	44 (38.6)	Referent
Urban	121 (65.4)	51 (72)	70 (61.4)	.15
Education status, n (%)				
No formal education	40 (21.6)	17 (24)	23 (20.2)	Referent
Formal education[d]	145 (78.4)	54 (76)	91 (79.8)	.55
Literate in English, n (%)				
No	116 (62.7)	43 (61)	73 (64.0)	Referent
Yes	69 (37.3)	28 (39)	41 (36.0)	.64
Employment status, n (%)				
Not gainfully employed	76 (41.1)	49 (69)	27 (23.7)	Referent
Gainfully employed	109 (58.9)	22 (31)	87 (76.3%)	<.001
Monthly income (INR[e])				
Median (IQR)	5000 (0-12000)	0 (0-5000)	9000 (2000-15000)	—
≥5000, n (%)	97 (52.4)	71 (100)	59 (51.8)	Referent
<5000, n (%)	88 (47.6)	0 (0)	55 (48.2)	<.001
Type of patient, n (%)				
New patient[f]	159 (85.9)	64 (90)	95 (83.3)	Referent
Others[g]	26 (14.1)	7 (10)	19 (16.7)	.20
Type of TB[h], n (%)				
Pulmonary	98 (52.9)	32 (45)	66 (57.9)	Referent
Extrapulmonary	87 (47.1)	39 (55)	48 (42.1)	.09
Microscopy (TB bacilli), n (%)				
Negative	100 (54.1)	47 (66)	53 (46.5)	Referent
Positive	85 (45.9)	24 (34)	61 (53.5)	.009
Treatment phase, n (%)				
Intensive	159 (85.9)	65 (92)	94 (82.5)	Referent
Continuation	26 (14.1)	6 (8)	20 (17.5)	.08
Treatment category, n (%)				
Category I	164 (88.6)	65 (92)	99 (86.8)	Referent
Others	21 (11.4)	6 (8)	15 (13.2)	.33
Treatment observation, n (%)				
In-person directly observed treatment.	44 (23.8)	16 (22)	28 (24.6)	Referent
Self-administered treatment	45 (24.3)	14 (20)	31 (27.2)	.60

Variables	Total (n=185)	Female (n=71)	Male (n=114)	*P* value
Not initiated	96 (51.9)	41 (58)	55 (48.2)	.48
Recruitment, n (%)				
Public health care facilities	26 (14.1)	10 (14)	16 (14.0)	Referent
St John's Medical College Hospital (private)	159 (85.9)	61 (86)	98 (86.0)	.99

[a]IQR: interquartile range.

[b]Not applicable.

[c]Referent: reference category.

[d]Formal education: this category includes middle school and above.

[e]1 INR (Indian Rupee)=0.014 US $, November 2018.

[f]New patient: a patient newly diagnosed with tuberculosis.

[g]Others: treatment after loss to follow-up or retreatment.

[h]TB: tuberculosis.

Table 2. Access to mobile phones and its association with demographic characteristics (N=185).

Variables	Mobile phone not used (n=34), n (%)	Mobile phone used (n=151), n (%)	Unadjusted OR[a] (95% CI)	Adjusted[b,c] OR (95% CI)
Sex, n (%)				
Female	22 (31)	49 (69)	Referent[d]	—[e]
Male	12 (10.5)	102 (89.5)	3.816 (1.747-8.338)	—
Education status, n (%)				
No formal education	14 (35)	26 (65)	Referent	Referent
Formal education	20 (13.8)	125 (86.2)	3.365 (1.508-7.513)	2.623 (1.118-6.153)
Literate in English, n (%)				
No	28 (24.1)	88 (75.9)	Referent	Referent
Yes	6 (9)	63 (91)	3.341 (1.306-8.546)	—
Employment status, n (%)				
Not gainfully employed	25 (33)	51 (67)	Referent	—
Gainfully employed	9 (8.3)	100 (91.7)	5.447 (2.367-12.531)	—
Monthly income (Indian Rupee), n (%)				
<5000	28 (32)	60 (68)	Referent	Referent
≥5000	6 (6)	91 (94)	7.078 (2.765-18.120)	6.288 (2.428-16.290)
Treatment observation, n (%)				
In-person directly observed treatment	8 (18)	36 (82)	Referent	—
Self-administered treatment	11 (24)	34 (76)	0.687 (0.247-1.913)	—
Not initiated	15 (16)	81 (84)	1.200 (0.467-3.083)	—
Recruitment center, n (%)				
Public health care facility	6 (23)	20 (77)	Referent	—
St John's Medical College Hospital (private)	28 (17.6)	131 (82.4)	1.404 (0.517-3.813)	—

[a]OR: odds ratio.

[b]Logistic regression model *P* value<.001 (Forward stepwise [conditional] method); Nagelkerke R^2: 0.215 (step 2); −2 Log-likelihood: 150.230 (step 2).

[c]Only variables retained in the final regression model have an adjusted OR.

[d]Referent: reference category.

[e]Not applicable.

Basic Functionality of Mobile Phones

Of the participants who used a mobile phone, 66/151 (43.7%) used the SMS feature, 72/151 (47.7%) used the alarm function, whereas 87/151 (57.6%) knew how to use the camera on their phone for photography and/or videography. Of those who used the alarm on the phone, only 2/73 (3%) used the alarm as a medication reminder.

Participants less than 32 years of age (adjusted OR 2.314, 95% CI 1.068-5.025) or those literate in English (adjusted OR 8.678, 95% CI 4.019-18.740) were more likely to use the SMS feature than their counterparts. In addition, those who were single (unadjusted OR 2.793, 95% CI 1.479-5.263), residing in an urban area (unadjusted OR 3.493, 95% CI 1.695-7.195), or formally educated (unadjusted OR 15.012, 95% CI 3.489-64.591) were more likely to use the SMS compared with those who were married, were from a rural area, or were not formally educated, respectively (Table 3).

Preferred Mobile Phone Interventions for Management of Tuberculosis

Of the 185 participants, 182 (98.4%) agreed to receive health information on their mobile phones. Topics that the participants preferred included information on available medications, advances in TB management, and medication reminders. Participants also requested communication with health care provider, motivational health messages, specific diet, and prevention of TB as additional features (Figure 1).

In response to specific queries about the preferred mode of health communication, 137/185 (74.1%) of the participants chose voice calls over the SMS (40/185, 21.6%). In addition, 8/185 (4.3%) participants preferred either. Most of those (43/48, 90%) who preferred the SMS requested to receive them in English. On the contrary, most of those who preferred voice calls (127/137, 92.7%) requested for communication in regional languages. Of the 185 participants, 78 (42.2%) chose to receive reminders as often as they required to take their medication, whereas the rest preferred reminders either once a week or less frequently. Similarly, with regard to the timing of the reminders, 123/185 (66.5%) preferred reminders at specific times, whereas 22/185 (11.9%) preferred reminders just before they took their medication. The remaining 40/185 (21.6%) participants were willing to receive reminders *anytime* as they were *free at home*. Overall, 183/185 (98.9%) of the participants did not perceive the reminders as an intrusion of their privacy.

Preference for voice calls was significantly associated with age, marital status, literacy in English, type of TB, and the ability to use camera on their phones (Table 4).

Most of the study participants preferred vDOT over the conventional in-person DOT. This preference was associated with being male, residing in an urban area, and being formally educated. Other factors such as literacy in English, the ability to use SMS and phone camera, although associated with preference for vDOT in bivariate analyses, were not found associated with this preference in an adjusted logistic regression model. Clinical characteristics were not associated with the preference for the mode of communication or adherence monitoring strategy (vDOT or in-person DOT; Table 5).

Table 3. Use of text messaging and its association with demographic characteristics (N=185).

Variables	SMS[a] not used[b] (n=119), n (%)	SMS used (n=66), n (%)	Unadjusted OR[c] (95% CI)	Adjusted[d,e] OR (95% CI)
Sex, n (%)				
Female	45 (63)	26 (37)	Referent[f]	—[g]
Male	74 (64.9)	40 (35.1)	0.936 (0.505-1.734)	—
Age (years)				
≥32, n (%)	78 (80)	19 (20)	Referent	Referent
<32, n (%)	41 (47)	47 (53)	4.716 (2.450-9.009)	2.314 (1.068-5.025)
Marital status, n (%)				
Married	89 (72.4)	34 (27.6)	Referent	—
Single	30 (48)	32 (52)	2.793 (1.479-5.263)	—
Residence, n (%)				
Rural	52 (81)	12 (19)	Referent	—
Urban	67 (55.4)	54 (44.6)	3.493 (1.695-7.195)	—
Education status, n (%)				
No formal education	38 (95)	2 (5)	Referent	—
Formal education	81 (55.9)	64 (44.1)	15.012 (3.489-64.591)	—
Literate in English, n (%)				
No	99 (85.3)	17 (14.7)	Referent	Referent
Yes	20 (29)	49 (71)	14.268 (6.865-29.654)	8.678 (4.019-18.740)
Monthly income (Indian Rupee), n (%)				
<5000	64 (73)	24 (27)	Referent	—
≥5000	55 (57)	42 (43)	2.036 (1.098-3.776)	—
Treatment observation, n (%)				
In-person directly observed treatment	28 (64)	16 (36)	Referent	—
Self-administered treatment	33 (73)	12 (27)	0.636 (0.258-1.569)	—
Not initiated	58 (60)	38 (40)	1.147 (0.548-2.398)	—
Recruitment center, n (%)				
Public health care facility	18 (69)	8 (31)	Referent	—
St John's Medical College Hospital (private)	101 (63.5)	58 (36.5)	1.292 (0.529-3.157)	—

[a]SMS: short message service.

[b]Comprised those who did not use SMSs as they did not have a phone and those who had a phone but did not use the feature.

[c]OR: odds ratio.

[d]Logistic regression model *P* value<.001 (Forward stepwise [Conditional] method); Nagelkerke R^2: 0.445 (step 3); −2 Log-likelihood: 168.595 (step 3).

[e]Only variables retained in the final regression model have an adjusted OR.

[f]Referent: reference category.

[g]Not applicable.

Figure 1. Type of health information requested over mobile phone (N=185). TB: tuberculosis.

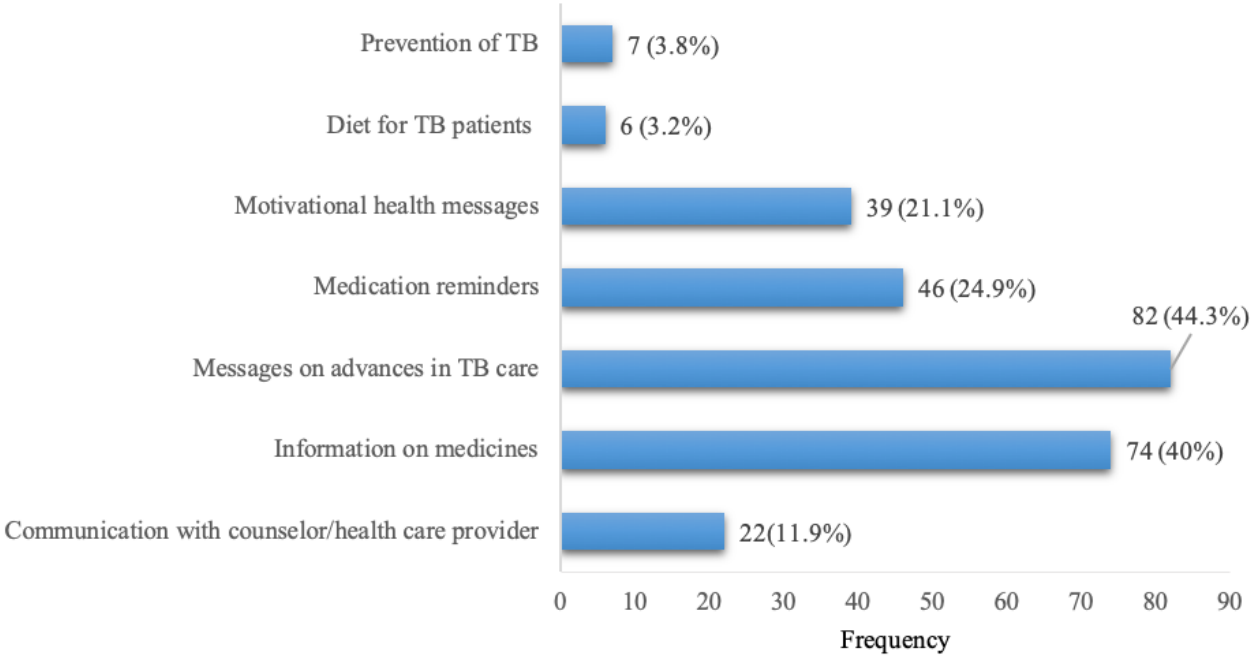


Table 4. Preference for an intervention and its association with clinical and demographic characteristics (N=185).

Variables	Prefer SMS[a] (n=48), n (%)	Prefer voice call (n=137), n (%)	Unadjusted OR[b] (95% CI)	Adjusted[c,d] OR (95% CI)
Sex, n (%)				
Female	20 (28)	51 (72)	Referent[e]	—[f]
Male	28 (24.6)	86 (75.4)	1.204 (0.616-2.354)	—
Age (years)				
<32, n (%)	40 (46)	48 (55)	Referent	Referent
≥32, n (%)	8 (8)	89 (92)	9.271 (4.017-21.396)	4.129 (1.557-10.947)
Marital status, n (%)				
Single	30 (48)	32 (52)	Referent	Referent
Married	18 (14.6)	105 (85.4)	5.469 (2.700-11.076)	2.934 (1.172-7.346)
Residence, n (%)				
Rural	8 (13)	56 (88)	Referent	—
Urban	40 (33.1)	81 (66.9)	0.289 (0.126-0.665)	—
Education status, n (%)				
No formal education	1 (3)	39 (97)	Referent	—
Formal education	47 (32.4)	98 (67.6)	0.053 (0.007-0.401)	—
Literate in English, n (%)				
No	12 (10.3)	104 (89.7)	Referent	Referent
Yes	36 (52)	33 (48)	0.106 (0.049-0.227)	0.265 (0.108-0.652)
Employment status, n (%)				
Not gainfully employed	17 (22)	59 (78)	Referent	—
Gainfully employed	31 (28.4)	78 (71.6)	0.725 (0.367-1.433)	—
Monthly income (Indian Rupee), n (%)				
<5000	16 (18)	72 (82)	Referent	—
≥5000	32 (33)	65 (67)	0.451 (0.227-0.898)	—
Type of patient, n (%)				
New patient	44 (27.7)	115 (72.3)	Referent	—
Others	4 (15)	22 (85)	2.104 (0.686-6.453)	—
Type of TB[g], n (%)				
Extrapulmonary	28 (32)	59 (68)	Referent	Referent
Pulmonary	20 (20)	78 (80)	1.851 (0.950-3.597)	3.205 (1.290-7.936)
Microscopy (TB bacilli), n (%)				
Negative	33 (33)	67 (67)	Referent	—
Positive	15 (18)	70 (82)	2.299 (1.146-4.611)	—
Treatment phase, n (%)				
Intensive	40 (25.2)	119 (74.8)	Referent	—
Continuation	8 (31)	18 (69)	0.756 (0.305-1.873)	—
Treatment category, n (%)				
Category I	45 (27.4)	119 (72.6)	Referent	—
Others	3 (14)	18 (86)	2.269 (0.638-8.075)	—
Treatment status, n (%)				

Variables	Prefer SMS[a] (n=48), n (%)	Prefer voice call (n=137), n (%)	Unadjusted OR[b] (95% CI)	Adjusted[c,d] OR (95% CI)
In-person directly observed treatment	15 (34)	29 (66)	Referent	—
Self-administered treatment	6 (13)	39 (87)	3.362 (1.163-9.721)	—
Not initiated	27 (28)	69 (72)	1.322 (0.615-2.843)	—
Recruitment center, n (%)				
Public health care facility	8 (31)	18 (69)	Referent	—
St John's Medical College Hospital (private)	40 (25.2)	119 (74.8)	1.322 (0.534-3.274)	—
Access to phone, n (%)				
No	4 (12)	30 (88)	Referent	—
Yes	44 (29.1)	107 (70.9)	0.324 (0.108-0.975)	—
SMS use, n (%)				
No	13 (10.9)	106 (89.1)	Referent	—
Yes	35 (53)	31 (47)	0.109 (0.051-0.230)	—
Camera use, n (%)				
No	9 (9)	89 (91)	Referent	Referent
Yes	39 (45)	48 (55)	0.124 (0.056-0.278)	0.243 (0.092-0.640)

[a]SMS: short message service.

[b]OR: odds ratio.

[c]Logistic regression model *P* value<.001 (Forward stepwise [conditional] method); Nagelkerke R^2: 0.491; −2 Log-likelihood: 136.497.

[d]Only variables retained in the final regression model have an adjusted OR.

[e]Referent: reference category.

[f]Not applicable.

[g]TB: tuberculosis.

Table 5. Preference for video-based directly observed treatment to in-person directly observed treatment and its association with demographics, clinical details, and mobile phone usage characteristics.

Variables	Prefer DOT[a] (n=43), n (%)	Prefer vDOT[b] (n=142), n (%)	Unadjusted OR[c] (95% CI)	Adjusted OR[d,e] (95% CI)
Sex, n (%)				
Female	26 (37)	45 (63)	Referent[f]	Referent
Male	17 (14.9)	97 (85.1)	3.297 (1.627-6.680)	4.004 (1.846-8.683)
Age (years)				
<32, n (%)	14 (16)	74 (84)	Referent	—[g]
≥32, n (%)	29 (30)	68 (70)	0.444 (0.216-0.909)	—
Marital status, n (%)				
Single	11 (18)	51 (82)	Referent	—
Married	32 (26.1)	91 (73.9)	0.613 (0.285-1.319)	—
Residence, n (%)				
Rural	22 (34)	42 (66)	Referent	Referent
Urban	21 (17.4)	100 (82.6)	2.494 (1.241-5.014)	2.626 (1.197-5.765)
Education status, n (%)				
No formal education	18 (45)	22 (55)	Referent	Referent
Formal education	25 (17.2)	120 (82.8)	3.927 (1.841-8.376)	3.391 (1.492-7.709)
Literate in English, n (%)				
Others	34 (29.3)	82 (70.7)	Referent	—
English	9 (13)	60 (87)	2.764 (1.234-6.193)	—
Employment status, n (%)				
Not gainfully employed	20 (26)	56 (74)	Referent	—
Gainfully employed	23 (21.1)	86 (78.9)	1.335 (0.672-2.655)	—
Monthly income (Indian Rupee), n (%)				
<5000	26 (30)	62 (70)	Referent	—
≥5000	17 (18)	80 (82)	1.973 (0.984-3.956)	—
Type of patient, n (%)				
New patient	37 (23.3)	122 (76.7)	Referent	—
Others	6 (23)	20 (77)	1.011 (0.378-2.704)	—
Type of TB[h], n (%)				
Pulmonary	24 (24)	74 (76)	Referent	—
Extrapulmonary	19 (22)	68 (78)	1.161 (0.585-2.305)	—
Microscopy (TB bacilli) , n (%)				
Negative	25 (25)	75 (75)	Referent	—
Positive	18 (21)	67 (78)	1.241 (0.623-2.473)	—
Treatment phase, n (%)				
Intensive	40 (25.2)	119 (74.8)	Referent	—
Continuation	3 (12)	23 (88)	2.577 (0.734-9.043)	—
Treatment category, n (%)				
Category I	39 (23.8)	125 (76.2)	Referent	—
Others	4 (19)	17 (81)	1.326 (0.421-4.175)	—
Treatment observation, n (%)				

Variables	Prefer DOT[a] (n=43), n (%)	Prefer vDOT[b] (n=142), n (%)	Unadjusted OR[c] (95% CI)	Adjusted OR[d,e] (95% CI)
In-person DOT	13 (30)	31 (70)	Referent	—
Self-administered	11 (24)	34 (76)	1.296 (0.507-3.315)	—
Not initiated	19 (20)	77 (80)	1.699 (0.749-3.857)	—
Recruitment center, n (%)				
Public health care facility	8 (31)	18 (69)	Referent	—
St John's Medical College Hospital (private)	35 (22.1)	124 (77.9)	1.575 (0.632-3.925)	—
Access to phone, n (%)				
No	15 (44)	19 (56)	Referent	—
Yes	28 (18.5)	123 (81.5)	3.468 (1.571-7.654)	—
Short message service use, n (%)				
No	36 (30.3)	83 (69.7)	Referent	—
Yes	7 (11)	59 (89)	3.656 (1.523-8.776)	—
Camera use, n (%)				
No	29 (30)	69 (70)	Referent	—
Yes	14 (16)	73 (84)	2.192 (1.069-4.492)	—

[a]DOT: directly observed treatment.

[b]vDOT: video-based directly observed treatment.

[c]OR: odds ratio.

[d]Final adjusted logistic regression model *P* value<.001 (Forward stepwise [conditional] method); Nagelkerke R^2: 0.22; −2 Log-likelihood: 171.465.

[e]Only variables retained in the final regression model have an adjusted OR.

[f]Referent: reference category.

[g]Not applicable.

[h]TB: tuberculosis.

Discussion

Principal Findings

Barriers such as stigma, medication side effects, and transport to the health care facility can affect adherence to ATT. In this light, exploiting the pervasiveness of mobile phone technology to overcome these barriers and support medication adherence is a promising solution. Motivational health messages and customized medication reminders via mobile phones are some interventions designed to support adherence to ATT [24]. Although several studies have explored the use of mobile phone interventions for adherence support in chronic infectious diseases such as HIV and TB, not all have shown favorable results [29,35]. Factors such as the complexity, personalization, and mode of communication could affect acceptability, uptake, and success of mobile phone interventions. We therefore chose to explore the acceptability of mobile phone adherence support interventions for ATT and identify the characteristics of such interventions, which the patients with TB would prefer, before developing such an intervention.

In this study, most participants were willing to receive adherence reminders via mobile phones and did not consider such interventions as an intrusion of their privacy. Randomized controlled trials have shown the effectiveness of SMS reminders in HIV infection and malaria in sub Saharan Africa [36]. However, the effectiveness of mobile phone reminders for antiretroviral treatment support in India was questionable [35]. Nevertheless, given that treatment for TB is for 6 months compared with HIV infection where the treatment is lifelong, mHealth interventions for TB are likely to face lesser intervention fatigue and are worth exploring. Although patients with TB from Salem in Tamil Nadu state, South India, considered communication via mobile phones useful, they preferred in-person contact with health care providers. The study, however, did not assess the preferred mode or type of ATT adherence support [37].

A study from Lesotho, Africa, reported a high uptake of SMS interventions in HIV/TB patients [28]. On the contrary and consistent with other studies from South India, most of our study participants preferred voice calls in comparison with the SMS [38]. The young, the employed, and the educated participants were more likely to use the SMS for communication. These participants probably preferred reading an SMS text as opposed to answering a phone call as it saved time and attracted lesser attention when received. Given the limited literacy in English, the roman script was used for both English and regional languages in SMS communication; it is not surprising that SMS communication was less popular than voice calls both in this study and in the literature in the Indian context [37,38]. However, with the availability of regional language options in mobile phones, developing interventions in regional languages

is an option worth exploring. Furthermore, given that some of the participants were not literate but could use the basic functionality on mobile phones, interventions that use videos and pictures with limited requirements for literacy are an option.

Asynchronous vDOT is an accepted alternative suggested to circumvent the barriers to in-person DOT [39]. Studies have shown that vDOT is more confidential, easy to use [40], and allows health care providers to efficiently monitor a larger number of patients at a distance when compared with in-person DOT [31]. Barriers to vDOT include interruption of data connectivity [31], loss or theft of phone [40], and an inability to confirm that medicine was actually taken in certain settings [41]. Over three-fourths of the participants preferred vDOT to in-person DOT, despite no experience with the intervention, citing vDOT as an optimal solution for saving time and money or minimizing hospital visits.

Few participants expressed concern with sending their video to the health care provider. Fear of disclosure of their illness to family, unknown people watching their videos, fear of the videos getting published via social media, and discomfort with video-recording themselves were reasons expressed for the concern. Therefore, counseling the beneficiaries of the measures taken to safeguard their videos along with reinforcing the importance of adherence to ATT is essential to ensure the uptake of vDOT. Smartphone apps, although nonintrusive, were found beneficial in the management of HIV infection despite their limited functionality [42].

In this light, the abundant features for patient-centric mHealth interventions for TB can be explored. Most of the participants suggested newer apps incorporating disease-related information and behavior change communication, which could be incorporated into existing or newer apps. The concept of *photovoice* [30], where patients cured of TB shared their treatment experiences and replaced health care personnel, can be considered an option. *Photovoice* can also be incorporated into vDOT for health education and communication with patients.

Methodological Issues

Given that standard TB care in India is based on geographic location, most of the patients were from urban areas and therefore more representative of urban patients with TB. Also, as one-third of the patients received ATT through the RNTCP (at SJMCH or at public health care facilities), the study mirrors the public-private mix in TB care in India. Furthermore, as many of the study participants were newly diagnosed with limited treatment experience, their opinions are also likely to change with treatment. This, along with the limited experience of the participants with mobile phone interventions may mean that opinions may change with actual interventions. The limitations in estimating a sample size and the nonprobability sampling technique used may affect the generalizability of our findings. Nevertheless, the results of the study cannot be undermined as they inform patient-centric intervention design in contexts little exposed to mHealth interventions.

In addition, in terms of implementation reality, factors such as poverty, smartphone penetrance, internet access, and level of education necessary for using mobile phones will have to be addressed in an integrated manner to maximize the potential benefit of vDOT in ensuring treatment success in TB.

Conclusions

This study sought to assess whether communication via mobile phones could be an acceptable form of health care delivery in the context of patients with TB. We found that adherence reminders and information disseminating apps were acceptable in the management of TB. Contrary to the popularity of SMS-based reminders elsewhere globally, most of the study participants preferred voice calls. Efficacy of mHealth interventions could be improved when components that enable the inclusion of all demographic groups are incorporated along with enabling customizations to individual needs. Given the popularity of voice calls, interventions should include a voice component along with various language options in the Indian context. Although facing interceptable barriers such as privacy and stigma, vDOT as an alternative to DOT appears to hold promise in the Indian context. The effectiveness of mobile phone apps such as vDOT may therefore be worth exploring in the Indian context, while ensuring privacy and confidentiality of the end user.

Acknowledgments

The authors would like to thank Ms Shubha Krishnamurthy, Ms Hannah DJ, Dr Nikitha Prabhakar, and Dr Naseer Ahmed who supported data collection and data management. The authors convey special thanks to Wellcome Trust/DBT India Alliance and the Swedish Research Council for funding support. The authors also appreciate the support from the offices of the district TB officers, Bruhat Bengaluru Mahanagara Palike, and Bangalore Urban District, Bangalore, India, through their respective public health care facilities in urban and rural Bangalore.

Authors' Contributions

The concept and design of this study were provided by RR, ADC, and GDS. Data management and analysis were done by AAK, AD, GAS, and RR. The drafting the manuscript was done by AAK, GAS and RR, and the critical review was done by ADC and GDS.

Conflicts of Interest

None declared.

References

1. World Health Organization. Global Tuberculosis Report 2017. Geneva: WHO; 2017. URL: https://www.who.int/tb/publications/global_report/gtbr2017_main_text.pdf [accessed 2019-04-01] [WebCite Cache ID 77JJxsVFF]
2. World Health Organization. Global Tuberculosis Report 2016. Geneva: WHO; 2016. URL: https://www.who.int/sorry/ [accessed 2019-04-01] [WebCite Cache ID 77JK40AzK]
3. Raviglione MC, Pio A. Evolution of WHO policies for tuberculosis control, 1948-2001. Lancet 2002 Mar 02;359(9308):775-780. [doi: 10.1016/S0140-6736(02)07880-7] [Medline: 11888605]
4. World Health Organization. A Brief History of Tuberculosis Control in India. Geneva: WHO; 2010. URL: https://apps.who.int/iris/bitstream/handle/10665/44408/9789241500159_eng.pdf?sequence=1
5. Arinaminpathy N, Batra D, Khaparde S, Vualnam T, Maheshwari N, Sharma L, et al. The number of privately treated tuberculosis cases in India: an estimation from drug sales data. Lancet Infect Dis 2016 Nov;16(11):1255-1260 [FREE Full text] [doi: 10.1016/S1473-3099(16)30259-6] [Medline: 27568356]
6. Gebreweld FH, Kifle MM, Gebremicheal FE, Simel LL, Gezae MM, Ghebreyesus SS, et al. Factors influencing adherence to tuberculosis treatment in Asmara, Eritrea: a qualitative study. J Health Popul Nutr 2018 Jan 5;37(1):1 [FREE Full text] [doi: 10.1186/s41043-017-0132-y] [Medline: 29304840]
7. Kulkarni P, Akarte S, Mankeshwar R, Bhawalkar J, Banerjee A, Kulkarni A. Non-adherence of new pulmonary tuberculosis patients to anti-tuberculosis treatment. Ann Med Health Sci Res 2013 Jan;3(1):67-74 [FREE Full text] [doi: 10.4103/2141-9248.109507] [Medline: 23634333]
8. Yellappa V, Lefèvre P, Battaglioli T, Narayanan D, van der Stuyft P. Coping with tuberculosis and directly observed treatment: a qualitative study among patients from South India. BMC Health Serv Res 2016 Dec 19;16:283 [FREE Full text] [doi: 10.1186/s12913-016-1545-9] [Medline: 27430557]
9. Adane AA, Alene KA, Koye DN, Zeleke BM. Non-adherence to anti-tuberculosis treatment and determinant factors among patients with tuberculosis in northwest Ethiopia. PLoS One 2013;8(11):e78791 [FREE Full text] [doi: 10.1371/journal.pone.0078791] [Medline: 24244364]
10. Chapman HJ, Veras-Estévez BA, Pomeranz JL, Pérez-Then EN, Marcelino B, Lauzardo M. Perceived barriers to adherence to tuberculosis infection control measures among health care workers in the Dominican Republic. MEDICC Rev 2017 Jan;19(1):16-22 [FREE Full text] [Medline: 28225541]
11. Negandhi H, Tiwari R, Sharma A, Nair R, Zodpey S, Reddy Allam R, et al. Rapid assessment of facilitators and barriers related to the acceptance, challenges and community perception of daily regimen for treating tuberculosis in India. Glob Health Action 2017;10(1):1290315 [FREE Full text] [doi: 10.1080/16549716.2017.1290315] [Medline: 28485693]
12. Bagchi S, Ambe G, Sathiakumar N. Determinants of poor adherence to anti-tuberculosis treatment in Mumbai, India. Int J Prev Med 2010;1(4):223-232 [FREE Full text] [Medline: 21566777]
13. Dandona R, Dandona L, Mishra A, Dhingra S, Venkatagopalakrishna K, Chauhan LS. Utilization of and barriers to public sector tuberculosis services in India. Natl Med J India 2004;17(6):292-299. [Medline: 15736548]
14. World Health Organization. Global Tuberculosis Report 2018. Geneva: WHO; 2018. URL: https://www.who.int/tb/publications/global_report/en/ [WebCite Cache ID 750OLwmun]
15. Mittal C, Gupta S. Noncompliance to DOTS: how it can be decreased. Indian J Community Med 2011 Jan;36(1):27-30 [FREE Full text] [doi: 10.4103/0970-0218.80789] [Medline: 21687377]
16. Sarpal SS, Goel NK, Kumar D, Janmeja AK. Treatment outcome among the retreatment tuberculosis (TB) patients under RNTCP in Chandigarh, India. J Clin Diagn Res 2014 Feb;8(2):53-56 [FREE Full text] [doi: 10.7860/JCDR/2014/6510.4006] [Medline: 24701481]
17. Srivastava K, Gupta A, Saxena R, Sharma RP, Midha T. A study on non-compliance in tuberculosis cases towards the directly observed treatment short course under RNTCP in Kanpur Nagar. Int J Community Med Public Health 2017 Nov 23;4(12):4485. [doi: 10.18203/2394-6040.ijcmph20175166]
18. Sakalle S, Waskel B, Dixit S, Pandey D, Sirohi S, Saroshe S. A study on patient compliance of tuberculosis enrolled under revised national tuberculosis control programme. Natl J Community Med 2014;5(1):96-99 [FREE Full text]
19. Saxena V, Saxena S. Study of the causes and factors which affect compliance of the patient in directly observed short course chemotherapy in pulmonary tuberculosis in Central India. Int J Res Med Sci 2017 Jun 24;5(7):2989. [doi: 10.18203/2320-6012.ijrms20172974]
20. Mistry N, Tolani M, Osrin D. Drug-resistant tuberculosis in Mumbai, India: an agenda for operations research. Oper Res Health Care 2012 Jun;1(2-3):45-53 [FREE Full text] [doi: 10.1016/j.orhc.2012.06.001] [Medline: 24501697]
21. Udwadia Z, Mullerpattan J, Banka R, Ganatra S, Udwadia Z. The health economics of treating MDR-TB in the Indian private sector. Eur Respir J 2016;48(60) [FREE Full text] [doi: 10.1183/13993003.congress-2016.PA1536]

22. Telecom Regulatory Authority of India. The Indian Telecom Services Performance Indicators: July-September, 2017 URL: https://www.trai.gov.in/sites/default/files/PIR_July_Sept_28122017.pdf [accessed 2018-05-11] [WebCite Cache ID 6xDZCQTys]
23. Abdulrahman SA, Rampal L, Ibrahim F, Radhakrishnan AP, Kadir Shahar H, Othman N. Mobile phone reminders and peer counseling improve adherence and treatment outcomes of patients on ART in Malaysia: a randomized clinical trial. PLoS One 2017;12(5):e0177698 [FREE Full text] [doi: 10.1371/journal.pone.0177698] [Medline: 28520768]
24. Kunawararak P, Pongpanich S, Chantawong S, Pokaew P, Traisathit P, Srithanaviboonchai K, et al. Tuberculosis treatment with mobile-phone medication reminders in northern Thailand. Southeast Asian J Trop Med Public Health 2011 Nov;42(6):1444-1451 [FREE Full text] [Medline: 22299414]
25. Nglazi MD, Bekker L, Wood R, Hussey GD, Wiysonge CS. Mobile phone text messaging for promoting adherence to anti-tuberculosis treatment: a systematic review. BMC Infect Dis 2013 Dec 2;13:566 [FREE Full text] [doi: 10.1186/1471-2334-13-566] [Medline: 24295439]
26. World Health Organization. mHealth: New Horizons for Health Through Mobile Technologies. Geneva: WHO; 2011. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2019-04-01] [WebCite Cache ID 77JKCBSRf]
27. Choun K, Achanta S, Naik B, Tripathy JP, Thai S, Lorent N, et al. Using mobile phones to ensure that referred tuberculosis patients reach their treatment facilities: a call that makes a difference. BMC Health Serv Res 2017 Aug 22;17(1):575 [FREE Full text] [doi: 10.1186/s12913-017-2511-x] [Medline: 28830542]
28. Hirsch-Moverman Y, Daftary A, Yuengling KA, Saito S, Ntoane M, Frederix K, et al. Using mHealth for HIV/TB treatment support in Lesotho: enhancing patient–provider communication in the START study. J Acquir Immune Defic Syndr 2017 Jan 1;74(Suppl 1):S37-S43 [FREE Full text] [doi: 10.1097/QAI.0000000000001202] [Medline: 27930610]
29. Mohammed S, Glennerster R, Khan AJ. Impact of a daily SMS medication reminder system on tuberculosis treatment outcomes: a randomized controlled trial. PLoS One 2016;11(11):e0162944 [FREE Full text] [doi: 10.1371/journal.pone.0162944] [Medline: 27802283]
30. Shelke SC, Adhav PS, Moonan PK, Willis M, Parande MA, Satyanarayana S, et al. Photovoice: a novel approach to improving antituberculosis treatment adherence in Pune, India. Tuberc Res Treat 2014;2014:302601 [FREE Full text] [doi: 10.1155/2014/302601] [Medline: 25374679]
31. Chuck C, Robinson E, Macaraig M, Alexander M, Burzynski J. Enhancing management of tuberculosis treatment with video directly observed therapy in New York City. Int J Tuberc Lung Dis 2016 May;20(5):588-593. [doi: 10.5588/ijtld.15.0738] [Medline: 27084810]
32. Majumdar A, Sujiv A, Palanivel C. Video directly observed treatment: how effective will it be in Indian setting? J Family Med Prim Care 2015;4(1):152-153 [FREE Full text] [doi: 10.4103/2249-4863.152279] [Medline: 25811013]
33. Chadha S, Trivedi A, Nagaraja SB, Sagili K. Using mHealth to enhance TB referrals in a tribal district of India. Public Health Action 2017 Jun 21;7(2):123-126 [FREE Full text] [doi: 10.5588/pha.16.0080] [Medline: 28695085]
34. Iribarren SJ, Schnall R, Stone PW, Carballo-Diéguez A. Smartphone applications to support tuberculosis prevention and treatment: review and evaluation. JMIR Mhealth Uhealth 2016 May 13;4(2):e25 [FREE Full text] [doi: 10.2196/mhealth.5022] [Medline: 27177591]
35. Shet A, De Costa A, Kumarasamy N, Rodrigues R, Rewari BB, Ashorn P, HIVIND Study Team. Effect of mobile telephone reminders on treatment outcome in HIV: evidence from a randomised controlled trial in India. Br Med J 2014 Oct 24;349:g5978 [FREE Full text] [doi: 10.1136/bmj.g5978] [Medline: 25742320]
36. Raifman JR, Lanthorn HE, Rokicki S, Fink G. The impact of text message reminders on adherence to antimalarial treatment in northern Ghana: a randomized trial. PLoS One 2014;9(10):e109032 [FREE Full text] [doi: 10.1371/journal.pone.0109032] [Medline: 25350546]
37. Elangovan R, Arulchelvan S. A study on the role of mobile phone communication in tuberculosis DOTS treatment. Indian J Community Med 2013 Oct;38(4):229-233. [doi: 10.4103/0970-0218.120158] [Medline: 24302824]
38. DeSouza SI, Rashmi MR, Vasanthi AP, Joseph SM, Rodrigues R. Mobile phones: the next step towards healthcare delivery in rural India? PLoS One 2014;9(8):e104895 [FREE Full text] [doi: 10.1371/journal.pone.0104895] [Medline: 25133610]
39. Guidelines for Treatment of Drug-susceptible Tuberculosis and Patient Care 2017. Geneva: World Health Organization; 2017. URL: http://apps.who.int/iris/bitstream/10665/255052/1/9789241550000-eng.pdf?ua=1 [WebCite Cache ID 73hyFsyit]
40. Garfein RS, Collins K, Muñoz F, Moser K, Cerecer-Callu P, Raab F, et al. Feasibility of tuberculosis treatment monitoring by video directly observed therapy: a binational pilot study. Int J Tuberc Lung Dis 2015 Sep;19(9):1057-1064 [FREE Full text] [doi: 10.5588/ijtld.14.0923] [Medline: 26260824]
41. Belknap R, Weis S, Brookens A, Au-Yeung KY, Moon G, DiCarlo L, et al. Feasibility of an ingestible sensor-based system for monitoring adherence to tuberculosis therapy. PLoS One 2013;8(1):e53373 [FREE Full text] [doi: 10.1371/journal.pone.0053373] [Medline: 23308203]
42. Himelhoch S, Kreyenbuhl J, Palmer-Bacon J, Chu M, Brown C, Potts W. Pilot feasibility study of Heart2HAART: a smartphone application to assist with adherence among substance users living with HIV. AIDS Care 2017 Dec;29(7):898-904. [doi: 10.1080/09540121.2016.1259454] [Medline: 28043176]

Abbreviations

ATT: antitubercular treatment
DBT: Department of Biotechnology
DOT: directly observed treatment
DOTS: directly observed treatment, short-course
H: isoniazid
MDR: multidrug resistance
mHealth: mobile health
OR: odds ratio
R: rifampicin
RNTCP: Revised National Tuberculosis Control Program
SJMCH: St John's Medical College Hospital
SMS: short message service
TB: tuberculosis
vDOT: video-based directly observed treatment

Edited by G Eysenbach; submitted 30.07.18; peer-reviewed by S Abdulrahman, V Dhikav; comments to author 27.09.18; revised version received 30.12.18; accepted 23.01.19; published 03.04.19.

Please cite as:
Kumar AA, De Costa A, Das A, Srinivasa GA, D'Souza G, Rodrigues R
Mobile Health for Tuberculosis Management in South India: Is Video-Based Directly Observed Treatment an Acceptable Alternative?
JMIR Mhealth Uhealth 2019;7(4):e11687
URL: http://mhealth.jmir.org/2019/4/e11687/
doi:10.2196/11687
PMID:30942696

Original Paper

Patients' and Doctors' Perceptions of a Mobile Phone–Based Consultation Service for Maternal, Neonatal, and Infant Health Care in Bangladesh: A Mixed-Methods Study

Mafruha Alam[1], MPH; Cathy Banwell[1], PhD; Anna Olsen[1], PhD; Kamalini Lokuge[1], PhD

National Centre for Epidemiology and Population Health, Research School of Population Health, The Australian National University, Canberra, Australia

Corresponding Author:
Mafruha Alam, MPH
National Centre for Epidemiology and Population Health
Research School of Population Health
The Australian National University
62 Mills Road
Canberra, 2601
Australia
Phone: 61 6125 5602
Email: mafruha.alam@gmail.com

Abstract

Background: A mobile-based consultation service, or telehealth, can be used for remote consultations with health care professionals for screening, self-care management, and referral. In rural Bangladesh, where there is high demand for scarce male and even scarcer female doctors, remote consultations may help women seeking maternal and child health care. Aponjon is a mHealth service in Bangladesh that provides weekly voice or text messages to pregnant women, new mothers, and family members on various aspects of maternal, neonatal, and infant health. Subscribers can also access a dedicated 24*7 call center to discuss maternal, neonatal, and infant health or emergencies with medically trained doctors. The service provides advice, primary diagnoses, prescriptions, and referrals to subscriber callers.

Objective: We investigated the Aponjon service to understand access, acceptability, usability, benefits, and challenges of a mobile phone-based consultation service.

Methods: We conducted call log data analysis for September to November 2015 to understand how many unique subscribers accessed the service, who accessed the service, the geographical distribution of callers, and the purpose of the calls. We also conducted a qualitative exploratory substudy of eight married women and eight married men who were subscribers to and accessed the service during this time to understand their experiences. We interviewed 11 doctors from the same service who provided phone consultations to subscribers.

Results: Approximately 3894 unique subscribers accessed the service for single or multiple consultations during the study period; 68.36% (2662/3894) of subscribers were from rural households, and 53.00% (2064/3894) of calls were made by pregnant women or new mothers. Approximately 96.08% (5081/5288) calls were nonurgent, 2.69% (142/5288) semiurgent, and 1.23% (65/5288) urgent. Almost 64.7% (134/207) semiurgent or urgent calls came between 8 PM and 8 AM. Callers found the consultation service trustworthy, cost-effective, and convenient. The doctors dispelled misconceptions and promoted good health care practices, regular health check-ups, and responsible use of medicine. They helped families understand the severity of sicknesses and advised them to seek care at health facilities for semiurgent or urgent conditions. The service lacked a pro-poor policy to support talk times of subscribers from poor households and a proper referral system to help patients find the right care at the right facilities.

Conclusions: Although a regular messaging service is constrained by a one-way communication system, this service using the same platform, gave subscribers access to an abbreviated "consultation" with medical doctors. The consultations provided subscribers with valued medical advice and support, although they were limited in their population reach and their integration into the wider medical system. Further research is required to understand the impact of advice and referral, cost-effectiveness, and willingness to pay for mHealth consultation services, but this research suggests that these services should be supported or even expanded.

(JMIR Mhealth Uhealth 2019;7(4):e11842) doi:10.2196/11842

KEYWORDS

mobile-based consultation; mHealth; remote diagnosis; referral

Introduction

Telehealth, or mobile-based (mHealth), remote consultation service has the potential to improve maternal and neonatal health outcomes in developing countries such as Bangladesh, where mobile phone subscription rates are very high [1] and there is disparity within the health system in the distribution of trained doctors and qualified medical professionals in urban and rural areas [2,3]. Although Bangladesh has made significant improvements in reducing the maternal mortality ratio by 40% between 2001 and 2010 to 194 per 100,000 live births [4], it is still higher than in other South Asian countries [5]. The neonatal mortality rate (28 per 1000 live births) is also alarmingly high and contributes to almost two-thirds of the total under-five child mortality rate in Bangladesh [6]. Delay in seeking care from a health facility or formal health care provider is a major contributor to maternal and neonatal fatalities [7,8]. Research suggests that delay in seeking care is related to cost, inability to understand the severity of symptoms, and use of informal health care providers [4,7,8]. Furthermore, it is considered culturally inappropriate for women to seek maternal health care from male providers [9], yet female doctors are scarce in remote settings [2,9]. With the potential to address many of the barriers to seeking timely health care, a mobile-based consultation service could be used as a health promotion tool in raising awareness among women and their families in remote settings [10].

Mobile phone consultations between pregnant women and nurses at obstetric call centers were found to be effective in providing triage for determining emergency cases, helping patients manage nonurgent symptoms, and increasing efficiency in emergency departments at local hospitals in selected rural states of North America [11]. Another study in Australia found telephone helplines have potential for improving mother's confidence and sustaining breastfeeding [12]. Evidence of interventions involving direct communication between client and provider is scarce in developing countries and is mostly reported in pilot studies [13,14]. Text and sometimes voice messages are the preferred interventions, seeking moderate positive behavioral outcomes in these settings [15,16] due to cultural contexts and operational challenges [13,17-20]. A pilot study in Southern Malawi reported a decrease in hospital attendance for child-related illnesses among an intervention group who were entitled to text or voice messages and a toll-free hotline service for maternal, neonatal, and child health care (MNCH) [21]. Community-skilled birth attendants who accessed toll-free mobile phones to consult with remote physicians were found to have improved confidence in treating and referring patients and gaining the trust of patients in a clustered randomized controlled trial in a subdistrict of Bangladesh [14]. Pilot interventions involving toll-free communication also allowed pregnant women to initiate calls to their local midwives regarding danger signs of pregnancy and delivery in Zanzibar [22] and Bangladesh [14]. Qualitative evidence from pilot projects in India suggested that 24-hour obstetric mobile phone-based helplines could potentially mitigate delays associated with seeking care, reaching proper health facilities, and receiving treatment [13].

Approximately 95% of urban and 84% of rural households in Bangladesh have access to a single mobile phone [1]. A study in a rural subdistrict reported an increase in household mobile phone ownership from 2% in 2004 to 81% in 2012 [23] indicating a rapid uptake of mobile phone technologies in a decade across the country [1]. Competitive pricing among telecom operators significantly dropped service prices for mobile phone owners and closed the "digital gap" of mobile phone ownership between low and high socioeconomic status (SES) households in Bangladesh; low SES households who were once behind the race were increasingly entering the mobile phone consumer market [24]. Understanding the multifaceted benefits of mobile phones in health care, the Government of Bangladesh adopted mHealth in its national Health Information Strategy to increase the performance of local facilities and access to health by all citizens [25].

In 2007, a private mHealth initiative in Bangladesh led to the development of a 24-hour helpline service run by the largest mobile phone operator, Grameenphone, in collaboration with Telemedicine Reference Centre Limited [26,27]. It initially charged 15 Bangladeshi taka (BDT; equivalent to US $0.18) for the first three minutes and then 5 BDT (US $0.06) for every subsequent minute [27]. In 2016, the helpline was made free to Grameenphone subscribers with additional benefits to support hospitalization costs under the new service name Tonic [28]. Following in the footsteps of Grameenphone, other telecom operators have introduced 24-hour remote counseling services to their subscribers [29]. There are also emergency mobile phone numbers provided for the district (*zilla*) and subdistrict (*upazilla*) public hospitals to reach physicians remotely free of charge [25]. The Ministry of Health and Family Welfare initiated a helpline service, "Shasthya batayan," operated by a private call center, which provided citizens access to a doctor's consultation and ambulance services through their mobile phones; customers were typically charged the usual call rates per minute [30]. However, there is a need for more scientific literature that correlates the effect of these teleconsultation services on the health outcomes of callers.

In addition to teleconsultation services, other mHealth initiatives in Bangladesh have involved the use of mobile phones to educate consumers about preventive health care behaviors and services. One such service, Aponjon, provides information to pregnant women and new mothers regarding pregnancy, delivery, and neonatal and infant health care since 2012 [31,32]. Registered female subscribers are entitled to receive two voice or text messages a week and an optional single message for their family member [31,32]. Based on feedback from subscribers who reported the need to talk to a medically trained doctor over the phone [31], Aponjon expanded its messaging service in 2013 to include consultations between patients and doctors [33,34]. The hotline service, 24*7, is available to

Aponjon subscribers and uses a pool of medically trained, mostly female doctors [33,34]. The hotline service was critically vetted by experts in the field to ensure adherence with the national drug regulatory policy and counseling guidelines before being rolled out [29]. Subscribers are typically charged 2.3 BDT (US $0.03) per minute for the consultation service [34].

Although the Aponjon messaging service has been evaluated [32], the effectiveness of the medical consultation service has not. The voice and text messages show no or minimum effect on delivery and neonatal care practices among subscribers [32], although a quasi-experimental evaluation study [35] suggested that prolonged exposure to messages (6 months) can bring about behavioral changes. Another pilot study tested the feasibility of the nutrition counseling service among pregnant women and new mothers along with a mobile cash transfer system; however, the efficacy of the intervention is yet to be determined [36]. Given that the mobile-based consultation service involves the consumer calling (as opposed to being called) for advice and interaction with a health care professional, the acceptability of the service and effect on health behaviors are likely to be different to that of one-way messaging services. There is a dearth of literature on the effectiveness of mHealth consultation services for maternal and newborn health in resource-limited settings including Bangladesh [16,37]. The objective of this study is to describe the call volumes, the purpose of the calls, the experiences of subscribers, and the perception of doctors who provided consultations through the Aponjon service with a focus on acceptability, usefulness, and health promotion in the setting of Bangladesh.

Methods

Setting

Bangladesh is a country situated in South Asia with a population of more than 150 million in a land size of 147,570 km^2, surrounded by India in the north, west, and east; Myanmar in the southeast; and the Bay of Bengal in the south [38]. Approximately 90% of the population is Muslim, 9% Hindu, and the remaining 1% are other religions such as Buddhists and Christians [39,40]. The economy of Bangladesh is rapidly progressing; the gross domestic product growth has been significant at a rate of 6% for the last decade, which has led Bangladesh to achieve lower middle-income country status in 2015 [41]. Administratively, the country is divided into 8 divisions, which are further divided into 64 districts (*zillas*) and 545 subdistricts (*upazillas*) [42]. The subdistricts, or *upazillas*, are further divided into urban areas (consisting of wards; *mohollas* or cluster of households make up each ward) and rural areas (consisting of union *parishads*; *mouzas* or clusters of villages make up each union *parishad*) [38].

Approximately 75% of the population resides in rural areas, although there is a trend of rapid urbanization [38]. The Ministry of Health and Family Welfare has the key responsibility to provide universal health coverage to urban and rural populations and operates a dual system of general health and family planning services through two Directorates General of Health Services and Family Planning in district hospitals, *upazilla* health complexes at the subdistrict level, union health and family welfare centers at the union level, and community clinics at the ward level [3]. Apart from government initiatives, private sector (formal and informal) and nongovernment organizations (NGOs) have established a network of facilities to provide health and family planning services [3]. There is no structured referral system or national health insurance policy; health services are mostly financed by each households' out-of-pocket payment system [3]. Treatment facilities and medicine at public health facilities are subsidized; however, the quality of service has largely been criticized for the shortage and inequitable distribution of skilled health professionals in rural areas, lack of proper equipment, staff absenteeism, and governance issues [2,3]. Private facilities that have major state-of-the-art treatment equipment are urban-centric, costly, and remain out of reach of poor people [2,3].

Aponjon Service

The Aponjon (meaning "the close or dear one" in Bangla) service was rolled out phase by phase across Bangladesh after being initially launched in 2012 [31,32]. The service underwent a year-long pilot before launching. Aponjon collaborated with the Government of Bangladesh and local NGOs to scale the service and recruit subscribers across the country [33]. The service is available through all telecommunication operators in the country [32].

The core service includes a messaging service—both interactive voice responses (IVRs) and text messages [31,32]. Typically, IVR messages are 1 minute long; the contents are structured around a female "doctor" (called *daktar apa*), who advises pregnant women, new mothers, and their husbands or caregivers on various aspects of pregnancy, delivery, postpartum and newborn care, illness symptoms, and nutrition [31]. The text messages were transliterated as Unicode and were not supported by basic mobile phones, resulting in it being popular among subscribers with higher education [31]. However, IVRs also became very popular among rural subscribers; 99% of them were receiving IVRs [32]. By mid-2014 the service had one million registered subscribers. Any pregnant woman can register in the service from 6 weeks of pregnancy if they have access to a mobile phone [31]. A potential subscriber may enroll her own phone or a phone that she shares with her husband or family members [31]. IVR messages are scheduled to be sent to subscribers according to their preferred time and days of the week to decrease the chance of them being missed [31]. Missed IVR messages can be retrieved by calling the service [31]. By the end of pregnancy, pregnant women are able to update their service to that for a new mother and receive messages for baby and mother until the child turns 1 year of age [31]. A new mother who did not receive the service for pregnancy can enroll in the new mother's service [31]. Prospective subscribers are recruited to the service in several ways. The most popular way is through the thousands of community health workers [31] working under the flagship of different NGOs in Bangladesh [43]; they register interested women in the service during their monthly door-to-door visits for antenatal and postpartum care. Pregnant women and new mothers can also enroll with the help of paramedics or doctors at health centers [31]. The other option is to directly call the service for enrollment; the short code of

the service had been popularized through advertisements in television, billboards, leaflets, and newspapers [31,32].

Since mid-2013, Aponjon subscribers can additionally talk to a doctor by dialing the service short code [34]. A pool of medically trained doctors, located in a central call center of Aponjon service in Dhaka city, handle calls regarding medical queries from subscriber women from across Bangladesh [34]. The purpose of the consultation service is to support subscribers to understand the severity of symptoms, help in decision making during emergencies, and improve understanding of maternal, neonatal, and infant health care practices at home through one-to-one consultation.

The Aponjon service has a pro-poor policy [31,32]; the socioeconomic data gathered during enrollment are assessed to consider the eligibility of marginalized subscribers to receive the messages free of charge. The rest of the subscribers pay 2.3 BDT (US$ 0.03) per message [32]. Aponjon subscribers accessing the service under any telecom operator will have similar pricing experiences. The “free of charge” policy is not applicable when subscribers access the consultation service; they pay the usual call rate for outgoing calls.

Study Design

We conducted a mixed methods study. In September 2015, the service had approximately 1,135,839 subscribers from across Bangladesh. Anyone registered in the service could access the 24*7 consultation service. We analyzed the call log data of the consultation service. We also conducted qualitative interviews with subscribers who contacted the consultation service and with Aponjon doctors who provided consultations.

All participants provided written and verbal consent for this study. The study received ethics approval from the Science and Medical Delegated Ethics Research Committee, the Australian National University, along with permission from Dnet, the implementing agency of Aponjon in Bangladesh.

Quantitative Data Collection

Each day, the Aponjon consultation service received an average of 150 calls from existing subscriber women’s cell numbers. All calls were logged on the service “customer relationship management” (CRM) database including date, time, callers’ mobile phone numbers, location, call duration, type of caller (subscriber woman/husband/family member), severity of the condition for which advice was sought (general, semiurgent, and urgent), and type of advice provided by the doctors. For the quantitative study, we retrospectively accessed the CRM database and extracted data for the period of September 1 to November 30, 2015.

Qualitative Data Collection

All interviews were conducted between December 2015 and January 2016. For this study, we accessed the CRM database for calls made in the 3 months prior to data collection to reduce users’ recall bias. Only subscriber women and their husbands who made calls to the Aponjon consultation service in the past 3 months for maternal-, neonatal-, and infant health care-related queries were eligible for interviews. Exclusion criteria included callers who made calls for nonmedical purposes, who were younger than 18 years of age, and family members other than husbands. From the CRM database, we identified unique callers, prepared a sampling frame for two different types of callers (subscriber women or husbands), and systematically called every fifth caller. Each mobile phone was called a maximum of two times; we reached a caller almost 40% of the time (after making 81 call attempts to 45 phone numbers). Of the contacted callers, 18 agreed to participate in the study, and we arranged a convenient time for phone interviews. Two potential participants were excluded because they did not fit the entry criteria. Overall, 16 different families—eight women subscribers and eight husbands of subscribers—were recruited. Recruitment continued until data saturation was achieved, meaning that very little new relevant information was likely to be gained from further interviews with callers [44,45].

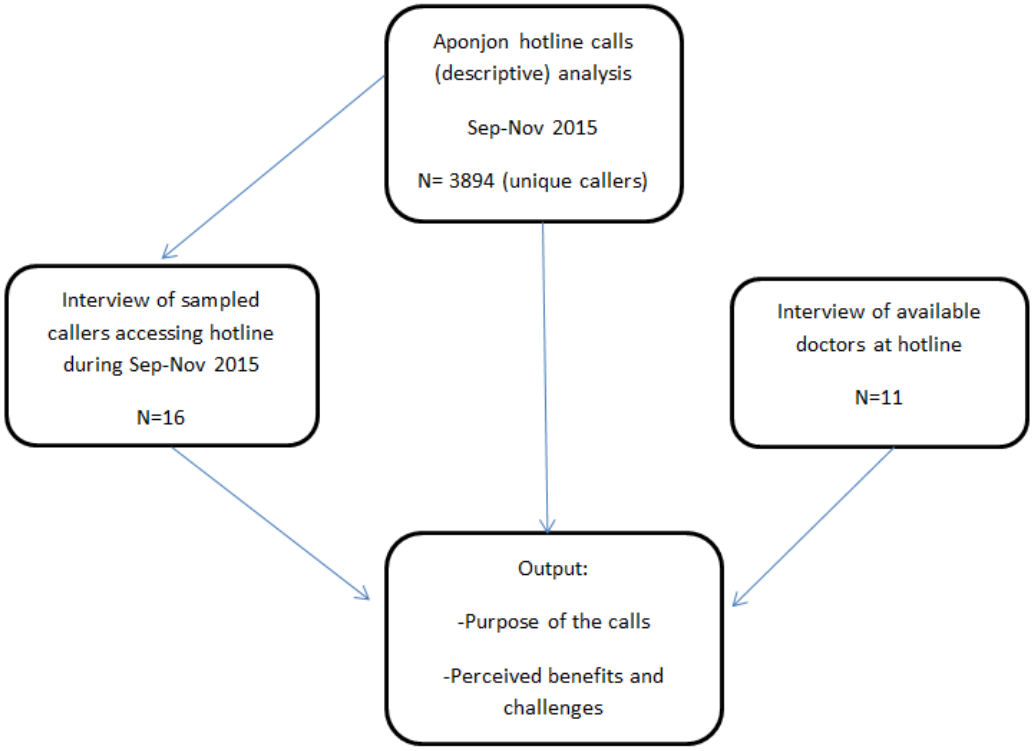
Recruitment of participant doctors was supported by Aponjon management. At the time of data collection, 16 medical doctors (12 females and 4 males) were registered for roster duties at the consultation service for morning, day, or night shifts. All doctors were contacted and informed about the research through Aponjon management. Among them, 11 doctors (9 females and 2 males) were available before or after their roster at the call center and were interviewed. The remaining doctors who could not be interviewed did not have a roster or were on leave.

We prepared interview guides for consumers (callers) and doctors which contained open-ended questions. Consumer participants were read the purpose of the research over the phone, and they gave verbal consent to participate and have their interviews recorded. Each phone interview lasted between 25 and 35 minutes. Participants were asked about their experience using the Aponjon hotline service—why they used it; the benefits, costs, and limitations of the service; and access to other health care providers. Participants’ socioeconomic information, such as age, education, monthly income, birth order, family size, occupation, and place of residence, was collected. Participants were also asked about how they individually accessed mobile phones, whether they had the capacity to receive or make calls, read, and send text messages without any help. Additionally, husbands (n=8) were asked how their wives accessed mobile phones.

All the doctors’ interviews were conducted face-to-face at the call center before or after their rosters. The doctors read the participation sheet before giving signed consent to participate and have their interviews recorded. Doctors were asked about their motivation to work in consultation service under mHealth initiatives, the type of advice they provided at Aponjon, callers’ expectations, and the benefits and barriers to a mobile-based consulting service. The interviews lasted between 35 and 50 minutes.

Author MA conducted all interviews in Bangla. All participants (callers and doctors) were given their rights to not answer any question or withdraw from the research at any point of time. Identities of all participants were masked by IDs.

Figure 1. Mixed methods data triangulation plan.



Data Analysis

The CRM database for the targeted period was analyzed with statistical software SPSS version 24.0 using a descriptive data analysis method. We ran frequency tests to describe patterns of call duration, type of callers, type of calls based on severity, and timing of calls; data were presented as frequencies, means, modes (where applicable), and percentages.

Professional transcribers and the author MA shared transcriptions of the interviews. All transcripts were converted to Bangla text. Author MA checked the recordings and transcripts. Transcripts were translated into English by author MA and then coded using Atlas Ti software for storage and to assist with analysis [35]. The transcripts were repeatedly read by authors to compare the doctors' and clients' experiences. A thematic analysis approach was used for identifying, analyzing, and reporting patterns (themes) within data [46]. Using this approach, a number of smaller subthemes [46] were identified inductively and then organized under two key metathemes identified by the authors: perceived benefits and perceived challenges. These two metathemes were developed after conducting a comparative analysis of concepts coded in different participant groups (participant callers and doctors). Codes and statements from participant groups were compared to assess positive and negative perceptions of the service [47], and were entered in a matrix. These similar perceptions from callers and doctors form the basis of the subthemes and are illustrated by quotes from both callers and doctors. A diagram (Figure 1) is provided on data triangulation for this study.

Results

Quantitative Findings

Who Were the Callers?

Approximately 3894 unique subscribers accessed the service during the 3 months for single or multiple consultations. There were three different caller groups based on who was making the calls from a single registered mobile number. There were 53.00% (2064/3894) who were subscriber women in the first group, who always made calls by themselves and consulted doctors directly. The second group contained callers other than subscriber women; 33.98% (1323/3894) of calls were made by husbands or family members who consulted doctors for subscriber women's or babies' symptoms. The third group, 13.02% (507/3894), were mixed callers (both subscriber women and husband or family members) from the same phone, inferring subscriber women shared mobile phones with family members.

The callers came from diverse areas across the country; call logs suggest that the service was accessed by callers from eight (8/13) districts in Dhaka, seven (7/11) districts in Chattogram, four (4/6) districts in Barisal, six (6/10) districts in Khulna, four (4/8) districts in Rajshahi, four (4/8) districts in Rangpur, one (1/4) district in Mymensingh, and four (4/4) districts in Sylhet division. In all, 68.36% (2662/3894) of calls came from rural households (Table 1).

Table 1. Caller type and location (N=3894).

Criteria	n (%)
Who called	
Only subscriber women (pregnant/new mother)	2064 (53.00)
Subscriber women and husband/family member	507 (13.02)
Husband/family member	1323 (33.98)
Place of residence	
Rural	2662 (68.36)
Urban	1232 (31.64)

Table 2. Category of calls (N=5288).

Category of calls	Subject of queries about neonates and infants	Subject of queries about pregnant women	Subject of queries about new mothers	n (%)
General	Queries on general cold, fever, nutrition, neonatal and infant health care, development milestones, growth, play, immunization, and medicine; primary diagnosis of neonatal and infant illnesses	Queries on pregnancy health care, nutrition, fetal development, delivery, general cold, lifestyle, and medication; primary symptoms of pregnancy discomfort and illnesses	Queries on family planning methods, maternal health, reproductive health, and breastfeeding; primary symptoms of reproductive health and postdelivery complications	5081 (96.08)
Semiurgent	Neonatal and infant diseases, infection, birth-related injury, and developmental delays	Pregnancy and delivery-related danger signs	Maternal health issues	142 (2.69)
Urgent	Neonatal and infant illnesses, developmental problems, and birth-related injury	Pregnancy and delivery-related danger signs	Maternal health issues	65 (1.23)

Type of Calls

A total of 13,101 incoming calls were handled during this period. The minimum call duration for a complete conversation was recorded at 1 minute 40 seconds and the maximum at approximately 16 minutes.

The problems stated by the patients were triaged into nonurgent, semiurgent, and urgent categories by the doctors depending on the severity of symptoms. Complete data on the problem statement, advice, and related notes by the doctor were available for 5288 calls. From the available data, most calls were nonurgent (5081/5288, 96.08%); callers were seeking general advice on nutrition, growth, development on pregnancy, delivery, neonatal and infant health, and postpartum maternal health. The rest of the calls were made for semiurgent (142/5288, 2.69%) and urgent conditions (65/5288, 1.23%). Nonurgent calls were provided with self-care advice, prescriptions, appointment advice with local physicians, or referral to local facilities. Semiurgent and urgent calls required an emergency visit to local health facilities for labor, delivery, and newborn illnesses. A schematic guideline of the type of calls is provided in Table 2.

The call center received most calls for urgent or semiurgent conditions at night. Out of 207 such calls, 64.7% (134/207) calls came between 8 pm and 8 am (Table 3). Almost 56.1% (116/207) of late night calls were made by husbands and 43.9% (91/207) by pregnant women or new mothers.

Table 3. Morphology of emergency (urgent and semiurgent) calls (N=207).

Criteria	Urgent (n=65), n	Semiurgent (n=142), n	Total (N=207), n (%)
Call time			
8 am-2 pm	18	15	33 (16.0)
2 pm-8 pm	17	23	40 (19.3)
8 pm-8 am	30	104	134 (64.7)
Who called			
Pregnant woman	19	21	40 (19.3)
New mother	5	46	51 (24.6)
Husband	41	75	116 (56.1)
Purpose of the calls			
Pregnancy or delivery related	38	63	101 (48.8)
New born health	24	75	99 (47.8)
Maternal health (postdelivery)	3	4	7 (3.4)

Call Interruption, Missed Calls, and Preferences

The CRM data for the study period indicated that 152 calls were ended by callers before the doctors could finish providing their advice, and 64 calls were disrupted due to a poor network. There were 342 missed calls reported in the call notes in which callers cut the call after one ring.

Special notes suggest that there were three incidents in which the subscriber called back to update on prognosis, one incident in which the female subscriber asked to talk to a female doctor, and one incident in which the female subscriber wanted to talk to the doctor she previously consulted.

Qualitative Findings

The callers, eight women and eight husbands representing 16 families (Table 4), came from diverse geographic locations in Bangladesh; they were a mix of rural (12/16) and urban residents (4/16). Most of the participating families were middle income. The women callers (Table 4) were mainly younger than 25 years (5/8); two were in their late teens (18 and 19 years). Educational levels differed between families of female callers and male callers. Female callers and their husbands had higher levels of education (most of them had completed higher secondary education) than male callers and their wives. Women callers reported more autonomy in accessing mobile phones (such as making and receiving calls, texting, and owning a personal phone) than women whose husbands called Aponjon.

Table 4. Background information of participants (Aponjon subscribers) (N=16)

Characteristics	Caller women n (n=8),	Caller husband (n=8), n
Age of participants (years)		
<20	2	0
20-24	5	1
≥25	1	7
Access to Aponjon messages		
Regular	7	5
Sometimes	1	2
Never	0	1
Type of messages received		
Voice	8	5
text	0	3
Who received Aponjon messages		
Husband	0	3
Wife	8	0
Both	0	4
Never received messages	0	1
Aponjon subscriber woman possesses a functional personal mobile phone		
Yes	8	2
No	0	6
Mobile phone literacy of Aponjon subscriber women		
Can receive and make calls		
Yes	8	5
No	0	3
Can send/read text messages		
Yes	8	3
No	0	5
Educational qualification of subscriber women/wives		
None or primary education	0	3
Junior secondary education	2	0
Secondary school or higher	6	4
Don’t know	0	1
Educational qualification (husbands)		
None or primary education	0	3
Junior secondary education	0	0
Secondary school or higher	6	5
Don’t know	2	0
Family monthly income (BDT)		
<10,000	0	1
10,000-50,000	8	6
>50,000	0	1
Occupation of woman		
Homemaker	6	8

Characteristics	Caller women n (n=8),	Caller husband (n=8), n
Teacher	1	0
Student	1	0
Occupation of husband		
Overseas work	4	0
Unemployed	0	1
Small or large business	3	5
Other paid jobs	1	2
Family type		
Nuclear	1	1
Extended	6	7
Stays with parents (husband overseas)	1	0
Number of children		
Expecting first child	2	0
1	3	5
2	3	1
3 or more	0	2
History of miscarriage/stillbirth/abortion		
Miscarriage/abortion	1	1
Child death (obstructed delivery)	1	0
Division		
Dhaka	5	3
Rangpur	1	0
Rajshahi	0	3
Chittagong	2	0
Khulna	0	2

Table 5. Demographics of participant physicians (N=11).

Characteristics	Physicians
Gender, n	
Female	9
Male	2
Age (years)	
Median (IQR)	28 (3)
Mean (SD)	28 (2)
Range	25-32
Duty hours per week at call center	
Median (IQR)	21 (7)
Mean (SD)	18.9 (4.7)
Range	11-28
Experience with Aponjon service, n	
<3 months	2
>6 months	1
>1 year	8
Experience after graduation (years), n	
1-3	2
3-5	5
>5	4
Special training, n	
Gynecology	4
Pediatrics	5
Medicine	2

Unlike a previous study [32], in which all Aponjon subscribers were enrolled in the service through health workers, only one-third (5/16) of participants were recruited to the service via this method. The remaining participants enrolled via a nurse or paramedics (6/16) during antenatal visits at hospitals and by self-enrollment (5/16) after seeing advertisements on television or in the newspaper.

All participating doctors (n=11) were medically trained and had received training in gynecology and pediatrics. The doctors were young with few years of experience after graduating from medical school (Table 5). In addition to working at the call center, they also worked at private and public health facilities.

Theme 1: Perceived Benefits

Access

All callers were satisfied with the consultation service because they could access a trained doctor any time they wanted. Women who were able to make calls on their own found the service convenient to access medical advice without relying on male members to organize a visit to a general practitioner (GP). Women had a sense of relief because they could discuss their pregnancy stages, newborn care practices, and symptoms of illnesses with Aponjon doctors. One teenaged new mother accessed the service frequently to discuss newborn care and disease symptoms. The doctors also reported receiving calls from young women and their husbands. The customer-centric approach of the call center encouraged doctors to use friendly greetings, to listen to queries with patience, and to provide practical advice. There was a policy that the "customers hang up calls, not the doctors," which encouraged callers to contact the service without any reservation. The doctors reported that they could allocate more time to patients' problems at call centers, which was impossible at crowded outdoor services at health facilities. Some doctors thought mobile consultation encouraged shy women to discuss their reproductive health. One doctor reported:

> *There are women who do not want to visit a doctor; they feel shy. But since we don't see the patients face-to-face some women think their privacy is protected and they share their private problems with us. I have received many calls like that. When we receive the calls, we tell our name like "I am Doctor X." Women who use the service regularly remember our names and sometimes would ask for the same doctor* [Doctor 6]

One anxious pregnant woman illustrated how the service gave her ready access to reassurance and information:

> *I had a miscarriage before this one, so naturally I am a bit tensed all the time and I want to clarify the symptoms with a doctor every now and then. I feel very low when my Gyne doesn't pick up the call or if she hurries through my calls. She won't suggest anything on phone, and asks me to visit her chamber right at that moment, which is not always possible especially if it is late at night. So, I call Aponjon doctors whenever I am feeling unwell. They listen to me patiently, and they advise me whether my conditions are severe. If they ask me to see my Gyne or visit a hospital for a check-up, I do that without any delay.* [Caller 8, rural pregnant woman]

The service was convenient for rural people living in remote areas who otherwise had two options: travel for several hours to discuss a health-related concern with a qualified doctor at a facility or consult the local drugstore salesman who has no professional qualifications. Husbands who were unable to allocate time, money, or leave from work to accompany their wives to health facilities thought a mobile phone-based consultation service was a convenient option for initial diagnosis. Sometimes husbands would organize phone calls so that their wives could talk to the doctors. One husband stated:

> *I have been accessing this service for almost a year now since the beginning of my wife's pregnancy. I often make a call to the service, I even called yesterday. It is difficult to find doctors (for a visit) all the time, and I work in the [...] department. So, I cannot always make time to take my wife to see a doctor. It takes a lot of time-waiting in the queue, traveling, and the fee [private doctor's] is between 300 and 400 taka [US$ 3.56 to 4.74] which is costly for us. So, I like this service for primary consultation. It's very useful as we can call anytime and get good advice.* [Caller 3, rural father]

The doctors reported receiving calls from both disadvantaged rural community and urban households, although more came from rural Bangladesh. This is not surprising as Aponjon recruited a large proportion of its client base through rural community health workers. The doctors considered their job rewarding and felt that they were useful in availing themselves to rural patients.

> *I have received calls from poorest of the poor (rural) to bankers (urban). But most of the time the callers are low-income people who are located in very remote places and live quite far away from health facilities or upazilla [subdistrict] health complexes. I think this [call from low-income households in rural areas] is because most of the recruitments to the service were done by NGO health workers. They [patients] are able to know where they have to go or what to do by making a call to us. Most of these calls are nonurgent, but we get urgent calls too mostly at night. I feel very happy at the end of the day that I can be of some help to rural people who need it most.* [Doctor 11]

The option to talk to a doctor at the call center regarding maternal and newborn health was advertised on TV channels and newspapers, which interested some participants to enroll in the service. Availability of female doctors at the call center encouraged women callers to access the consultation service; some participant women would rather wait for female doctors to be rostered on if the problem was not acute. The service was useful to clients at night time when visiting a doctor was not possible. The doctors reported that they were getting increased calls regarding problems that required urgent attention mostly late at night. One such incident noted by a participant woman:

> *We were very worried one night when my son [infant] became ill. My husband contacted Aponjon. At that time the doctor gave us a good solution which improved his condition. In the morning we took him to a hospital.* [Caller 6, rural mother]

Trust

Participating callers explained that they considered Aponjon's consultation service to be a trusted source of medical information and clinical advice. Aponjon doctors confirmed that they graded the severity of the illnesses after careful consideration of the symptoms. Often, on request of the women who made the call, they motivated reluctant family members to seek treatment from health facilities. The doctors reported that they received many calls in which callers sought a "second opinion" from them regarding cesarean delivery or other treatment at hospitals.

One doctor stated:

> *I once referred a mother for further investigation to nearby medical hospital; her baby suffered from brain injury during delivery which required proper treatment to avoid impairment. Doctors at Mymensingh Medical College Hospital referred her to Centre for Rehabilitation of Paralyzed in Savar, specialized for this treatment. But the mother was in a dilemma and called me from the hospital for assurance.* [Doctor 9]

Aponjon's mobile phone-based consultation service offered a convenient medium to families through which they could demystify cultural myths and practices with someone who they believed to have superior knowledge and medical training—*pash kora daktar* with a Bachelor of Medicine, Bachelor of Surgery (MBBS) degree. Doctors reported receiving calls from women, their husbands, and even mothers-in-law who sought advice on the care of pregnant women, lactating mothers, and newborn babies. Mothers-in-law are powerful household members who have considerable influence on the birthing and health care practices of their daughters-in-law. For example, some female callers (Aponjon subscribers) rang on behalf of their mothers-in-law to inquire about commonly practiced rituals, such as rubbing mustard oil on the newborn baby's umbilical stump. "I could overhear the mother-in-law asking her [daughter-in-law] to ask me these questions," said a doctor. Some anxious mothers-in-law called to inquire about nutrition for pregnant women, how to increase the production of breast milk, or the illness symptoms of the mother and infant. The doctors supported new fathers, who then felt empowered to

resist the traditional practices of elderly family members. One new father reported:

> *The doctor told me not to feed newborn babies mustard oil. He also told me that we should not perform shak dewa [vapor treatment] around umbilical stump of newborn baby. That was so helpful! You know how murubbi [elderly family members] ask us to do that. I prevented them [from performing on my newborn]. Instead, we applied a powder [chlorhexidine] which the doctor told me over phone.* [Caller 2, rural father]

Satisfied husbands wanted to expand the service beyond the current scope (pregnancy, postpartum maternal health, newborn and infant health) and expressed an intention to use the consultation service for both men's and women's health in all age groups.

Cost

The service price (2.3 taka per minute or US $0.03) did not seem to be a barrier to access direct consultation service for the participant callers; they were willing to pay for the consultation service. As this participant suggests, the cost of the service is lower than the cost of attending a face-to-face medical consultation:

> *We [he and his wife] practically are bokolom [illiterate]. They send us recorded messages, we both listen.The consultation service with doctors through mobile phone is very helpful for us. They give us direct advice. I cannot close my shop and spend time traveling and waiting to see a doctor all the time. Every day I sell products worth 1000 to 1200 taka [US $11.85 to $14.22], my profit remains 100 to 150 taka [US $1.19 to $1.78] and with that money the four of us have to survive. How can I close my shop for a moment, tell me?...I don't find the service expensive. If you [Aponjon] need money, tell me, I will flexi [mobile cash] some money [to Aponjon], no problem.* [Caller 1, rural father]

Another participant weighed the cost against the value of a child's health:

> *I don't think the charge [for consultation service] is too much. You can't ignore your child's health, right? I visit my baby's pediatrician quite often who is available 2 days a week and located only 10 to 15 minutes away from my home. But still I think there are so many queries I should ask, and I prefer calling Aponjon. I don't have to book tickets and go through all the hassles [at hospitals] for phone consultations.* [Caller 16, urban mother]

These quotations suggest that participants made thoughtful decisions about cost by considering a number of elements.

Raising Awareness and Assisting with Decision Making

Some participants found the consultation service to be very useful besides the weekly messages they received; they had the scope to discuss their confusion regarding any messages with the doctors. Doctors reported receiving calls from patients requesting clarification on weekly messages. Some participants reported problems with receiving messages due to a network problem or inadequate balance issues in their mobile phones; they thought calling the doctors was always a good option to catch up with what they had missed. One participant husband stated:

> *I tell my friends about the service. We get shundor shundor [nice] messages every week from this service. My wife and I both receive the messages and have learnt new information. The messages have made us conscious about a lot of things. And then whenever we have a query regarding our baby I straightaway call and talk to the [Aponjon] doctor who is again helping us with proper advice and (sometimes) medication. I can share everything [with the doctor] which is a great thing.* [Caller 9, urban father]

Interviews with both the callers and doctors suggest that doctors were able to raise awareness about the dangerous signs of pregnancy, delivery, and neonatal illnesses. Doctors reported a common "inertia of families toward seeking help from outside" and in response worked to educate callers on a range of issues, such as routine antenatal checks, counting of fetal movements, and symptoms of eclampsia (one of the most prominent causes of maternal mortality in Bangladesh) [4].

> *My wife is expecting our fourth child...it was not a planned pregnancy...our first child is an adult. Never before [for other pregnancies], my wife had an ultrasound, but this time when Aponjon doctor told us to do that, we performed the test. [Now] we know the baby is okay [inside the womb] and we also know the expected date of delivery.* [Caller 17, rural father]

The doctors reportedly interfered with callers' decisions to seek care from untrained providers, such as the local drugstore salesman or traditional healers. One incident reported by a doctor:

> *Once a newborn baby was having convulsions, and the family decided to take the baby to a Kabiraj [nonformal provider] to perform jhar-phoonk [rituals believed to have magical effects] which is harmful [to the baby]. The father of the child called Aponjon...I asked him to remove the patient [child] to a hospital immediately. I explained the illness of the child and he understood it.* [Doctor 8]

A telephone consultation with doctors during an emergency such as a prelabor rupture of a membrane (water break) or prolonged labor can help family members make a safer decision and ultimately save the life of the mother and the unborn child. One such incident reported by a doctor magnifies the need for such guidance during emergencies:

> *My shift was almost over that night when the husband called. He was confused as the village doctor gave saline [at home] and left saying that in the morning everything will be fine [a vaginal delivery]...I suspected it could be a case of an obstructed labor and instructed him to remove his wife to a hospital that night...We later received a call from the husband*

> *that an emergency cesarean procedure had saved his wife and baby. The doctors at the hospital said he was just in time.* [Doctor 4]

Authentic Prescribing

Callers from both urban and rural households reported purchasing medicine for simple or severe complications from local drugstores. Drugstore salespeople are typically unqualified allopaths with no formal degree, having learned their trade via apprenticeship with a qualified or unqualified practitioner [48]. Medication dispensing is often negotiated depending on the severity of the complaint, the economic situation of the patient, and the range of available medicines [48]. Aponjon doctors reported alerting callers to the inappropriate use of medications. Additionally, the service has the capacity to provide callers with prescriptions for over-the-counter drugs if required. The doctors took great care to spell the names of prescribed medicines correctly. Usually, if the patients were unable to follow the doctor's instruction, they sought help from a literate person in the family to write down the names of medicines. Sometimes patients would call the doctors from the pharmacy so that the doctors could talk directly to the salesperson about the particular medicine. One doctor reported:

> *Patients expect that we will give them solutions for everything through medicines. They want names of medicines which they can purchase from pharmacies and get cured without visiting health facilities. We do not support that. As we cannot see the patients, we provide advice and prescribe over-the-counter drugs if necessary. We ask them to notify us about their condition within a few days. If the condition persists, we ask them to visit local health facilities.* [Doctor 5]

One participating mother explained the importance of receiving trained medical advice:

> *My son is only 4 months. There is no pediatrician in my area. So, I saw a provider who gave three medicines. Then I called to Aponjon and told the doctor the name of the medicines. The doctor asked me not to follow those medicines as my son is too young and insisted to see a pediatrician. I did not give the medicine [and] saw a pediatrician from the district hospital.* [Caller 4, teenaged rural mother]

Theme 2: Perceived Challenges

Logistical Issues

The doctors and callers both mentioned problems with consultations when the network dropped out. For example, an anxious husband reportedly was desperately trying to contact a doctor when his wife's labor progressed, but "the calls failed again and again." Apart from network issues, the call center line can be busy forcing callers to try several times:

> *Overall, I like the service. For consultations, I spend approximately 20 taka [US $0.24]. Sometimes getting the line through to the doctors takes time. There is network problem, and then often the lines are busy. But still, I make the calls.* [Caller 12, rural mother]

Equal Access

As noted previously, most participants considered the service call rate (2.3 taka or US $0.03 per minute) "cheap." However, we do not know if the service was inaccessible to low-income households. Even among households who could afford the service, doctors reported that the quality of consultation was compromised when clients kept the consultation time short to manage the expense. One doctor reported:

> *Usually, most patients want us to explain the cause of their symptoms and often the discussion can lead up to 15 to 20 minutes depending on the patient's query. Sometimes the network can be problematic; the phone line can get disrupted. Clients call back and finish the discussion. But some clients cannot afford long conversations; they are in a rush. They don't even want us to greet them, they straight jump to their problems and want us to give them a solution in the quickest possible time. They tell us "hurry up" [taratari bolen]. Sometimes before we have finished, they just hang up. We cannot even record their case history properly.* [Doctor 1]

Visual Examination

The doctors reported that their foremost challenge in virtual consultation was the difficulty diagnosing symptoms without the ability to conduct a visual examination. For example, the doctors could not verify whether a person actually had "jaundice." They often had to innovate ways to understand the temperature of a "feverish" child over the phone if the parent did not have a thermometer or was not confident in using one. Doctors sometimes faced challenges understanding patients' colloquial language and had to spend time seeking clearer descriptions of symptoms. Similarly, prescription of over-the-counter drugs was difficult without knowing the baby's weight and checking for other fatal illnesses such as pneumonia. As one doctor said, "We cannot always rely on how they explain the symptoms. What is a 'mild' condition according to them can be 'severe' in our diagnosis." Consequently, referral to a health facility for diagnosis was a common outcome of calls.

Referral and Follow Ups

Aponjon doctors identified the lack of connection between the consultation service and available care and emergency services as a key weakness in the service. The doctors were unable to follow up with patients about referrals and expressed concerns about whether patients could navigate the right care at the right facilities. Indeed, according to patient callers, delay in finding the right care often varied depending on the distance between their residence and a quality health facility, their financial situation, and social connections. For example, a female rural participant had to organize overnight stays at her parents' house in the city when she needed to visit a pediatrician at a public hospital. Similarly, another participant, a first-time father, had relocated his pregnant wife from his ancestral village home to his residence near his workplace because his work colleagues had alerted him to the availability of quality emergency obstetric services, which were hard to find in rural areas. A doctor also reported his experience of receiving patients at urban facilities

with much worse symptoms due to a significant delay in finding the right care:

> *I have seen parents from remote areas seeking treatment at private hospitals for neonates with severe illnesses. They travel from very remote areas to the city. The neonates get admitted at neonatal intensive care [at the private hospital]. What can those [rural] parents do? The local [public] facilities probably don't have the support.* [Doctor 2]

The service also lacked a policy to follow up on how patients were complying with medical advice at home regarding issues such as prescriptions. Doctors stressed the need for a database to track patients' case histories, prognoses, and referral outcomes.

Discussion

We investigated the access, acceptability, usability, benefits, and challenges of a mHealth consultation service in Bangladesh. Our research is novel because we have combined analyses from both call log data and perceptions of patients and providers. We analyzed call center data on frequency, duration, urgency level, and advice sought by subscriber callers of a mHealth consultation service who also accessed the regular messaging service. We also conducted in-depth interviews with subscriber callers and doctors of the mHealth consultation service regarding their perceptions, and the perceived benefits and barriers in using such a service for maternal, neonatal, and infant health. Our analyses of call log data and qualitative interviews are complementary. Overall, our findings suggest that the consultation service was trusted among callers and contributed to significant changes in health care-seeking behavior (Table 6). The call log data suggest that more than half the calls were from subscriber women (pregnant women or new mothers) and that majority of all callers were rural residents. This suggests that mobile phones may help address culturally determined gender norms in Bangladesh by enabling "shy" women to discuss their personal health issues with female doctors who are hard to find in rural communities [2,49]. The service may lead to a paradigm shift in how women seek health care. Young married women in low-income households are generally guided by their husbands and mothers-in-law, who may deter them from accessing skilled maternal health care services [49,50]. This research shows that some women were reluctant to visit maternal health care services due to distance, cost, fear of hospitals, privacy concerns, unavailability of female physicians in local facilities, need for a male chaperone, and being refused approval of husbands or mothers-in-law who thought the symptoms were not so severe or could be cured at home [9,49]. We find that Aponjon's remote consultation service has the potential to empower women in rural communities to discuss their queries regarding maternal, neonatal, and infant health. Findings from qualitative interviews suggest almost all participant caller women and their husbands found the consultation service convenient, accessible, and cost-effective for any medical advice and primary diagnosis. Trust in the service encouraged husbands to make phones available to their wives to consult Aponjon physicians; a step toward reducing the gender gap in women's ownership and access to phones in Bangladesh [51]. Our findings provide useful evidence for mHealth implementers in similar socioeconomic and cultural settings of other South Asian countries, such as Nepal, India, and Pakistan [52,53].

This study suggests that the mHealth comprehensive system of messages and the consultation service is moving in the right direction to respond to the growing consciousness of maternal and newborn's health and aligns well with other national interventions to improve infant health and well-being [3,6]. Talking to a doctor by phone is a plausible solution for subscribers with limited mobile phone literacy who may find it difficult to retrieve missed IVR messages [20,58].

Remote mobile consultation may improve people's decision making during emergencies and reduce delay in reaching health facilities [10,59]. Call center data and qualitative interviews suggest that the real-time advice was instrumental in helping families understand the graveness of symptoms and encouraging them to seek immediate help, especially at night when visiting a doctor is technically impossible. The Aponjon consultation service was able to influence mothers-in-law and husbands, particularly, in advising the family to attend a medical service rather than a traditional healer.

Table 6. Identified benefits and ways forward.

Benefits	Way forward
The service was available 24*7, especially at night	Since there is a need for consultation late at night when physical visits to doctor is not possible, Aponjon service could open the consultation service to non-Aponjon subscribers (pregnant women and new mothers)
Women could talk with female doctors	During enrollment, community health workers or call center agents should inform women about the option to talk to a female doctor regarding symptoms and illnesses
Consultation service was cost-effective because it saved time, money, and absence from work	Each consultation should be documented in electronic medical records to improve consultation time and quality in following sessions; the service may include electronic prescriptions which patients can access through Aponjon apps [54]
Advice of doctors relieved pregnant women and new mothers of anxiety; dispelled myths and wrong practices	The service may extend counseling services to women who experienced miscarriage, stillbirth, and postpartum depression with accredited psychologists [55]
The service helped families to understand graveness of symptoms and making decisions during emergencies	The consultation service needs to improve referral system linked with local facilities so that the families may seek treatment at local facilities without delay [56]
Doctors were instrumental in raising awareness about unnecessary medication from unauthorized personnel	Drug Procurement Act [57] should be strengthened so that purchase of antibiotics and other medicines without a prescription is prohibited

Table 7. Identified challenges and recommendations.

Challenges	Recommendations
Consultation gets disrupted due to network problem	There is no immediate solution for such disruption, while continued investment in infrastructure of telecommunication network will improve the quality of calls in future
Talk time is expensive for the poor	Aponjon may create a toll-free option for selected communities where subscribers have low wealth index [14]; Aponjon may also explore other ideas such as mobile cash transfer to marginalized customers for talk times [36]
Doctors struggle to diagnose certain symptoms without visual examination	In this context, a training program of local providers, such as village doctors, community health workers, and *Dai* s, should be revived. Since the local providers are the foremost people to be contacted during delivery or any other illnesses, their explanation of symptoms could provide a realistic picture of the condition of the patients to the doctors over phone [14,58]
Referrals of the doctors to seek treatment at health facilities or a specialist service is not specific	A proper integrated system plan is required; the consultation service requires to be linked with existing local facilities; a smooth connection for emergency transport needs to be organized between consultation service and existing facilities; data interoperability between service centers and a national patient database is warranted [61-66]
The consultation service did not follow up patients on medication and referrals	Toll-free follow-up calls and reminder text messages to patients regarding appointments and medication need to be integrated in the service [67]

Despite several benefits recounted by callers and doctors regarding the consultation service, we find that the service has potential to improve and thus have included recommendations (Table 7). In terms of equity, the consultation service may still be too costly for low-income households who reportedly kept conversations short to minimize talk times. Under such circumstances, the current pro-poor policy of the service [31] is not effective in securing access for low-income households free of charge, pushing them to rely on unqualified health professionals for low-cost treatment [2,3]. Customized approaches of previous interventions, such as a toll-free communication line [14] or mobile cash transfer [36], for "talk times" remain options to keep the service affordable to low-income households. The service may build on the successes of the national maternity voucher system in selected districts by increasing utilization of delivery facilities in marginalized communities and providing similar support to subscriber women from low-income households to be able to receive discounted MNCH services at local facilities [60].

Our study findings suggest that the doctors provided triage based on the severity of symptoms, but lacked an operational referral service to support their decisions, leaving patients to their own resources to find a physician or health facility. A study in rural India suggested there was an increase in utilization of MNCH services and growing demand for facility-based services by pregnant women after sessions of mHealth-supported counseling service by community health workers [68]. To respond to similar circumstances, we propose that an m-referral linked with the consultation service is essential to increase the efficacy of Aponjon service. In fact, a well-functioning mHealth service is incomplete without a proper referral service and is undeniably reliant on the capacity and readiness of the existing health system to provide quality care to citizens [69]. The government is progressing major reforms in the health system to ensure access to universal health coverage by all citizens [3], and m-referral can play a pivotal role in these reforms by offering patients real-time information on existing health care and referrals to quality medical care. An integrated referral system could provide routine toll-free follow-up calls to referred patients to ensure patient well-being and adherence to treatments [67]. In particular, the service could assist users from rural low-income households navigate the health system and

potentially avoid the ingrained practice of bribery by poorly trained staff and *dalals* (middlemen) in public health facilities [2,70]. Referrals from Aponjon to specific destinations would require thorough information on the existing capacity of local facilities. For example, adherence with a referral advice was high in a recent pilot study [56] in Bangladesh in which the efficacy of an emergency transport system for referring sepsis patients to nearby local facilities was tested, after assessment by a call center with the aid of a geographic information system. Although this study was conducted among a small population, lessons from this study may guide the Aponjon consultation service in establishing linkages with local health facilities for referrals. The referral system also needs to build a framework for electronic medical health records and a robust platform for data interoperability between consultation services and health facilities, which will allow transfer of patient data such as symptoms, primary assessment, past medical history, findings, treatment, medication, care plan, special notes, and immunization [61,62]. The need for a national database for patient electronic health records is indispensable to accommodate heterogeneous patient data from different sources. This will improve the referral system, treatment, and allocation of medical professionals accordingly, and enhance the prospects of longitudinal studies evaluating disease causation and prevention [63-65]. Due to growing concerns about security breaches of electronic records worldwide, the system will need to ensure the protection of patient identity, maintenance and monitoring of firewalls, training of medical staff on confidentiality and ethics, and access of data to appropriate personnel only [66].

The doctors expressed difficulties in understanding the symptoms of patients who could not express themselves well over the phone and felt the necessity for a visual examination. The need for visual examination may be mitigated by images sent from smartphone apps by patients [71]; however, this may not be a practical solution in Bangladesh where smartphones are still not a cheap option for rural disadvantaged households [72], and maneuvering a smartphone requires soft skills beyond the limited ability to receive or make calls. We suggest that local resources, such as the village doctors, *Dai* s (traditional birth attendants), or community health workers, the first people to be contacted during emergencies, may help explain the symptoms to remote doctors and arrange emergency referrals in rural areas [14]. Capacity building of local resources is a prerequisite for this model to work and requires a well-structured and coordinated national capacity building program. mHealth-linked call centers established between village doctors and professional medical doctors were suggested as a model to build the capacity of semiqualified professionals in Chakaria, Bangladesh [58]. Semiqualified or unqualified village doctors account for approximately 80% of local providers in Bangladesh and are known for inappropriate medication, polypharmacy (prescribing five or more medicines in one prescription), and inappropriate injections at the "patient's demand" [73]. mHealth-linked professional training of village doctors could improve remote diagnosis and adherence to an appropriate prescription policy in Bangladesh [57]. We could benefit from mHealth doctors' advocacy in raising awareness among callers about appropriate pharmaceutical prescriptions and use, thus contributing toward a change of behavior in consuming self-medicated antibiotics purchased without prescription at local drugstores [74]. Additionally, a national strategy and initiatives to train traditional *Dai* s and community health workers should also be revived [3]. Evidence from another study that showed improved performance of community health workers in screening high-risk patients with the aid of mHealth algorithms and training modules may be adopted in programs targeted to build the capacity of community health workers [68,75].

Our study is limited by a number of factors. First, the sample size for interviews was small, which may impact the generalizability of our findings; however, data triangulation (analysis of call log data) was adopted to minimize the limitation [45]. Second, due to the systematic sampling of interviewees from the call log list, low-income households are underrepresented in the qualitative study [76]. Third, we had access to limited data in the call list; a dataset containing clients' age, education, occupation, economic status, and follow-up results could improve analysis on adoption of mobile phone consultation service among women. Fourth, our research does not address sustainability issues such as operational costs, technological challenges, and revenue generation by the call center. This is vital information for understanding the possibilities of service expansion and escalation. A future quantitative study including a cost-effectiveness analysis, willingness to pay among callers, and the effects of doctors' remote consultations on health outcomes and referrals is warranted.

Despite the limitations, our study is significant for a number of reasons. We presented experiences of two groups of families—one where women could access the service themselves and the other where women had to rely on their husbands—to understand how gender, autonomy, and mobile phone literacy play a role in women's utilization of telehealth care. These are important considerations when designing future mHealth interventions. We advocate that the inclusion of husbands in health interventions would accelerate women's access to health care for positive outcomes. Our study reports on the experiences of doctors who recounted the benefits and challenges of remote diagnosis in comparison to health facilities. The call log data inform us about the call time, frequencies, level of urgency, and access of subscribers. These insights offer the opportunity to improve our understanding of the benefits and challenges of mHealth as services develop and expand across Bangladesh and other low-income countries.

Mobile phone consultations or telehealth for maternal, neonatal, and infant health care are a feasible primary health care activity to increase access to professional medical advice, especially in remote areas of Bangladesh. A remote consultation service may influence some of the cultural and economic factors that delay maternal and child health care-seeking practices. A collaborative well-designed referral system connected with specialists and emergency MNCH services may improve health outcomes for urgent conditions. A central government-run database to collect heterogeneous patient data from public-private initiatives will ensure prompt management of emergency referrals generated by mHealth services. We recommend the systematic use of

quantitative surveys to understand health care-seeking behaviors and health outcomes of users, as well as trials on an integrated referral system between call centers and health facilities.

Acknowledgments

The authors acknowledge Dnet's (the implementing agency of Aponjon in Bangladesh) support to use their data on the Aponjon service and permission to conduct this research. Dnet received catalytic support from the United States Agency for International Development for launching the service. We also thank all participants in this research, and the call center management for organizing the place and time to conduct the interviews of the doctors. We acknowledge the critical reviews and intellectual inputs of the reviewers, which have improved the quality of the manuscript. This research was supported by Research School of Population Health, The Australian National University, and an Australian Government Research Training Program Scholarship. The views in this publication are those of the authors only and do not necessarily reflect the views of the institutions involved in this publication.

Authors' Contributions

MA had a key role in the design, conception, data collection, transcription, translation, coding, analysis, and interpretation of data, and was a major contributor in writing the manuscript. CB had a substantial contribution in interpreting the data. CB, AO, and KL were involved in designing the work and critically revising the manuscript for important intellectual content. All authors proofread, approved the manuscript, and are accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Bangladesh Bureau of Statistics. Dhaka: Bangladesh Bureau of Statistics with support from International Labour Organization; 2015 Oct. Report on ICT Use and Access by Individuals and Households Bangladesh 2013 URL: http://203.112.218.65:8008/WebTestApplication/userfiles/Image/LatestReports/ICTUseAccessSurvey2013.pdf [accessed 2019-03-17] [WebCite Cache ID 76xH9l6V6]
2. Mannan A. Access to public health facilities in Bangladesh: a study on facility utilisation and burden of treatment. Bangladesh Dev Stud 2013;36(4):25-80 [FREE Full text]
3. Ahmed SM, Alam BB, Anwar I, Begum T, Huque R, Khan JA, et al. Bangladesh Health System Review. Geneva: World Health Organization; 2015.
4. National Institute of Population Research and Training (NIPORT). Dhaka, Bangladesh: MEASURE Evaluation, and ICDDR,B; 2012. Bangladesh Maternal Mortality and Health Care Survey 2010 URL: https://www.measureevaluation.org/resources/publications/tr-12-87 [WebCite Cache ID 76xJ3vsTb]
5. Roy L, Biswas TK, Chowdhury ME. Emergency obstetric and newborn care signal functions in public and private facilities in Bangladesh. PLoS One 2017;12(11):e0187238. [doi: 10.1371/journal.pone.0187238] [Medline: 29091965]
6. National Institute of Population Research and Training, Ministry of Health and Family Welfare, Mitra and Associates, The DHS Program, ICF International. Bangladesh Demographic and Health Survey 2014. Dhaka, Bangladesh, and Rockville, MD: NIPORT, Mitra and Associates, ICF International; 2016. URL: https://dhsprogram.com/pubs/pdf/FR311/FR311.pdf [accessed 2019-03-17] [WebCite Cache ID 76xJUjd7n]
7. Killewo J, Anwar I, Bashir I, Yunus M, Chakraborty J. Perceived delay in healthcare-seeking for episodes of serious illness and its implications for safe motherhood interventions in rural Bangladesh. J Health Popul Nutr 2006 Dec;24(4):403-412 [FREE Full text] [Medline: 17591337]
8. Nonyane BA, Kazmi N, Koffi AK, Begum N, Ahmed S, Baqui AH, et al. Factors associated with delay in care-seeking for fatal neonatal illness in the Sylhet district of Bangladesh: results from a verbal and social autopsy study. J Glob Health 2016 Jun;6(1):010605 [FREE Full text] [doi: 10.7189/jogh.06.010605] [Medline: 27350876]
9. Sarker BK, Rahman M, Rahman T, Hossain J, Reichenbach L, Mitra DK. Reasons for preference of home delivery with traditional birth attendants (TBAs) in rural Bangladesh: a qualitative exploration. PLoS One 2016;11(1):e0146161 [FREE Full text] [doi: 10.1371/journal.pone.0146161] [Medline: 26731276]
10. Oyeyemi SO, Wynn R. The use of cell phones and radio communication systems to reduce delays in getting help for pregnant women in low- and middle-income countries: a scoping review. Global Health Action 2015 Sep 10;8(1):28887. [doi: 10.3402/gha.v8.28887]
11. Rhoads SJ, Eswaran H, Lynch CE, Ounpraseuth ST, Magann EF, Lowery CL. High-risk obstetrical call center: a model for regions with limited access to care. J Matern Fetal Neonatal Med 2018 Apr;31(7):857-865. [doi: 10.1080/14767058.2017.1300645] [Medline: 28316278]

12. Gallegos D, Cromack C, Thorpe KJ. Can a phone call make a difference? Breastfeeding self-efficacy and nurse responses to mother's calls for help. J Child Health Care 2018 Dec;22(3):433-446. [doi: 10.1177/1367493518757066] [Medline: 29514511]
13. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. Matern Child Health J 2012 Jul;16(5):1092-1101. [doi: 10.1007/s10995-011-0836-3] [Medline: 21688111]
14. Huq NL, Azmi AJ, Quaiyum MA, Hossain S. Toll free mobile communication: overcoming barriers in maternal and neonatal emergencies in rural Bangladesh. Reprod Health 2014;11:52 [FREE Full text] [doi: 10.1186/1742-4755-11-52] [Medline: 25015126]
15. Sondaal SF, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mHealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. PLoS One 2016;11(5):e0154664 [FREE Full text] [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
16. Abaza H, Marschollek M. mHealth application areas and technology combinations. A comparison of literature from high and low/middle income countries. Methods Inf Med 2017 Aug 08;56(7):e105-e122. [doi: 10.3414/ME17-05-0003] [Medline: 28925418]
17. LeFevre AE, Mohan D, Hutchful D, Jennings L, Mehl G, Labrique A, et al. Mobile technology for community health in Ghana: what happens when technical functionality threatens the effectiveness of digital health programs? BMC Med Inform Decis Mak 2017 Dec 14;17(1):27. [doi: 10.1186/s12911-017-0421-9] [Medline: 28292288]
18. Poorman E, Gazmararian J, Elon L, Parker R. Is health literacy related to health behaviors and cell phone usage patterns among the text4baby target population? Arch Public Health 2014;72(1):13 [FREE Full text] [doi: 10.1186/2049-3258-72-13] [Medline: 24872883]
19. Barron P, Peter J, LeFevre AE, Sebidi J, Bekker M, Allen R, et al. Mobile health messaging service and helpdesk for South African mothers (MomConnect): history, successes and challenges. BMJ Glob Health 2018;3(Suppl 2):e000559 [FREE Full text] [doi: 10.1136/bmjgh-2017-000559] [Medline: 29713503]
20. Crawford J, Larsen-Cooper E, Jezman Z, Cunningham SC, Bancroft E. SMS versus voice messaging to deliver MNCH communication in rural Malawi: assessment of delivery success and user experience. Glob Health Sci Pract 2014 Feb;2(1):35-46 [FREE Full text] [doi: 10.9745/GHSP-D-13-00155] [Medline: 25276561]
21. Watkins SC, Robinson A, Dalious M. Evaluation of the Information and Communications Technology for Maternal, Newborn and Child Health Project. Malawi: Invest in Knowledge Initiative; 2013. URL: http://www.villagereach.org/wp-content/uploads/2017/07/ICT_for_MNCH_Report_131211md_FINAL.pdf [accessed 2019-03-18] [WebCite Cache ID 76xREdJVZ]
22. Lund S, Nielsen BB, Hemed M, Boas IM, Said A, Said K, et al. Mobile phones improve antenatal care attendance in Zanzibar: a cluster randomized controlled trial. BMC Pregnancy Childbirth 2014 Jan 17;14:29 [FREE Full text] [doi: 10.1186/1471-2393-14-29] [Medline: 24438517]
23. Khatun F, Hanifi SM, Iqbal M, Rasheed S, Rahman MS, Ahmed T, et al. Prospects of mHealth services in Bangladesh: recent evidence from Chakaria. PLoS One 2014;9(11):e111413 [FREE Full text] [doi: 10.1371/journal.pone.0111413] [Medline: 25375255]
24. Tran MC, Labrique AB, Mehra S, Ali H, Shaikh S, Mitra M, et al. Analyzing the mobile "Digital Divide": Changing Determinants of Household Phone Ownership Over Time in Rural Bangladesh. JMIR Mhealth Uhealth 2015;3(1):e24 [FREE Full text] [doi: 10.2196/mhealth.3663] [Medline: 25720457]
25. Directorate General of Health Services (DGHS), Ministry of Health & Family Welfare (MOHFW). Health Information system and eHealth. Dhaka: Government of the People's Republic of Bangladesh; 2011. URL: http://www.dghs.gov.bd/images/docs/OP/HIS%20and%20e-health%20OP%20approved%2015-12-2011.pdf [accessed 2019-03-18] [WebCite Cache ID 76yYjD3Wh]
26. Ahmed T, Lucas H, Khan AS, Islam R, Bhuiya A, Iqbal M. eHealth and mHealth initiatives in Bangladesh: a scoping study. BMC Health Serv Res 2014;14:260 [FREE Full text] [doi: 10.1186/1472-6963-14-260] [Medline: 24934164]
27. Grameenphone. 2007. Grameenphone's healthline service re-launched. URL: https://cdn01.grameenphone.com/about/media-center/press-release/grameenphones-healthline-service-re-launched [accessed 2019-03-18] [WebCite Cache ID 76yZQ8N1i]
28. Grameenphone. 2016. Free health service TONIC launched for Grameenphone customers. URL: https://www.grameenphone.com/about/media-center/press-release/free-health-service-tonic-launched-grameenphone-customers [accessed 2019-03-18] [WebCite Cache ID 76yaat11E]
29. Khatun F, Heywood AE, Ray PK, Bhuiya A, Liaw ST. Community readiness for adopting mHealth in rural Bangladesh: a qualitative exploration. Int J Med Inform 2016 Sep;93:49-56. [doi: 10.1016/j.ijmedinf.2016.05.010] [Medline: 27435947]
30. Ministry of Health and Family Welfare Bangladesh. Mobile phone consultation service. URL: http://16263.dghs.gov.bd/ [accessed 2019-03-18] [WebCite Cache ID 76ybTMNz4]
31. Rajan R, Raihan A, Alam M, Agarwal S, Ahsan A, Bashir R, et al. MAMA Aponjon Formative Research Report. Baltimore, MD: Johns Hopkins University Global Health Initiative; 2013 Dec. URL: http://www.aponjon.com.bd/pdf_view_report/aponjon_report_56a092170a5b9.pdf [accessed 2019-03-18] [WebCite Cache ID 76ybuL6F1]

32. Alam M, D'Este C, Banwell C, Lokuge K. The impact of mobile phone based messages on maternal and child healthcare behaviour: a retrospective cross-sectional survey in Bangladesh. BMC Health Serv Res 2017;17(1):434. [doi: 10.1186/s12913-017-2361-6] [Medline: 28645278]
33. Aponjon. Dhaka: Dnet URL: http://www.aponjon.com.bd/ [accessed 2019-03-18] [WebCite Cache ID 76ydL0IM3]
34. Aponjon News. Dhaka: Dnet; 2013 Apr. URL: http://www.aponjon.com.bd/pdf_view/aponjon_newsletter_55ed2ec62cd3c.pdf [accessed 2019-03-18] [WebCite Cache ID 76ycALbTQ]
35. Chowdhury ME, Roy L, Wahed T, Mangxaba J. What is the minimum exposure to mobile messaging needed to improve maternal and newborn health knowledge, practices and demand for services in Bangladesh? 2015 Presented at: Global Maternal Newborn Health Conference; Oct 18-21, 2015; Mexico City URL: https://cdn2.sph.harvard.edu/wp-content/uploads/sites/32/2015/12/Mahbub-Elahi-Chowdhury.pdf
36. Huda TM, Alam A, Tahsina T, Hasan MM, Khan J, Rahman MM, et al. Mobile-based nutrition counseling and unconditional cash transfers for improving maternal and child Nutrition in Bangladesh: pilot study. JMIR Mhealth Uhealth 2018;6(7):e156. [doi: 10.2196/mhealth.8832] [Medline: 30021707]
37. Lee SH, Nurmatov UB, Nwaru BI, Mukherjee M, Grant L, Pagliari C. Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: Systematic review and meta-analysis. J Glob Health 2016 Jun;6(1):010401 [FREE Full text] [doi: 10.7189/jogh.06.010401] [Medline: 26649177]
38. World Health Organization, Ministry of Health and Family Welfare Bangladesh. Success Factors for Women's and Children's Health. Geneva: World Health Organization; 2015. URL: https://apps.who.int/iris/handle/10665/178623 [accessed 2019-03-18] [WebCite Cache ID 76yd6kww6]
39. Directorate General of Health Services. Health Bulletin 2012. Dhaka: Ministry of Health and Family Welfare Bangladesh; 2012. URL: http://dghs.gov.bd/bn/licts_file/images/Health_Bulletin/HealthBulletin2012_full.pdf [accessed 2019-03-18] [WebCite Cache ID 76ydOjnTl]
40. Bangladesh Bureau of Statistics. 2012. Population and housing census 2011: socio-economic and demographic report. URL: http://203.112.218.65:8008/WebTestApplication/userfiles/Image/National%20Reports/SED_REPORT_Vol-4.pdf [accessed 2019-03-19] [WebCite Cache ID 76ydVxk7O]
41. The World Bank. Bangladesh Overview 2018. URL: https://www.worldbank.org/en/country/bangladesh/overview [accessed 2019-03-18] [WebCite Cache ID 76ydtNs0h]
42. Bangladesh Bureau of Statistics. District statistics 2019. URL: http://www.bbs.gov.bd/site/page/2888a55d-d686-4736-bad0-54b70462afda/- [accessed 2019-03-18] [WebCite Cache ID 76ydwGVM2]
43. El Arifeen S, Christou A, Reichenbach L, Osman FA, Azad K, Islam KS, et al. Community-based approaches and partnerships: innovations in health-service delivery in Bangladesh. Lancet 2013 Dec 14;382(9909):2012-2026. [doi: 10.1016/S0140-6736(13)62149-2] [Medline: 24268607]
44. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
45. Fusch PI, Ness LR. Are we there yet? Data saturation in qualitative research. Qual Report 2015;20(9):1404-1416 [FREE Full text]
46. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
47. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
48. Ashraf A, Chowdhury S, Streefland P. Health, disease and health-care in rural Bangladesh. Soc Sci Med 1982;16(23):2041-2054 [FREE Full text] [Medline: 7157035]
49. Shanmugavelan M, Alam M, Raihan A, Shoemacker E. Mobile Phones for Millennium Development Goal: Qualitative Need Analysis to Overcome Communication Barriers to Improve Maternal Health. Dhaka, Bangladesh: Dnet and Panos London; 2011.
50. Shahabuddin A, Nostlinger C, Delvaux T, Sarker M, Delamou A, Bardaji A, et al. Exploring maternal health care-seeking behavior of married adolescent girls in Bangladesh: a social-ecological approach. PLoS One 2017;12(1):A. [doi: 10.1371/journal.pone.0169109] [Medline: 28095432]
51. Khatun F, Heywood AE, Hanifi SM, Rahman MS, Ray PK, Liaw ST, et al. Gender differentials in readiness and use of mHealth services in a rural area of Bangladesh. BMC Health Serv Res 2017 Aug 18;17(1):573 [FREE Full text] [doi: 10.1186/s12913-017-2523-6] [Medline: 28821243]
52. Osamor PE, Grady C. Women's autonomy in health care decision-making in developing countries: a synthesis of the literature. Int J Womens Health 2016;8:191-202. [doi: 10.2147/ijwh.S105483] [Medline: 27354830]
53. Qureshi RN, Sheikh S, Khowaja AR, Hoodbhoy Z, Zaidi S, Sawchuck D, et al. Health care seeking behaviours in pregnancy in rural Sindh, Pakistan: a qualitative study. Reprod Health 2016;13:34. [doi: 10.1186/s12978-016-0140-1] [Medline: 27356863]

54. Bouri N, Ravi S. Going mobile: how mobile personal health records can improve health care during emergencies. JMIR Mhealth Uhealth 2014;2(1):e8 [FREE Full text] [doi: 10.2196/mhealth.3017] [Medline: 25098942]
55. Williams A, Sarker M, Ferdous ST. Cultural attitudes toward postpartum depression in Dhaka, Bangladesh. Med Anthropol 2018 Dec;37(3):194-205. [doi: 10.1080/01459740.2017.1318875] [Medline: 28409647]
56. Chowdhury AI, Haider R, Abdullah AY, Christou A, Ali NA, Rahman AE, et al. Using geospatial techniques to develop an emergency referral transport system for suspected sepsis patients in Bangladesh. PLoS One 2018;13(1):e0191054. [doi: 10.1371/journal.pone.0191054] [Medline: 29338012]
57. Directorate General of Drug Administration. Dhaka: Ministry of Health and Family Welfare Bangladesh; 2017. National Drug Policy 2016. URL: http://www.dgda.gov.bd/index.php/laws-and-policies/261-national-drug-policy-2016-english-version [accessed 2019-03-18] [WebCite Cache ID 76yhVyKNw]
58. Khan NU, Rasheed S, Sharmin T, Ahmed T, Mahmood SS, Khatun F, et al. Experience of using mHealth to link village doctors with physicians: lessons from Chakaria, Bangladesh. BMC Med Inform Decis Mak 2015 Aug 05;15:62 [FREE Full text] [doi: 10.1186/s12911-015-0188-9] [Medline: 26242574]
59. Thaddeus S, Maine D. Too far to walk: maternal mortality in context. Soc Sci Med 1994 Apr;38(8):1091-1110. [Medline: 8042057]
60. Hunter BM, Harrison S, Portela A, Bick D. The effects of cash transfers and vouchers on the use and quality of maternity care services: A systematic review. PLoS One 2017;12(3):e0173068 [FREE Full text] [doi: 10.1371/journal.pone.0173068] [Medline: 28328940]
61. Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. Int J Med Inform 2008 May;77(5):291-304. [doi: 10.1016/j.ijmedinf.2007.09.001] [Medline: 17951106]
62. Gurupur VP, Wan TT. Challenges in implementing mHealth interventions: a technical perspective. Mhealth 2017;3:32 [FREE Full text] [doi: 10.21037/mhealth.2017.07.05] [Medline: 28894742]
63. Directorate General of Health Services. Health Informatics Standards and Data Structure for Bangladesh. Dhaka: Ministry of Health and Family Welfare Bangladesh; 2012 Oct. URL: http://www.dghs.gov.bd/images/docs/eHealth/Standards_and_interoperability_document_final_5.01.14.pdf [accessed 2019-03-18] [WebCite Cache ID 76ygSuexY]
64. Shen F, Liu S, Wang Y, Wen A, Wang L, Liu H. Utilization of electronic medical records and biomedical literature to support the diagnosis of rare diseases using data fusion and collaborative filtering approaches. JMIR Med Inform 2018 Oct 10;6(4):e11301. [doi: 10.2196/11301] [Medline: 30305261]
65. Kocbek S, Kocbek P, Stozer A, Zupanic T, Groza T, Stiglic G. Building interpretable models for polypharmacy prediction in older chronic patients based on drug prescription records. PeerJ 2018 Oct 12;6:e5765. [doi: 10.7717/peerj.5765] [Medline: 30345175]
66. Ozair FF, Jamshed N, Sharma A, Aggarwal P. Ethical issues in electronic health records: a general overview. Perspect Clin Res 2015;6(2):73-76 [FREE Full text] [doi: 10.4103/2229-3485.153997] [Medline: 25878950]
67. Choun K, Achanta S, Naik B, Tripathy JP, Thai S, Lorent N, et al. Using mobile phones to ensure that referred tuberculosis patients reach their treatment facilities: a call that makes a difference. BMC Health Serv Res 2017 Aug 22;17(1):575 [FREE Full text] [doi: 10.1186/s12913-017-2511-x] [Medline: 28830542]
68. Prinja S, Nimesh R, Gupta A, Bahuguna P, Gupta M, Thakur JS. Impact of m-health application used by community health volunteers on improving utilisation of maternal, new-born and child health care services in a rural area of Uttar Pradesh, India. Trop Med Int Health 2017 Jun 13;22(7):895-907. [doi: 10.1111/tmi.12895]
69. Jo Y, Labrique AB, Lefevre AE, Mehl G, Pfaff T, Walker N, et al. Using the lives saved tool (LiST) to model mHealth impact on neonatal survival in resource-limited settings. PLoS One 2014;9(7):e102224 [FREE Full text] [doi: 10.1371/journal.pone.0102224] [Medline: 25014008]
70. Killingsworth JR, Hossain N, Hedrick-Wong Y, Thomas SD, Rahman A, Begum T. Unofficial fees in Bangladesh: price, equity and institutional issues. Health Policy Plan 1999;14(2):152-163. [Medline: 10538718]
71. Gunter R, Fernandes-Taylor S, Mahnke A, Awoyinka L, Schroeder C, Wiseman J, et al. Evaluating patient usability of an image-based mobile health platform for postoperative wound monitoring. JMIR Mhealth Uhealth 2016 Sep 28;4(3):e113 [FREE Full text] [doi: 10.2196/mhealth.6023] [Medline: 27683059]
72. Rogers M. Country Overview: Bangladesh Mobile Industry Driving Growth and Enabling Digital Inclusion. London: GSM Association; 2018. URL: https://www.gsmaintelligence.com/research/?file=a163eddca009553979bcdfb8fd5f2ef0&download
73. Rasu RS, Iqbal M, Hanifi S, Moula A, Hoque S, Rasheed S, et al. Level, pattern, and determinants of polypharmacy and inappropriate use of medications by village doctors in a rural area of Bangladesh. Clinicoecon Outcomes Res 2014;6:515-521. [doi: 10.2147/ceor.S67424] [Medline: 25506232]
74. Biswas M, Roy MN, Manik MI, Hossain MS, Tapu SM, Moniruzzaman M, et al. Self medicated antibiotics in Bangladesh: a cross-sectional health survey conducted in the Rajshahi City. BMC Public Health 2014 Aug 14;14:847 [FREE Full text] [doi: 10.1186/1471-2458-14-847] [Medline: 25124712]
75. Alam M, Khanam T, Khan R. Assessing the scope for use of mobile based solution to improve maternal and child health in Bangladesh: A case study. 2010. Presented at: ICTD 2010: 4th ACM/IEEE International Conference on Information and Communication Technologies and Development; Dec 13-16, 2010; London. [doi: 10.1145/2369220.2370755]

76. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544 [FREE Full text] [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]

Abbreviations

BDT: Bangladeshi taka (currency)
CRM: customer relationship management
GP: general physician
IVR: interactive voice responses
MNCH: maternal, neonatal, child health care

Edited by G Eysenbach; submitted 05.08.18; peer-reviewed by YY Wang, J Apolzan, C Perrin, T Woods; comments to author 12.10.18; revised version received 07.12.18; accepted 04.03.19; published 22.04.19.

Please cite as:
Alam M, Banwell C, Olsen A, Lokuge K
Patients' and Doctors' Perceptions of a Mobile Phone–Based Consultation Service for Maternal, Neonatal, and Infant Health Care in Bangladesh: A Mixed-Methods Study
JMIR Mhealth Uhealth 2019;7(4):e11842
URL: http://mhealth.jmir.org/2019/4/e11842/
doi:10.2196/11842
PMID:31008716

Original Paper

Needs Analysis for a Parenting App to Prevent Unintentional Injury in Newborn Babies and Toddlers: Focus Group and Survey Study Among Chinese Caregivers

Peishan Ning[1*], MD; Deyue Gao[1*], BA; Peixia Cheng[1], BA; David C Schwebel[2], PhD; Xiang Wei[1], BA; Liheng Tan[1], BA; Wangxin Xiao[1], BA; Jieyi He[1], BA; Yanhong Fu[1], BA; Bo Chen[1], BA; Yang Yang[3], PhD; Jing Deng[1], PhD; Yue Wu[4], PhD; Renhe Yu[1], BA; Shukun Li[5], MA; Guoqing Hu[1], PhD

[1]Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Changsha, China

[2]Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, United States

[3]Department of Biostatistics, College of Public Health and Health Professions, University of Florida, Gainesville, FL, United States

[4]Department of Occupational and Environmental Health, Xiangya School of Public Health, Central South University, Changsha, China

[5]Information and Network Center, Central South University, Changsha, China

[*]these authors contributed equally

Corresponding Author:
Guoqing Hu, PhD
Department of Epidemiology and Health Statistics
Xiangya School of Public Health
Central South University
110 Xiangya Road
Changsha, 410078
China
Phone: 86 073184805414
Fax: 86 073184805414
Email: huguoqing009@gmail.com

Abstract

Background: With the growing popularity of mobile health technology, app-based interventions delivered by smartphone have become an increasingly important strategy toward injury prevention.

Objective: This study aimed to develop a framework supporting the design of an app-based intervention to prevent unintentional injury, targeted for caregivers of Chinese children aged 0 to 6 years.

Methods: A theory-based mixed-method study, including focus groups and Web-based quantitative survey, was performed. Adult caregivers who care for children aged 0 to 6 years and own a smartphone were recruited into 2 sequential stages of research. First, focus groups were conducted among the caregivers at community health care centers and preschools from December 2015 to March 2016. Focus groups (8-10 participants per group) explored awareness, experiences, and opinions of caregivers toward using an app to prevent unintentional injury among children. Second, based on the focus groups findings, a Web-based quantitative survey was designed and distributed to caregivers in November 2016; it collected information on specific needs for the app-based intervention. Thematic analysis and quantitative descriptive analyses were performed.

Results: In total, 12 focus groups were completed, involving 108 caregivers. Most participants expressed a strong desire to learn knowledge and skills about unintentional child injury prevention and held positive attitudes toward app-based interventions. Participants expressed multiple preferences concerning the app-based intervention, including their contents, functions, interactive styles, installation and registration logistics, and privacy protection and information security. Following the focus groups, 1505 caregivers completed a WeChat-based quantitative survey, which generated roughly similar results to those of focus groups and added numerical metrics concerning participants' preferences on what to learn, when to learn it, and how to learn it. A detailed framework was established involving 5 components: (1) content design, (2) functional design, (3) interactive style, (4) installation and registration logistics, and (5) privacy protection and information security, and 15 specific requirements.

Conclusions: We developed a framework that can be used as a guide to design app-based interventions for parents and caregivers, specifically for unintentional injury prevention of children aged 0 to 6 years.

(JMIR Mhealth Uhealth 2019;7(4):e11957) doi:10.2196/11957

KEYWORDS

injury; child; mobile health; education; intervention; parenting; mhealth

Introduction

Background

Unintentional injury is a serious public health problem among children aged 0 to 6 years in China. In 2017, over 20,000 Chinese children younger than 5 years died of unintentional injury, and nearly 2.4 million children required emergency or outpatient care because of unintentional injury [1].

Parenting interventions can substantially reduce child unintentional injury risk by improving caregivers' safety knowledge and perceptions about the risk for injury to their children and through adoption of safety equipment and practices [2-5]. For a variety of reasons, implementation of such interventions has been inadequate in low- and middle-income countries such as China [6,7]. In fact, 2 domestic studies from China indicate that less than 20% of caregivers had ever received education concerning unintentional injury prevention for children aged 0 to 6 years or attended a professional prevention course on the topic [8,9].

Encouragingly, the mobile health (mHealth) movement to deliver health interventions through technology has potential to overcome this barrier to parent education for child injury prevention. The number of smartphone users in China has grown quickly over the past decade, with recent estimates suggesting that about 403 million adults aged 20 to 39 years (95% of the Chinese population in that age group) accessed the internet through mobile phones in December 2017 [10,11]. Compared with traditional health education methods delivered in person by professionals, mHealth interventions are low-cost and easy-to-implement, allow interactions between users and providers, and can be accessed anonymously with flexibility at any time and any place [12-14].

Recently, app-based interventions have been developed to prevent or reduce unintentional injury risk in specific injury domains, including sports injuries [15], fire injuries [16], road traffic injuries [16,17], falls [18], and burns [19]. Some app-based interventions have been critiqued, however, for failing to meet basic principles for effective prevention programs such as being based in theory or being tailored to risks in the target population [20].

Multiple health behavior change theories can be used in the development of an app to reduce injuries. The Theory of Planned Behavior (TPB) offers a social psychological theory that fits nicely, as it is used to interpret and predict why individuals perform specific behaviors [21]. A recent systematic review concluded that TPB is an effective framework to identify and understand child and adolescent nutrition–related behaviors. It found, for example, that attitude was strongly related to dietary behavioral intention (mean r=0.52) and intention was the most common predictor of behavior performance (mean r=0.38) [22]. Interventions grounded in TPB must attend to topics such as the individual's attitudes toward using an app-based intervention and changing their behavior as a result of that app, the individual's subjective norms about behavior that should be conducted, and the individual's perception of how much they can control their own behavior surrounding the health area of interest.

Computer science theories are also critical to the development of an app-based intervention. The Framework for the Rational Analysis of Mobile Education (FRAME) model offers a theoretical grounding in computer science to guide the development of mHealth learning programs. The FRAME model considers 3 aspects of mobile learning: usability of the device and app, capacities of the learners, and social interaction between users. Regarded as a comprehensive model to develop and implement mobile learning, attention to the components of the FRAME model allows app development to proceed with attention to all relevant components of user learning [23].

Objectives

We, therefore, conducted a mixed-method study to establish a theory- and need-based framework that would support and lead to the design of an app-based intervention for child injury prevention to be used by Chinese caregivers of children aged 0 to 6 years. All assessment protocols were grounded in TPB and the FRAME model and were designed to gather information that would be valuable for the design of app-based health interventions focused on improving parenting among Chinese caregivers to reduce child injury risk.

Methods

Study Design

Grounded in TPB and the FRAME model [21,23], we designed a 2-step sequential mixed-method study. The TPB suggests that an effective app parenting intervention must offer strategies to alter the attitudes, subjective norms, and perceived control to engage in actions that will improve the safety of children. We targeted several aspects of these strategies in our inquiries to caregivers concerning their preferences in the design of an app-based intervention. The FRAME model addresses users' preferences for app usability, the capacity of the app users, and the social interactions between app users; we addressed these aspects of app functioning also in the inquiries to caregivers.

To gather qualitative data first, focus groups were conducted to explore the experiences and preferences of adult caregivers on using an app-based intervention to prevent unintentional injury among children younger than 7 years. Next, focus group findings were used to guide development and implementation of a Web-based survey to quantify key needs for the app-based interventions (eg, frequency and length of content, variety and types of forms of learning, and duration of app-based learning).

Focus Group

Participants

Participants were recruited for focus groups using purposive sampling. Eligible participants included primary caregivers of 1 or more children younger than 7 years and who used a smartphone regularly. No exclusion criteria applied beyond the age of children that caregivers looked after.

Caregivers of children aged 0 to 3 years were recruited primarily from community health centers. Caregivers of children aged 4 to 6 years were recruited from preschools. To maximize sampling variation, we sampled from a range of preschool types (including both public and private) and geographic locations (including varied socioeconomic status and living areas). Focus group members represented both sexes and a range of household incomes, ages, and levels of education.

Setting

Focus groups were completed between December 2015 and March 2016 in Changsha, China. A semistructured discussion guide was developed and refined through pilot testing first among the research group and then with an independent focus group of caregivers of children aged 0 to 3 years (Multimedia Appendix 1). Informed consent was obtained from all participants before each focus group began.

Broadly, discussions were organized to concentrate on perceptions and opinions about unintentional child injury prevention and app-based injury interventions as well as on issues of designing app-based intervention (eg, content, function, interfaces, and data security). To support discussion about intervention interfaces, the facilitators prepared a series of slides that illustrated different styles of interface designs.

When each focus group discussion concluded, participants completed short paper-based questionnaires that collected information on sociodemographic characteristics for both caregivers and their children and unintentional injury history for their children.

Following scholarly recommendations to organize focus groups [24,25], each focus group consisted of 8 to 10 participants and lasted for 60 to 90 min. A total of 3 trained facilitators led the discussion of all groups; 1 served as moderator and the other 2 took extensive discussion notes and supported the moderator as needed to organize the discussion. All focus group discussions were audio-taped, transcribed, and then reviewed before the next focus group. This allowed facilitators to include new concepts or opinions and exclude old but less relevant subtopics iteratively. Focus groups were concluded when no new concepts and opinions emerged.

Data Analysis

Focus group audiotapes were transcribed verbatim and checked for accuracy. All participant names were removed to permit anonymized analysis. The transcripts were then analyzed using thematic analysis strategies [26]. First, 2 researchers spent several days familiarizing themselves with all transcripts and generating initial codes independently. Second, the 2 coders combined initial codes and organized them into themes based on the discussion guideline, concordance between codes, and the underlying theory. Third, the 2 researchers' proposed themes were combined and re-evaluated to ensure no major themes from the discussions were omitted. Finally, group discussions of the research group (including 2 coders) were held to achieve consensus on any themes where the 2 researchers held differing views. Transcript analysis and coding were completed using MAXQDA 12.0 (qualitative data analysis software by VERBI GmbH).

Web-Based Survey

Participants

Study participants for the Web-based survey component of the study were recruited through snowball sampling. Inclusion criterion matched those of the focus groups: caregivers of children younger than 7 years who owned a smartphone. No exclusion criteria applied beyond the age of children that caregivers looked after.

To maximize reach to the target population, we used WeChat, the most popular social media communication platform in China, with an average of 889 million active users each month [27], to conduct the survey. Specifically, the research team sent a study recruitment message through an official WeChat account known as La Ma Xue Yuan (School for Young Mothers), which includes 9290 users throughout China (mostly caregivers of young children). This message included an invitation letter and the questionnaire. Participants consented to participate online and then completed and submitted the questionnaire. Participants were also encouraged to share the message with their social network, many of whom met the inclusion criterion, allowing us to "snowball" to a larger sample size. The Web-based survey remained open until the number of completed questionnaires plateaued.

Setting

The Web-based survey was completed in November 2016. It was developed based on the results of the focus groups, grounded in TPB and especially the FRAME model, and finalized through pilot testing with 20 caregivers. The survey included items assessing demographic characteristics, history of child unintentional injury, prior learning experiences about child unintentional injury prevention, and preferred learning contents (Multimedia Appendix 2). To avoid repeated questionnaire completion, we restricted the survey to a single response from each smartphone. To encourage participation in the Web-based survey portion of the study, raffle prizes with cash incentives were distributed at a probability of 1/3 after participants completed the survey.

Data Analysis

The collected questionnaires were screened to exclude those completed by caregivers who did not meet the inclusion criteria. Proportions were then calculated to describe preferred contents of the app-based intervention, forms of app-based learning, preferred learning time, and frequency and duration for using app-based learning. Chi-square test examined group differences across groups. Data analysis was performed using the Statistical

Product and Service Solutions (SPSS 18.0; IBM Corporation). $P<.05$ was considered to be statistically significant.

Ethical Considerations

The protocol was approved by the ethics committee of the Institute of Clinical Pharmacology of Central South University. This study was conducted, analyzed, and reported according to the consolidated criteria for reporting qualitative research (COREQ) [28] and the checklist for reporting results of internet E-survey (CHERRIES [29]; Multimedia Appendix 3). All participants were informed about the study and provided informed consent before participating in the research. All data were analyzed anonymously.

Results

Focus Group Results

Participants

In total, 12 focus groups were organized; together they included 108 caregivers (90 parents, 6 grandparents, and 12 preschool teachers; Table 1). The largest portions of participants were female (96/108, 88.9%) and aged 20 to 39 years (94/108, 87.0%). Among the parent and grandparent participants, the caregivers supervised 96 children who ranged in age from 0 to 6 years, with the largest portions of children aged 4 to 6 years (35/96, 37%) and 0 to 1 year (34/96, 35%). The ratio of boys to girls supervised was 1.04 (Table 1).

Attitude and Behavioral Intention Toward Child Unintentional Injury Prevention

Most caregivers reported a strong desire to learn unintentional injury prevention strategies for children aged 0 to 6 years. Many participants mentioned that they felt unintentional injuries were largely preventable and that they had a strong intention to learn knowledge and skills that would help them prevent unintentional injuries to children under their care. A typical opinion was as follows:

> *We definitely want to learn it[knowledge about unintentional injury prevention for children ages 0-6 years]. From my point of view, it would be very useful and valuable for us [to learn that information].* [Participant #DWT-A-01]

Table 1. Demographic characteristics of caregiver participants.

Variable	Focus group participants, n (%)	Web-based survey participants, n (%)
Caregivers		
Total	108 (100.0)	1505 (100.00)
Sex		
Male	12 (11.1)	687 (45.65)
Female	96 (88.9)	818 (54.35)
Age group (years)		
≤19	2 (1.9)	33 (2.19)
20-29	48 (44.4)	805 (53.49)
30-39	46 (42.6)	604 (40.13)
40-49	8 (7.4)	62 (4.12)
≥50	4 (3.7)	1 (0.07)
Relationship with children		
Parent	90 (83.3)	—[a]
Grandparent	6 (5.6)	—[a]
Preschool teacher	12 (11.1)	—[a]
Children[b]		
Sex		
Male	49 (51)	859 (57.08)
Female	47 (49)	646 (42.92)
Age group (years)		
≤1	34 (35)	205 (13.62)
2-3	27 (28)	602 (40.00)
4-6	35 (37)	698 (46.38)

[a]Information was not collected.

[b]Children of the participant supervisors (90 parents and 6 grandparents) in the focus groups, excluding participants who were preschool teachers.

Beyond wanting to learn knowledge about child injury prevention, the majority of focus group participants expressed the opinion that they could use a smartphone app to learn that knowledge. The participants believed that an app developed by a professional team to address child injury prevention would be credible and could be created to include sufficient content for their learning. They also expressed the opinion that an app may be a more convenient educational method than alternatives, such as written brochures or social networks like WeChat, to search for injury prevention content when they wanted it:

> *Compared to other options, an app-based intervention developed by a professional team would provide rich content. Thus, it will be convenient to search for the knowledge we need. For example, knowledge we can obtain in WeChat is hard to retrieve when we need it again.* [Participant #YL-B-06]

A few participants expressed some concerns about the utility of a child injury prevention app. They worried about the possible difficulty in the installation and registration process and the possibility that it would use too much memory space on their smartphones, but they said they would accept the app if these problems were addressed.

Content of App-Based Intervention

Participants mentioned with some frequency 10 major causes of child unintentional injury that would be worth including in the app: exposure to animate mechanical forces (including animal bites and being trampled or bumped by other people); exposure to inanimate mechanical forces (including being pinched between 2 surfaces, such as in the doors of elevators, or cut or punctured by sharp objects); falls; contact with heat and hot substances; exposure to smoke, fire, and flames; transport crashes; unintentional threats to breathing that create suffocation risk; unintentional poisoning by and exposure to noxious substances; unintentional drowning and submersion; and exposure to electrical currents. Example statements appear below:

> *I really worry about injury from falls because children always like jumping from high places to lower places, which highly increases the risk of getting injured. In addition, road crashes are another important injury that I want to learn about, because many children run across or play in the street even when they see a car nearby.* [Participant #DWT-B-03]

> *Falls and burns are common types of unintentional injury for children ages 0-6 years at home. Relatively speaking, their harm to children is not as great as other types, such as electricalcurrents, which is one of the most dangerous injury causes for children.* [Participant #HHY-B-04]

Participants also recommended 3 facets of content design that could potentially increase the use of app-based intervention by caregivers: (1) providing professional and believable content to gain the trust of app users, (2) using plain language to make the contents easy to learn, and (3) providing easy-to-implement interventions to improve the practicability of applying lessons from the app-based intervention.

Forms of Learning Delivered by an App-Based Intervention

Several forms of learning from an app were proposed by the participants. Most frequently mentioned were short written alarms or warnings with pictures, cartoon vignettes, video testimonials, and interactive games. The participants believed these 4 forms would make learning easier and more engaging. Participants also mentioned the desire to learn the knowledge interactively with their children, a strategy they felt might maximize the effectiveness of the injury prevention program:

> *Pictures and video are easy to understand for both adults and children and might be the best way to disseminate the prevention knowledge. I also think both caregivers and children need to learn unintentional injury prevention. The easier the learning form, the better the learning outcome. In addition, it is better to include short warning words in the pictures to more effectively stop the children from adopting dangerous behaviors.* [Participant #DWT-B-01]

Timing for App-Based Learning

There was some variation across caregivers, but the general consensus was that they preferred the opportunity to use the app twice a week, for between 2 to 5 min at each session. As an example, a participant made the following statement:

> *The duration of the app-based learning should not be too long. I think "no more than 5 minutes" is fine for me because over five minutes learning would make me dizzy.* [Participant #MWD-B-10]

Participants also recommended evening as the best time of the day to interact with the app:

> *For me, it is particularly suitable to learn injury prevention knowledge at 8 or 9 o'clock in the evening. At that time, I have finished housework and my child hasfallen asleep.* [Participant #MWD-B-09]

Some participants felt the length of time they used the app would depend greatly on how much they enjoyed using it and that tangible cases and visual attraction would improve the authority and authenticity of the training, increasing their desire to use the app.

Functional Design of the App-Based Intervention

Almost all focus group participants mentioned interactive features as an indispensable function of the app-based intervention. They explained that interactive portions would maintain the attention of app users and increase the effectiveness of the intervention. Web-based chats, forums, and message boards were suggested as ways to implement interactive processes:

> *Web-based chatting is the best way to solve the problems that we encounter in our lives.* [Participant #HHY-A-01]

> *As we know, our kids may encounter various unintentional injuries in daily life. If the app intervention can set up a module listing possible ways*

> *to prevent common injury causes, it would be great. In addition, it would be attractive to me if the app had a forum in which I could discuss these topics with other caregivers who confront the same questions.* [Participant #YL-C-06]

In addition, a customer service agent who was accessible through Web-based chatting and a frequently asked questions module was suggested by the caregivers participating in the focus groups.

Beyond interactive portions of the app, most caregivers in the focus groups felt that a survey with feedback would be an important function of the app-based intervention. Such a component would allow app users to be aware of children's unintentional injury risk in their homes and to obtain tailored professional recommendations to prevent injuries to their children. Both Web-based and printed questionnaire surveys were recommended as appropriate approaches to obtain feedback from professionals.

In addition, some participants suggested surveys and feedback should be scheduled to be brief (eg, 5 min to complete each survey) and to be repeated regularly (eg, 3-4 times per year). Parents felt their responses might change as children develop new skills with older age, and therefore, tailored information that coincides with their children's development would be valuable. Participants recommended that questionnaires be based on items with categorical response options (eg, true or false) that could be responded to quickly. A few caregivers suggested providing survey participants with small gifts or bonuses to maintain their adherence to the intervention. Furthermore, to increase compliance in completing Web-based surveys, participants suggested the use of short message service text message reminders:

> *I'd like to use Web-based questionnaires to send feedback. I think survey questions with single options are much better than open-ended survey questions because I really do not know what should be filled in in many cases.* [Participant #HHY-B-03]

> *Both Web-based and printed questionnaire surveys are acceptable for me, but a Web-based questionnaire survey is preferred since it is more convenient and takes less time than a printed questionnaire survey. I probably would agree to complete a paper questionnaire survey if I received small gifts after I completed it.* [Participant #HHY-B-04]

Some caregivers extended the survey idea further, suggesting that it would be valuable to have the app customized to their preferences and priorities in terms of content, form, frequency, and interface of the app. As an example, preventive content concerning riding a bicycle may not be of interest or relevant to caregivers whose children are younger than 3 years, as most Chinese children do not learn to ride bicycles until they are at least 4 years old. Thus, parents suggested the app be tailored so that such segments would be omitted in their version of the program.

Design of the Interface

Participants stated that they would like to choose the app interface (eg, color and style) based on their preferences. During the group discussion, the moderator demonstrated various interfaces, including the layout of 3 existing apps, to obtain participants' opinions on the app design. Of the 3 options offered by the research team, most caregivers chose a cartoon style. A few caregivers preferred a simple interface with 3 replaceable pages, and fewer still selected a simple interface with only 1 page.

Installation and Registration

Caregivers strongly suggested the app should take up only a small amount of memory space on their smartphone and that it should require a simple registration procedure. They believed a user manual or help module might assist them with app use, and they requested an informational module that described the purpose, details, and benefits of the project before starting to use the app-based intervention.

Privacy and Data Security

Many caregivers expressed privacy and security concerns about using the app, especially if they were requested to provide sensitive information about their children (eg, home address and activities and locations where children engage in those activities). They did state that they would trust apps downloaded from officially certified app shops or promoted by official agencies such as preschools.

Web-Based Survey Results

Participants

In total, 1505 valid questionnaires were collected through the Web-based survey, including 687 from men (45.65%) and 818 from women (54.35%). The respondents supervised 807 children aged 0 to 3 years (53.62%) and 698 children aged 4 to 6 years (45.97%). The proportion of male to female children supervised was 1.33 (Table 1).

Content of App-Based Intervention

Of the 1505 participants who completed the Web-based survey, 1313 (87.24%) expected the app would teach them relevant knowledge concerning unintentional injury prevention for children aged 0 to 6 years. Participants felt it would be valuable to learn knowledge about preventing the 10 major causes of child unintentional injury at different rates: contact with heat and hot substances (534/1313, 40.67% felt it would be valuable to learn); inanimate mechanical forces (including being pinched between 2 surfaces, such as in the doors of elevators, or cut or punctured by sharp objects; 520/1313, 39.60%); falls (449/1313, 34.20%); transport crashes (362/1313, 27.57%); unintentional threats to breathing (361/1313, 27.49%); exposure to animate mechanical forces (317/1313, 24.14%);exposure to smoke, fire, and flames (210/1313, 15.99%); unintentional poisoning by and exposure to noxious substances (146/1313, 11.12%); unintentional drowning and submersion (113/1313, 8.61%); and exposure to electrical currents (108/1313, 8.23%). The differences were statistically significant (χ^2_9=989.7; P<.05; Table 2).

Table 2. Number of participants who expressed a desire to learn knowledge about preventing major types of child unintentional injury causes.

Cause of injury	n (%)
Contact with heat and hot substances	534 (40.67)
Exposure to inanimate mechanical forces	520 (39.60)
Falls	449 (34.20)
Transport crashes	362 (27.57)
Unintentional threats to breathing	361 (27.49)
Exposure to animate mechanical forces	317 (24.14)
Exposure to smoke, fire, and flames	210 (15.99)
Unintentional poisoning by and exposure to noxious substances	146 (11.12)
Unintentional drowning and submersion	113 (8.61)
Exposure to electrical currents	108 (8.23)

Forms of App-Based Learning

The 5 most frequent learning forms that participants stated they would like to use were as follows: (1) short written alarms or warnings with pictures, (2) video testimonials, (3) cartoon vignettes, (4) pictures, and (5) interactive games, which were, respectively, preferred by 62.30% (818/1313), 54.53% (716/1313), 40.67% (534/1313), 38.31% (503/1313), and 28.03% (368/1313) of respondents (note that respondents were permitted to select multiple preferred learning forms). Differences were significant (χ^2_4=393.0; P<.05).

Preferred Times, Frequency, and Duration for Using App-Based Learning

Caregivers varied in their preferred frequency for using an app to learn about unintentional child injury prevention. Just over half the sample (705/1313, 53.69%) preferred to use the app twice a week, followed by once a week (337/1313, 25.67%) and once a day (215/1313, 16.37%). Differences were significant (χ^2_2=455.4; P<.05). Participants were approximately evenly split concerning their preference to use the app in the evening (376/1313, 28.64%), afternoon (332/1313, 25.29%), or noontime (265/1313, 20.18%; χ^2_2=25.2; P<.05). Preferred durations for each session of app-based learning were 6 to 10 min (488/1313, 37.17%), greater than or equal to 11 min (418/1313, 31.84%), and 3 to 5 min (334/1313, 25.44%; χ^2_2=42.0; P<.05).

Suggested Framework to Design the App Based on Focus Group and Web-Based Survey

On the basis of responses from the focus groups and the Web-based survey, as well as the 3 principles of TPB (attitudes toward the behavior, subjective norms, and perceived behavioral control) and the 3 aspects of the FRAME model (learner, devices, and social aspects), we developed a framework to guide design of an app to teach caregivers knowledge about preventing child unintentional injury. The design included 5 primary components: (1) content design, (2) functional design, (3) interface design, (4) installation and registration, and (5) privacy and data security, plus 15 subcomponents. Details, including the theoretical basis for each recommendation of the participants, appear in Table 3.

Table 3. Framework for developing an app to teach knowledge about preventing child unintentional injury to Chinese caregivers of children aged 0 to 6 years.

Requirements	Description	Theoretical basis
Content, form, time, frequency, and duration of knowledge dissemination		
Content	Ten major unintentional injury causes: contact with heat and hot substances; inanimate mechanical forces; falls; transport crashes; unintentional threats to breathing; animate mechanical forces; exposure to smoke, fire, and flames; unintentional poisoning by and exposure to noxious substances; unintentional drowning and submersion; and exposure to electrical currents, with the suggested proportions for knowledge disseminations of these 10 injury causes at the suggested proportions of 5:5:4:4:4:3:2:1:1:1	Learner aspect (FRAME[a] model)
	Tied to theory-based goals to improve attitudes, alter subjective norms, and increased perceived control to perform child unintentional injury prevention behaviors	Attitudes, subjective norms, and perceived control (TPB[b])
Form	Short written alarms or warnings with pictures, cartoon vignettes, video testimonials, and interactive games, with the suggested proportions of 2:2:1:1 across the 4 forms	Learner aspects (FRAME model)
Time	Preferred time to learn (evening, afternoon, or noon)	Learner aspects (FRAME model)
Frequency	Twice a week	Learner aspects (FRAME model)
Duration	No more than 5 min per time	Learner aspects (FRAME model)
Other attributes	Professionally disseminated contents	Learner aspects (FRAME model)
	Using plain language	Learner aspects (FRAME model)
	Easy to implement	Learner aspects (FRAME model)
Functions		
Interactive style	Regular communication between users and experts through online chat, forums, and message boards; frequently asked questions module; and Web-based customer service agents and ask-and-answer service	Social aspects (FRAME model); attitude and subjective norms (TPB)
Survey and feedback	Use Web-based and printed questionnaires and short message service text message reminders for surveys and to collect feedback, no more than once every 2 months	—[c]
	After each survey, motivate users through virtual rewards	—
Personalized customization	Allow personalized customization of contents, forms, frequency, and interface of app intervention	Learner aspect (FRAME model)
Design of the interface		
Interface design	Offer several choices of app interfaces: cartoon interface preferred for default	Device aspect (FRAME model)
Installation and registration		
Registration	Simple app registration procedure	Device aspect (FRAME model)
Memory space	Minimize the size of app so smartphone storage is not used excessively	Device aspect (FRAME model)
Informed background	Provide informational background to users' informed consent process, so they understand before downloading and using the app intervention	Device aspect (FRAME model)
Manual and help module	Provide a user manual and help module	Device aspect (FRAME model)
Privacy protection and data security		
Personal privacy and data security	Do not collect sensitive individual or family information such as name, home address, or family income	Device aspect (FRAME model)
	Ensure safe sharing and storage of data, for example, by using an individualized password	Device aspect (FRAME model)

[a]FRAME: Framework for the Rational Analysis of Mobile Education.

[b]TPB: Theory of Planned Behavior.

[c]Not applicable.

Discussion

Principal Findings

Using a mixed-method study, we explored the preferences of caregivers of Chinese children aged 0 to 6 years for an app offering knowledge about child unintentional injury prevention. Study results helped us establish a framework for the app that includes 5 primary components (content design, functional design, interface design, installation and registration, and privacy and security) and 15 subcomponents focused on educating Chinese caregivers on child unintentional injury prevention strategies. The framework in Table 3 lists details to design the app, including the theoretical basis for those details.

Comparison With Prior Work

In addition to replicating previous work underlying the urgency to implement unintentional injury prevention education strategies for caregivers of young children in China [9], this study explored caregivers' attitudes and preferences for app-based learning, offering comprehensive and valuable guidance for the design of an app-based intervention. A few previous studies, generally based on small-sample qualitative designs, proposed some fragmented recommendations on the design of health-related apps that parallel some of our findings [30-34]. For example, Gkatzidou et al [30] suggested that privacy and security, credibility, user journey support, and the task-technology-context fit from the patient's perspective should be considered in the design of health care apps to maximize their acceptability. Similarly, Curtis et al [34] reported preferences for a healthy eating app targeting parents about child weight management; they identified 4 main themes for app design: app features, time saving and convenience, aesthetics, and gamification.

Our findings agree to a large extent with previous work in the design of app interventions, as we uncovered user preferences to emphasize a user-centered design [35,36], including particular emphasis on the importance of inclusion of games [35] and on the protection of user privacy [36].

Implications

This study has important implications. First, the research will lead to the development of an app designed to help caregivers learn how to reduce injury risk among their children. Although mobile device interventions are most effective when they are based upon theory [37], many existing health-related mobile apps do not incorporate features derived from evidence-based theoretical frameworks, health behavior change theories, and clinical guidelines [38,39]. We propose research grounded in theory that will lead to the development of an evidence-based intervention program. More broadly, the framework we present may prove valuable to guide the design of app-based interventions to target other diseases and injuries in China, as users' preferences are likely to be similar across content areas. With cultural tailoring, it might also be useful in other countries, and ultimately, the recommendations could facilitate using smartphone technology to improve public health in China and beyond. Furthermore, our methodological strategies proved useful and could be replicated. The 2-step mixed-method approach could be extended to guide the development of theory-driven frameworks or interventions for various health issues in various cultural contexts.

Strengths

Our study has 2 strengths. First, we adopted a mixed-method design with large sample sizes. We conducted a rigorous mixed-method study and analyzed data and reported results in accordance with COREQ [28] and CHERRIES [29] guidelines. In particular, we adopted the principles of an effective intervention program (eg, implementing various learning methods, sufficient dosage, theory driven, and appropriately timed) [20] and used the TPB [21] and the FRAME model [23] to design the semistructured discussion guide for focus groups. Use of the TPB and the FRAME model ensured that the framework we proposed both addresses the needs of users (FRAME) and creates a situation that will optimally encourage appropriate behavior change among Chinese caregivers and yield reduced risk of unintentional injury among their young children (TPB).

Second, we implemented rigorous science to develop the focus group guides and the questionnaire for the Web-based survey. Methodological strategies included 2-round discussions within the research member group and pilot tests to improve the experimental stimuli and ensure they were grounded in theory and constructed to yield the information we desired.

Our results also extend fragmented recommendations from previous publications [30,32,33] and provide systematic qualitative and quantitative guidance to design an effective app-based intervention. We expanded previous recommendations for privacy and information security by providing detailed suggestions concerning handling of sensitive information.

Limitations

This study has several limitations. First, focus groups were conducted using standard procedures and moderated by skilled individuals, but like all focus groups, the conversation was susceptible to bias and to swaying of individual opinions by dominant participants [40]. Second, the app framework was generated under the context of Chinese culture and the primary factors for child injury in China. The framework might need tailoring and adjustment if it were applied to other age groups, other diseases, or other countries.

Conclusions

A theory-driven and evidence-based framework was established to guide the design of an app-based unintentional injury intervention program for the caregivers of Chinese children aged 0 to 6 years.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (No. 81573260). The funding body had no role in the design, collection, analysis, or interpretation of this study.

Authors' Contributions

GH and PN contributed to the design of the study. PN performed the data analysis. PN, DG, and GH drafted the manuscript. All authors finalized and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Procedure and discussion guide for focus groups.

[DOCX File, 21KB - mhealth_v7i4e11957_app1.docx]

Multimedia Appendix 2

Web-based survey questionnaire.

[DOC File, 39KB - mhealth_v7i4e11957_app2.doc]

Multimedia Appendix 3

Data reporting guidelines, checklist for reporting results of internet E-Surveys (CHERRIES).

[DOCX File, 20KB - mhealth_v7i4e11957_app3.docx]

References

1. Institute for Health Metrics and Evaluation Viz Hub. GBD Compare URL: https://vizhub.healthdata.org/gbd-compare/ [accessed 2019-03-02] [WebCite Cache ID 76ZKxKh3K]
2. Kendrick D, Mulvaney CA, Ye L, Stevens T, Mytton JA, Stewart-Brown S. Parenting interventions for the prevention of unintentional injuries in childhood. Cochrane Database Syst Rev 2013(3):CD006020. [doi: 10.1002/14651858.CD006020.pub3] [Medline: 23543542]
3. Omaki E, Rizzutti N, Shields W, Zhu J, McDonald E, Stevens MW, et al. A systematic review of technology-based interventions for unintentional injury prevention education and behaviour change. Inj Prev 2017 Dec;23(2):138-146. [doi: 10.1136/injuryprev-2015-041740] [Medline: 26787740]
4. Shields WC, McDonald EM, Stepnitz R, McKenzie LT, Gielen AC. Dog bites: an opportunity for parent education in the pediatric emergency department. Pediatr Emerg Care 2012 Oct;28(10):966-970 [FREE Full text] [doi: 10.1097/PEC.0b013e31826c6c13] [Medline: 23023457]
5. van Beelen ME, Beirens TM, den Hertog P, van Beeck EF, Raat H. Effectiveness of web-based tailored advice on parents' child safety behaviors: randomized controlled trial. J Med Internet Res 2014;16(1):e17 [FREE Full text] [doi: 10.2196/jmir.2521] [Medline: 24463421]
6. Hu G, Baker TD, Baker SP. Injury control in China: priorities and actions. Lancet 2009 Jan 17;373(9659):214. [doi: 10.1016/S0140-6736(09)60068-4] [Medline: 19150701]
7. Hu G, Baker TD, Li G, Baker SP. Injury control: an opportunity for China. Inj Prev 2008 Apr;14(2):129-130. [doi: 10.1136/ip.2007.017822] [Medline: 18388234]
8. Jiang W, Wei J, Liu W, He L, Mo X. The situation of pre-school children's injuries and the status of parents' cognition. Journal of Preventive Medicine Information 2011;27:659-662 (In Chinese).
9. Zhou W, Jiang Y, Guan L, Li Q, Huang L. Researches about the prevalence of pre-school children's injuries and the status of their parents/guardians' kap in a kindergarten, shanghai. China Practical Medical 2009;4:257-259 (In Chinese). [doi: 10.14163/j.cnki.11-5547/r.2009.20.008]
10. China Internet Network Information Center. The 41st China statistical report on internet development URL: http://www.cac.gov.cn/2018-01/31/c_1122347026.htm [accessed 2018-07-28] [WebCite Cache ID 71FMMOMG6]
11. National Bureau of Statistics of China. China Statistical Yearbook. Beijing: China Statistics Press; 2017.
12. Mehl G, Labrique A. Prioritizing integrated mHealth strategies for universal health coverage. Science 2014 Sep 12;345(6202):1284-1287. [doi: 10.1126/science.1258926] [Medline: 25214614]
13. Hampton T. Recent advances in mobile technology benefit global health, research, and care. J Am Med Assoc 2012 May 16;307(19):2013-2014. [doi: 10.1001/jama.2012.4465] [Medline: 22665083]

14. Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. E-health in low- and middle-income countries: findings from the Center for Health Market Innovations. Bull World Health Organ 2012 May 1;90(5):332-340. [doi: 10.2471/BLT.11.099820] [Medline: 22589566]
15. van Mechelen MD, van Mechelen W, Verhagen EA. Sports injury prevention in your pocket?! Prevention apps assessed against the available scientific evidence: a review. Br J Sports Med 2014 Jun;48(11):878-882. [doi: 10.1136/bjsports-2012-092136] [Medline: 23511697]
16. Omaki E, Shields WC, McDonald E, Aitken ME, Bishai D, Case J, et al. Evaluating a smartphone application to improve child passenger safety and fire safety knowledge and behaviour. Inj Prev 2017 Dec;23(1):58. [doi: 10.1136/injuryprev-2016-042161] [Medline: 27597399]
17. Wilson H, Stoyanov SR, Gandabhai S, Baldwin A. The quality and accuracy of mobile apps to prevent driving after drinking alcohol. JMIR Mhealth Uhealth 2016 Aug 08;4(3):e98 [FREE Full text] [doi: 10.2196/mhealth.5961] [Medline: 27502956]
18. Rasche P, Mertens A, Bröhl C, Theis S, Seinsch T, Wille M, et al. The "Aachen fall prevention App" - a smartphone application app for the self-assessment of elderly patients at risk for ground level falls. Patient Saf Surg 2017;11:14 [FREE Full text] [doi: 10.1186/s13037-017-0130-4] [Medline: 28503199]
19. Burgess JD, Cameron CM, Watt K, Kimble RM. Cool Runnings - an app-based intervention for reducing hot drink scalds: study protocol for a randomised controlled trial. Trials 2016 Dec 03;17(1):388 [FREE Full text] [doi: 10.1186/s13063-016-1521-z] [Medline: 27488411]
20. Nation M, Crusto C, Wandersman A, Kumpfer KL, Seybolt D, Morrissey-Kane E, et al. What works in prevention. Principles of effective prevention programs. Am Psychol 2003;58(6-7):449-456. [Medline: 12971191]
21. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckman L, editors. Action Control: From Cognition to Behavior. Berlin: Springer; 1985:11-39.
22. Riebl SK, Estabrooks PA, Dunsmore JC, Savla J, Frisard MI, Dietrich AM, et al. A systematic literature review and meta-analysis: the Theory of Planned Behavior's application to understand and predict nutrition-related behaviors in youth. Eat Behav 2015 Aug;18:160-178. [doi: 10.1016/j.eatbeh.2015.05.016] [Medline: 26112228]
23. Ally M. Mobile Learning: Transforming the Delivery of Education and Training. Edmonton, AB: Athabasca University Press; 2009.
24. Kitzinger J. Qualitative research. Introducing focus groups. Br Med J 1995 Jul 29;311(7000):299-302 [FREE Full text] [Medline: 7633241]
25. Morgan DL. Focus Groups as Qualitative Research. London: Sage Publications; 1998.
26. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. Technode. 2017. WeChat User & Business Ecosystem Report 2017 URL: https://technode.com/2017/04/24/wechat-user-business-ecosystem-report-2017/ [accessed 2019-04-03] [WebCite Cache ID 77LbkH9DX]
28. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
29. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
30. Gkatzidou V, Hone K, Sutcliffe L, Gibbs J, Sadiq ST, Szczepura A, et al. User interface design for mobile-based sexual health interventions for young people: design recommendations from a qualitative study on an online Chlamydia clinical care pathway. BMC Med Inform Decis Mak 2015 Aug 26;15:72 [FREE Full text] [doi: 10.1186/s12911-015-0197-8] [Medline: 26307056]
31. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
32. Teo CH, Ng CJ, White A. What do men want from a health screening mobile app? A qualitative study. PLoS One 2017;12(1):e0169435 [FREE Full text] [doi: 10.1371/journal.pone.0169435] [Medline: 28060953]
33. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
34. Curtis KE, Lahiri S, Brown KE. Targeting parents for childhood weight management: development of a theory-driven and user-centered healthy eating app. JMIR Mhealth Uhealth 2015;3(2):e69 [FREE Full text] [doi: 10.2196/mhealth.3857] [Medline: 26088692]
35. Tonkin E, Jeffs L, Wycherley TP, Maher C, Smith R, Hart J, et al. A smartphone app to reduce sugar-sweetened beverage consumption among young adults in Australian remote indigenous communities: design, formative evaluation and user-testing. JMIR Mhealth Uhealth 2017 Dec 12;5(12):e192 [FREE Full text] [doi: 10.2196/mhealth.8651] [Medline: 29233803]
36. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Preferences for a mobile HIV prevention app for men who have sex with men. JMIR Mhealth Uhealth 2014;2(4):e47 [FREE Full text] [doi: 10.2196/mhealth.3745] [Medline: 25355249]

37. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
38. Al-Durra M, Torio M, Cafazzo JA. The use of behavior change theory in internet-based asthma self-management interventions: a systematic review. J Med Internet Res 2015;17(4):e89 [FREE Full text] [doi: 10.2196/jmir.4110] [Medline: 25835564]
39. West JH, Hall PC, Hanson CL, Barnes MD, Giraud-Carrier C, Barrett J. There's an app for that: content analysis of paid health and fitness apps. J Med Internet Res 2012;14(3):e72 [FREE Full text] [doi: 10.2196/jmir.1977] [Medline: 22584372]
40. Wong LP. Focus group discussion: a tool for health and medical research. Singapore Med J 2008 Mar;49(3):256-260 [FREE Full text] [Medline: 18363011]

Abbreviations

CHERRIES: checklist for reporting results of internet E-survey
COREQ: consolidated criteria for reporting qualitative research
FRAME: Framework for the Rational Analysis of Mobile Education
mHealth: mobile health
TPB: Theory of Planned Behavior

Edited by G Eysenbach; submitted 16.08.18; peer-reviewed by R Biviji, Z Gao; comments to author 13.12.18; revised version received 30.01.19; accepted 24.03.19; published 30.04.19.

Please cite as:
Ning P, Gao D, Cheng P, Schwebel DC, Wei X, Tan L, Xiao W, He J, Fu Y, Chen B, Yang Y, Deng J, Wu Y, Yu R, Li S, Hu G
Needs Analysis for a Parenting App to Prevent Unintentional Injury in Newborn Babies and Toddlers: Focus Group and Survey Study Among Chinese Caregivers
JMIR Mhealth Uhealth 2019;7(4):e11957
URL: http://mhealth.jmir.org/2019/4/e11957/
doi:10.2196/11957
PMID:31038465

Original Paper

A Facial Recognition Mobile App for Patient Safety and Biometric Identification: Design, Development, and Validation

Byoungjun Jeon[1], BS; Boseong Jeong[2], BS; Seunghoon Jee[2], BS; Yan Huang[3], BS; Youngmin Kim[3], MS; Gee Ho Park[3], BS; Jungah Kim[4], MS; Maierdanjiang Wufuer[5], MD, MS; Xian Jin[5], MD; Sang Wha Kim[5], MD; Tae Hyun Choi[6], MD, PhD

[1]Interdisciplinary Program for Bioengineering, Graduate School, Seoul National University, Seoul, Republic of Korea

[2]Oezsoft Inc, Seoul, Republic of Korea

[3]Interdisciplinary Program in Stem Cell Biology, College of Medicine, Seoul National University, Seoul, Republic of Korea

[4]Department of Chemistry, Seoul National University College of Natural Sciences, Seoul, Republic of Korea

[5]Department of Plastic and Reconstructive Surgery, Seoul National University College of Medicine, Seoul, Republic of Korea

[6]Institute of Human Environment Interface Biology, Department of Plastic and Reconstructive Surgery, Seoul National University College of Medicine, Seoul, Republic of Korea

Corresponding Author:
Tae Hyun Choi, MD, PhD
Institute of Human Environment Interface Biology
Department of Plastic and Reconstructive Surgery
Seoul National University College of Medicine
6th floor
101 Daehak-ro, Jongno-gu
Seoul, 03080
Republic of Korea
Phone: 82 220721978
Fax: 82 27665829
Email: psthchoi@snu.ac.kr

Abstract

Background: Patient verification by unique identification is an important procedure in health care settings. Risks to patient safety occur throughout health care settings by failure to correctly identify patients, resulting in the incorrect patient, incorrect site procedure, incorrect medication, and other errors. To avoid medical malpractice, radio-frequency identification (RFID), fingerprint scanners, iris scanners, and other technologies have been implemented in care settings. The drawbacks of these technologies include the possibility to lose the RFID bracelet, infection transmission, and impracticality when the patient is unconscious.

Objective: The purpose of this study was to develop a mobile health app for patient identification to overcome the limitations of current patient identification alternatives. The development of this app is expected to provide an easy-to-use alternative method for patient identification.

Methods: We have developed a facial recognition mobile app for improved patient verification. As an evaluation purpose, a total of 62 pediatric patients, including both outpatient and inpatient, were registered for the facial recognition test and tracked throughout the facilities for patient verification purpose.

Results: The app was developed to contain 5 main parts: registration, medical records, examinations, prescriptions, and appointments. Among 62 patients, 30 were outpatients visiting plastic surgery department and 32 were inpatients reserved for surgery. Whether patients were under anesthesia or unconscious, facial recognition verified all patients with 99% accuracy even after a surgery.

Conclusions: It is possible to correctly identify both outpatients and inpatients and also reduce the unnecessary cost of patient verification by using the mobile facial recognition app with great accuracy. Our mobile app can provide valuable aid to patient verification, including when the patient is unconscious, as an alternative identification method.

(JMIR Mhealth Uhealth 2019;7(4):e11472) doi:10.2196/11472

KEYWORDS

facial recognition; patient identification systems; biometric identification; patient safety; smartphone; mobile applications

Introduction

Background

For patient safety, the minimum requirements for patient identification in a hospital setting include 2 different patient-provided verifications that usually include patient name, the registered patient number, or date of birth [1-3]. Every visit to any hospital location for examination, treatment, or other health care services requires patient identification.

The purpose of patient verification is to minimize the potential for medical malpractice and other risks to patient safety. Several incidents because of lack of patient identification during treatment have been reported. In 2014, a physician got confused with 2 patients' surnames and gave the wrong patient's name to the surgeon over the phone, which resulted in heart surgery on the wrong patient [4]. In 2000, the New York State Health Department cited Staten Island University Hospital for failing to monitor and discipline its chief of neurosurgery, who, according to the state, had operated on the wrong side of a patient's brain for the second time in 5 years [5]. It was because of lack of confirmation on all the necessary information, in this case, computed tomography (CT) scans, before making an incision on the left cerebellum. The case could be prevented if all the information of the patient needed before surgery was confirmed.

Furthermore, in 1994, a similar error occurred in a lung cancer patient in Texas [6]. The surgeon admitted that he negligently removed the wrong lung, and it was discovered 1 week after the surgery when the patient himself found this after reviewing his medical records. These cases could be easily prevented by surgeons or other faculties by identifying the patient and checking all the necessary information before the operation. According to the Joint Commission, the number of reported sentinel events in 2017 was 400, which is only a small fraction of actual events [7].

To resolve the current issues with patient verification, some hospitals have adapted newer technology, such as disposable, scannable radio-frequency identification (RFID) bracelets, for more precise and easier patient verification [8-11]. As the typical identification method is to ask patients for their name; patient number, which is a complicated series of numbers; or other appropriate information, it is impossible to verify a patient when they are unconscious or when they are not able to remember them. The RFID bracelet, which can simply be scanned, facilitates an improved verification process. However, the limitation of employing physical objects such as RFID bracelets, cards, tokens, or keys is that the object must always be presented and kept secure [10,12,13]. In other words, if the physical objects such as RFID bracelets are lost or unable to scan, patient identification is not possible with the objects. Therefore, biometric measure, which represents an alternative method for patient verification without the necessity for physical objects, has emerged.

Biometrics refers to the recognition of individuals based on their anatomical, physiological, and/or behavioral characteristics, which permits identification without physical objects. Biometric options are not limited to fingerprint scanners but also include palm scanners, iris scanners, etc for patient verification [14,15]. However, there are limitations and disadvantages associated with these measures. Fingerprint collection requires patients to physically place their finger onto the scanner every time, which may facilitate disease transfer through physical contact with the scanner. Disease transfer through fingerprint includes a risk of transfer of infectious microorganisms that are enteric and respiratory pathogens [16]. Furthermore, unconscious patients are not able to place their finger onto the scanner by themselves. Although iris scans seem to be the ideal option for patient verification compared with other various verification methods, the main drawback is that it is difficult to scan the iris of an unconscious patient, and therefore, it is virtually impossible to verify unconscious patients with an iris scan [15,17]. Furthermore, when a surgical operation is required, patients must be put under anesthesia, negating the iris scanner as a valid option.

One interesting biometric is facial recognition, which does not have the disadvantages of the other biometrics options described above. Facial recognition does not require any physical contact with a device for recognition, and patients can be recognized even when they are unconscious. The comparison of facial characteristics obtained from a patient with stored or preregistered facial records in a database allows the recognition process to verify the patient. Previous research using Microsoft Kinect v2 sensor for patient verification demonstrated that facial features could be implemented for patient verification [18]. Furthermore, other prior work revealed that the CT scans can be used as facial recognition with moderately high accuracy [19]. However, certain sensors and specific actions were required to acquire and process facial information in prior studies. In addition, facial recognition systems require simple and direct processes to be useful in identifying patients in various conditions and hospital settings.

Objectives

Here, we have applied a facial recognition system to develop a hospital-friendly mobile app for patient verification. We evaluated the performance of the mobile app on a total of 62 hospital inpatients and outpatients. The aim of this study was to see if the performance of the mobile app is suitable in a hospital setting as an alternative method for patient verification.

Methods

The Facial Recognition System–Embedded Mobile App

We developed an Android-based app in the Eclipse environment using Java (Oracle Corporation). A facial recognition engine powered by Oezsoft Inc was modified to leverage the Native Development Kit (NDK) library and was used in Java through

the Java Native Interface. The NDK library was developed in Linux using a GNU Compiler Collection compiler. A mobile camera captures the characteristics of individuals' facial features as if a mobile phone was used to photograph the individual's face. The facial recognition engine extracts 27 major landmarks from the 2-dimensional facial image captured by the mobile camera. Then, the 3-dimensional facial learning data are used to compile the 3-dimensional portrait by extracting the major landmarks to create a comparable facial template, which is then registered in the mobile phone.

Mobile App Interface Design

The interface is designed to be easy to use and familiar to medical staff because the facial data collection time is short, and the patient would want to move quickly. The interface of the app contains 5 main recording menus: the registration, medical records, examinations, prescriptions, and appointments.

Registration

The registration part contains facial information acquiring step and basic personal information recording step to register a patient on the app environment. With a mobile camera in operation, the facial data are extracted from the image, which is stored in a 3-dimensional vector form and is used as a facial template. The basic personal information such as the name, date of birth, sex, and phone number was recorded in the app.

Medical Records

The medical record part contains selecting a department, outpatient, inpatient, and surgery. As each option was selected, the medical recording part is ready to be entered. Examination and prescription options were also selected for later review by other medical staff. All the records were uploaded to the app system and the hospital's electronic medical record (EMR) separately because the systems are not interworking yet. The records in the app were compared against the information recorded in the EMR to see if the information in the app is preserved throughout the facilities and matches with the EMR. After the facial recognition test of each patient in each facility, the recorded information in the app was compared with the EMR as a preliminary test for future application and EMR interlocking system.

Examinations and Prescriptions

The examination and prescription parts show whether the patient has been to an ordered facility for examination or prescription. A medical staff at the facility can check whether the examination or prescription has been done. For instance, the examination list shows a red mark if the ordered examination or prescription has not been done and green mark if the ordered examination or prescription has been done.

Appointments

The appointments let the medical staff arrange the next schedule on the app so that the different medical staff can see the record when the patient arrives for the next visit. The listed options for the appointment are outpatient, inpatient, surgery, and the list of departments the patient would visit.

Ethics Statement

This project was approved by the Institutional Review Board (IRB) of Seoul National University Hospital (IRB approval number: 1701-027-821). All experiments dealing with human or human products were conducted with informed consent and carried out in accordance with the relevant guidelines and regulations. Every patient and their parents were told about the facial recognition process of how the verification using a mobile facial app would be used for their identification. After giving them detailed information about identification, permission from both patients and their parents were obtained by signing the informed consent form.

Outpatient Studies

Medical staff from the plastic surgery department of Seoul National University Children's Hospital were instructed to generate facial profiles of registered outpatient plastic surgery patients by taking pictures with the provided mobile device. All patients were first verified using routine verification methods such as matching patient number or name. Then, a medical staff registered patients for facial recognition upon admission. When a patient entered the examination room of the outpatient clinic, the patient was then verified using the facial recognition mobile app. The mobile app was used as a secondary method to verify the patient and to record patient information, necessary examinations, prescription information, and to make a reservation for the patient's next visit. When the patient moved through the facility to additional examination or treatment rooms, medical staff used the facial recognition mobile app to verify the patient and to determine the required examinations and/or treatments. Facility visits were recorded, which indicates the number of different facilities where each patient was verified by both traditional identification and the facial recognition mobile app. Each time the facial recognition mobile app was operated to verify a patient, it was recorded as pass verification when the mobile app identified the patient correctly, otherwise it was recorded as fail verification. Male and female verification numbers were counted, which indicates the sum of facility visits by each patient for whom facial verification was performed. A total of 30 outpatients were all pediatric patients, and the patients were aged between 1 and 13 years. All patients were conscious because they only visited the hospital for diagnostic or preoperative examination purposes.

Inpatient Studies

Inpatients scheduled for surgical operation in the plastic surgery department as well as intensive care unit patients were also selected for facial verification. Similar to outpatients described above, hospital faculty verified inpatients first with the routinely used patient verification method and then with facial recognition mobile app at the entrance of the operating room, in the operating room, and in the recovery room. After each facial recognition verification, patients were treated, examined, or prescribed medication. As stated in the outpatient section above, facility visits and pass or fail facial verification numbers were recorded separately. Facility visits indicate the number of different facilities visited by each patient. If the patient was correctly identified at the facility, it was recorded as a pass verification, and if the patient was identified incorrectly, it was

recorded as a fail verification. Among inpatients, a total of 24 visited the hospital for day surgery, which requires only a single day for hospitalization, surgical operation, and recovery. Of the remaining 8 inpatients, 1 was admitted to the intensive care unit. Facial features of day-surgery inpatients were registered when they arrived at the patient day surgery center (PDSC). Before entering the surgical operating room and during the operation, a member of the hospital staff confirmed the patient number for verification purposes and then used the mobile app for facial recognition. Nonday-surgery inpatients were registered the day before the operation and tracked on operation day as described for day-surgery patients. As they did not leave the hospital the same day, identification was confirmed in the recovery room after completion of surgery. A total of 32 inpatients were all pediatric patients and were aged between 1 and 17 years. Among 32 inpatients, 31 of them were unconscious in the operating room because of anesthesia. All inpatient facial data were captured when conscious before entering the operating room.

Statistical Analysis

Total comparison trials are calculated based on one-to-one matching with the number of sequentially registered patients and each patient's facility visits. The sum of each patient's facility visits is considered as true value. Here, the sensitivity is the percentage of correct match of a patient at various facilities (true-positive value). The specificity is the percentage of incorrect match against all previously registered patients (true-negative value). Sensitivity and specificity are calculated as below:

Sensitivity = (True Positive) / (True Positive + False Negative) (1)

Specificity = (True Negative) / (True Negative + False Positive) (2)

Results

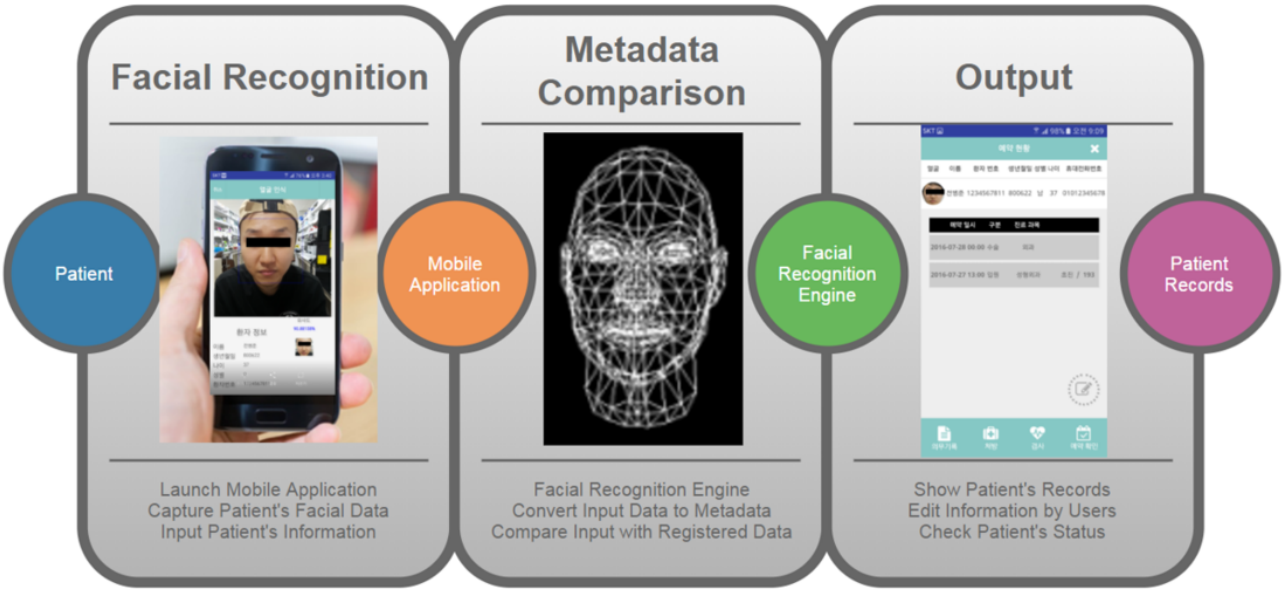
Mobile App Development

With a Samsung Galaxy Note 3, the mobile app was able to capture facial features and process facial recognition within 0.5 seconds. The minimum matching percentage of facial recognition to verify a patient was set to 95%. The minimum facial width and length between irises was set to 30 pixels. Recognizable maximum vertical and horizontal facial posture angles were 15 degrees and 40 degrees, respectively.

When verifying a patient, the newly extracted facial image is compared with the initially stored facial template. If the similarity between the 2 templates is over 95%, the facial recognition engine recognizes the images as belonging to the same patient. When recognizing a patient, the app calculates the Euclidean distances between 3-dimensionally repositioned landmarks of templates, weighing the distances of the eyes and nose. The accuracy of the facial recognition mobile app tested at Seoul National University Children's Hospital was 99%.

The facial recognition engine supports the Android, IOS, Embedded Linux, Windows, IBM AIX, and Sun Solaris operating systems. As the facial template size is minuscule, it can be stored in a secured server, and a copy of the template can be stored in a mobile device for offline use. The overview of the facial recognition app process is as shown in Figure 1. When the patient is first enrolled, the mobile camera captures the patient's face image, and the facial recognition engine creates and stores the facial template. Then, when patient identification is needed at another facility, the facial recognition engine captures the patient's face image, generates a new facial template, and compares the input (newly generated facial template) with the registered data (initially stored facial template). After facial recognition successfully identifies the patient, the output shows the patient's records, and detailed information of the patient can be edited. Detailed hospital-wide studies and design of the app are described in the section below.

Figure 1. Process of facial recognition application.



Outpatient Studies

A total of 30 hospital outpatients were registered and tracked (Figures 2 and 3). Figures 2 and 3 demonstrate the patient registration process and the facial recognition process. Registration of a patient requires an initial process of capturing the patient's facial information. There is a registration button at the bottom of the main screen of the app (Figure 2). To register a patient's facial data, patient number, name, date of birth, and phone number are required to be typed as shown in the second image of Figure 2. Capturing the patient's facial data and an example of the registered patient's facial image are shown in Figure 2.

Tracking a patient requires capturing patient's facial image to compare with registered data. There is a camera button at the upper right corner of the main screen of the app (Figure 3). The steps involved in capturing the patient's facial data and comparing the input with the registered data are shown in Figure 3 (see center). The last image of Figure 3 is an example screen of a successfully identified patient.

Average patient age was 5 years because of the fact that the research was done at the children's hospital where the main plastic surgery is to correct burns, birth defects, etc. Each patient visited an average of 4 different hospital facilities and verified their identification with the facial recognition mobile app (Table 1).

Figure 2. Registration of facial template.

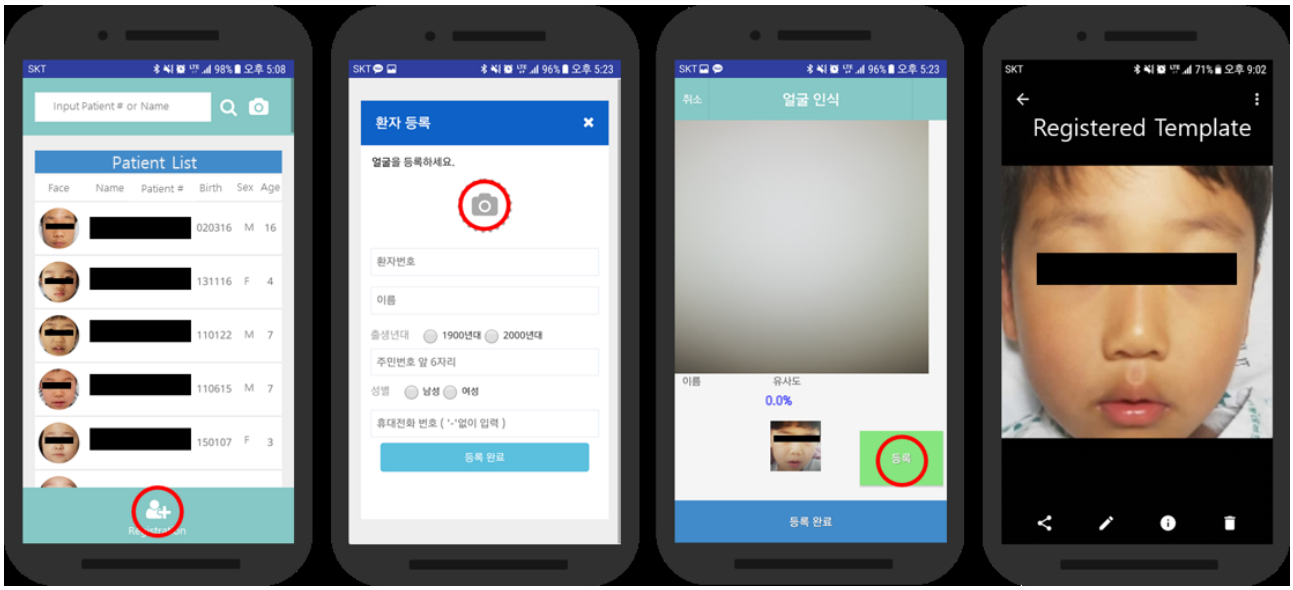


Figure 3. Facial recognition process.

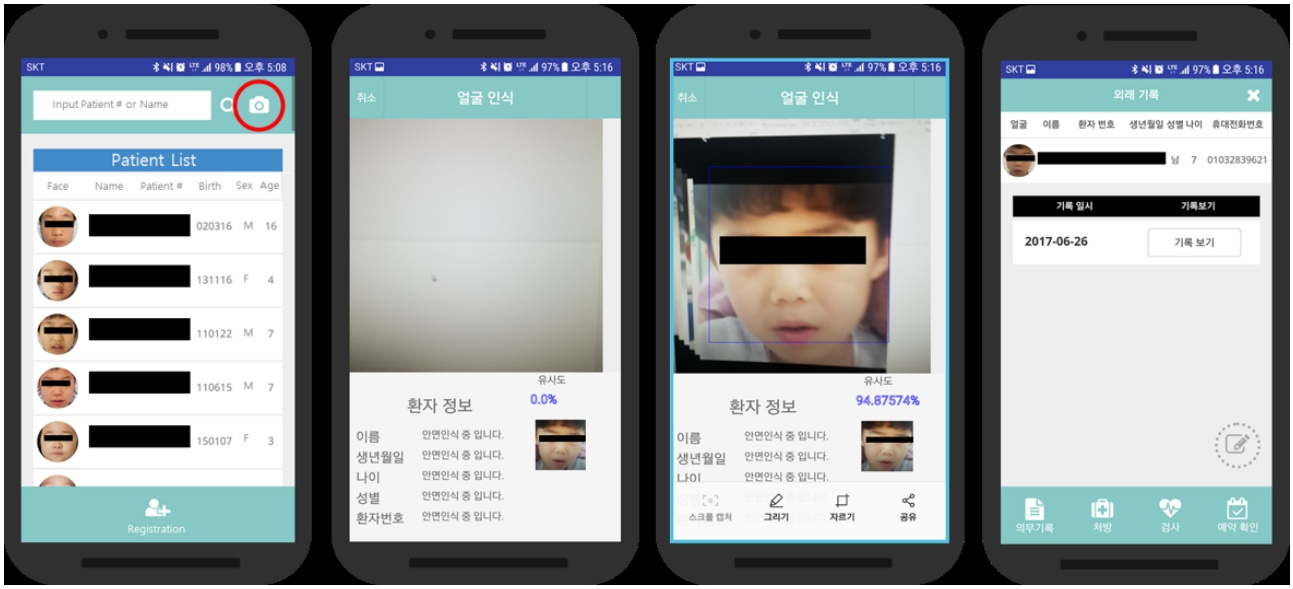
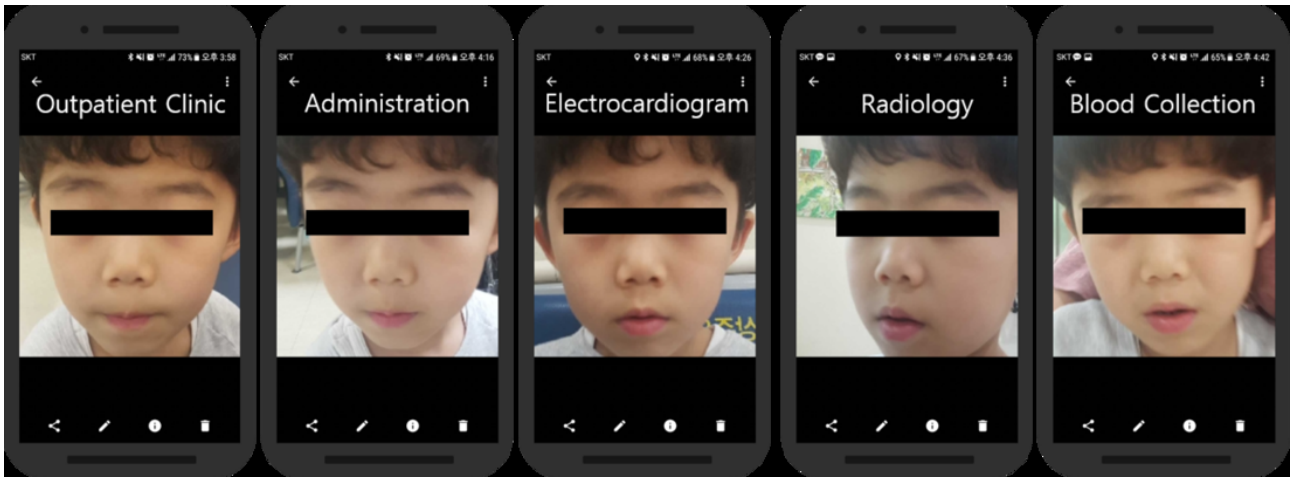


Table 1. Number of outpatient verification trial results.

Characteristic	Pass verification, n	Fail verification, n
Age (years)		
0-10	28	0
11-20	2	0
Sex		
Male	19	0
Female	11	0
Facility visit[a]		
Male	86	0
Female	50	0

[a]Total number of different facilities where each patient was verified using both traditional methods and the facial recognition mobile app.

Figure 4. Verified outpatient's input images in various facility.



When visiting different facilities, all 30 patients were successfully verified by the facial recognition system, and the images recognized as *correct* were captured and stored separately in a secure hard drive. An example of different facility visits by an outpatient and the images captured when verifying a patient using the facial recognition system is shown in Figure 4.

Ordered patient examinations and treatments were also recorded using the app and uploaded to the database so that hospital staff could determine whether the patient had been treated or examined appropriately. As shown in Figure 5, green color indicates that the examination was successfully confirmed by hospital staff; X-ray image, complete blood count data, and electrocardiography data were shown here as an example of records.

Appointment records were also updated in the app to facilitate scheduling and verification of future visits. The record includes options for inpatient, outpatient, department, and the date (Figure 6). All the information recorded in the app program is compared with the EMR, and it showed that the previously recorded information can be maintained elsewhere after the facial recognition process and can be utilized in the EMR interworking system in the future.

Inpatient Studies

A total of 32 inpatients were registered and tracked (Figures 2 and 3). Average patient age was 5 years because of the same reason stated in the outpatient section above. In the operating room, 31 out of 32 inpatients were identified with the mobile facial recognition app under anesthesia. In addition, the app was used to confirm the surgical details. During recovery, day-surgery inpatients were verified with the facial recognition mobile app and treated or prescribed in the recovery room as indicated by the app. When patients were ready to leave the hospital, their identification was confirmed via the facial recognition mobile app once again at the PDSC. Each patient was identified an average of 4 times after their facial features had been registered before entering the surgical operating room, which was recorded as a facility visit. The single unrecognizable patient was not recognized after surgery because of the compression bandage covering the patient's face (Table 2).

A total of 31 of 32 inpatients were recognized and verified by the facial recognition mobile app even after facial surgery. An example of different facility visits by an inpatient and the images captured when verifying a patient using the facial recognition system are shown in Figure 7.

Figure 5. Medical recordings and examination.

Figure 6. Appointment record.

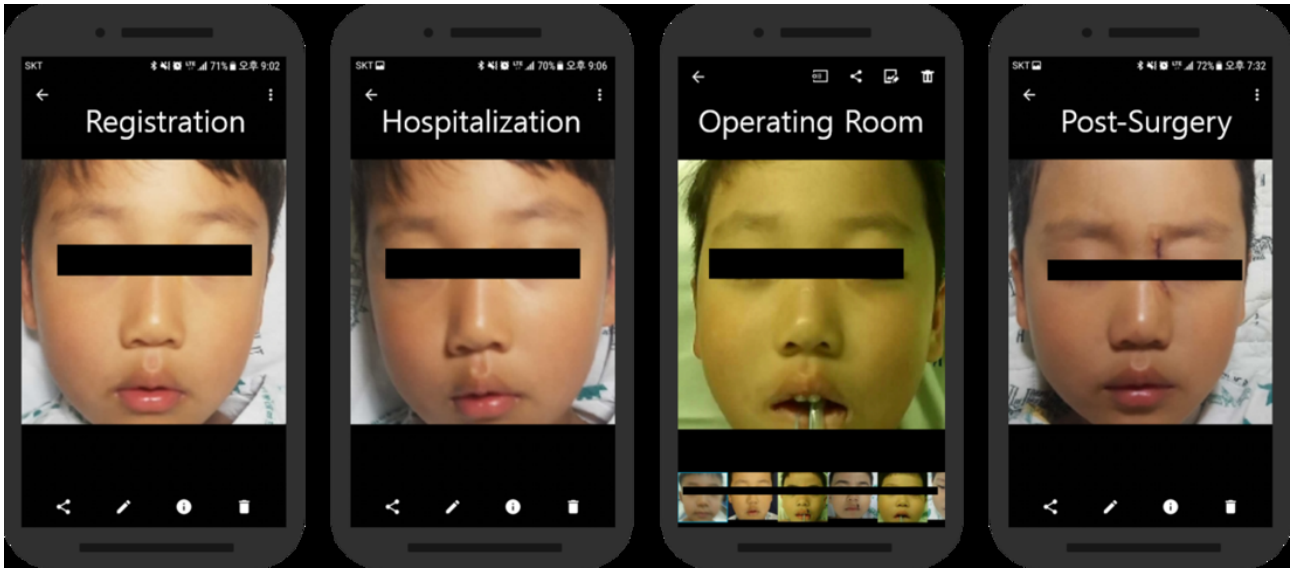
Table 2. Number of inpatient verification trial results.

Characteristic	Pass verification, n	Fail verification[a], n
Age (years)		
0-10	27	0
11-20	4	1
Sex		
Male	18	0
Female	13	1
Facility visit[b]		
Male	89	0
Female	60	1

[a]One failed verification was because of a compression bandage covering the patient's face after surgery.

[b]Total number of different facilities where each patient was verified using both traditional methods and the facial recognition mobile app.

Figure 7. Verified inpatient's input images in various facility.



Statistical Analysis

As the total number of patients is 62 and each patient's facility visits varies from 2 to 5, the total individual comparison trials based on one-to-many matching resulted in 8899. As shown in Tables 1 and 2, the sum of each patient's facility visits is 286. The sensitivity and specificity resulted in 99.7% (285/286) and 100.00% (8613/8613), respectively, with an accuracy of 99.99% (8898/8899).

Discussion

Principal Findings

Medical malpractice caused by failure in patient verification is a critical problem when considering patients' safety, as it can lead to operations on incorrect body sites and other further treatments. It was estimated that as many as 98,000 people die every year from medical errors [20]. Thus, accurate and precise patient identification is essential before every procedure in a hospital setting [21,22]. Traditional methods for patient identification include patient verification of name, date of birth, or hospital-provided patient number. This approach is not secure enough for patient verification when considering patient safety [23,24].

To mitigate current risks to patient safety, physical objects and biometric technologies such as RFID bracelets, fingerprint scanners, and iris scanners have been employed, but each approach has important limitations. Physical objects must always be present and kept secure to successfully confirm a patient's identity [10,25]. Fingerprint scanners may facilitate infection transmission, and iris scanners cannot be used on unconscious patients [26,27]. If these limitations are addressed, biometrics are one of the most reliable alternative options for patient verification under the minimum hospital requirements for patient verification. Facial recognition is another promising biometric approach to patient verification. It does not require any physical contact and can be used even when a patient is unconscious.

Our implementation of a facial recognition system through a mobile app resulted in 99.99% accuracy in patient verification. The accuracy of the facial recognition app would have been 100% had it not failed on a single patient whose face was covered with a compression bandage post surgery. With the exception of this case, the facial recognition mobile app was able to recognize and verify all other facial surgery patients. Moreover, patients under anesthesia in the operating room were recognizable and verified via the facial recognition mobile app.

For both patients and hospital staff, the traditional verification process of calling a patient's name, matching the recorded patient number by eye, or confirming the date of birth is time consuming and a suboptimal use of human resources [23]. Furthermore, asking patients for their patient number, requesting them to put their finger on a fingerprint scanner, or requesting to scan their eyes with an iris scanner are impossible under circumstances that often occur in a hospital setting. In contrast, mobile facial recognition systems can verify random patients at any time simply by taking pictures of the patient (whether conscious or unconscious) in lieu of making specific requests of the patient. When patient verification is imperative during serious circumstances, such as in the operating room, facial recognition systems provide a quick and simple way to identify patients, especially because the mobile app system does not require a space-demanding and expensive device such as a scanner. Through the mobile app, facial recognition technology is a convenient and secure alternative compared with other biometrics for patient verification because the app requires only a mobile phone. Thus, the facial recognition mobile app described herein is easily accessible and an accurate patient identification tool in the hospital setting.

Limitations

In the case of distance, there is a correlation with the size of the face captured by the camera. For the facial recognition app of this research, the camera resolution was fixed at 480×720 pixels, and it was developed considering the distance of less than 60 cm in terms of the usage form of a mobile phone. However, the distance is adjustable, and if the size of the face detected on the screen exceeds 120×120 pixels, face recognition is possible. For example, if we apply the developed app to the interactive

advertisement of New York Times Square, the app can recognize the faces of people 70 m away by using a high-resolution zoom camera and filtering out participants. As the example mentioned, distance is adjustable with a better camera, but at the same time, it requires greater amount of the central processing unit (CPU) resources for facial recognition when the image resolution is high. Thus, as the CPU of the mobile phone is slower than the personal computer, we aimed at recognizing the face within 0.5 seconds by fixing it at 480×720 resolution in this research.

As other studies on facial recognition stated, varying levels of light decreases the accuracy of the facial recognition process [18]. Currently, our facial recognition system is also limited by lighting. When the registered facial features were recorded under a bright light, patient recognition in darker places or a place where the light entered from a different angle, creating shade, resulted in delayed verification requiring up to 5 min for light correction to match images in the database. In addition, it was not possible to link the system to the EMR because the app was developed for test purposes only.

Comparison With Prior Work

Previous researches have implemented facial recognition technology for various purposes with a main focus on identifying an individual. However, one of the limitations is that it requires specific sensors and motions to be used in patient verification [18]. In a hospital setting, the identification process should be simple and effective, as patients could be in various conditions such as in an unconscious state. The facial recognition mobile system is effective and simple as it only requires a mobile phone and does not require specific motion such as turning and tilting the head in various directions for patient identification.

The other limitation of the facial recognition research conducted is the cost of processing facial recognition. Other research revealed that facial images derived from CT scans can be accurate [19]. However, the process of extracting the facial features of a patient and the cost of CT scans are not adequate for patient verification purpose. Even though CT scans are higher resolution compared with a mobile camera, simplicity and cost-effectiveness of the mobile facial recognition system are more suitable for the patient verification process.

Future Research

Further research is necessary to resolve the light sensitivity suffered by the app and to more accurately evaluate its performance for patient verification in the hospital, as well as to link the app with the hospital EMR to improve accessibility by hospital staff. Moreover, if a hospital-wide stationary camera-based facial recognition system is developed as a patient verification method, it could be used as a surveillance camera to monitor and verify patients entering various hospital facilities.

Conclusions

The facial recognition mobile app described here has been developed for patient verification purposes. The developed app contains 5 main parts suited for hospital usage: registration, medical records, examinations, prescriptions, and appointments. Hospital staff registered the facial feature of both outpatients and inpatients in the facial recognition database. The implementation of the facial recognition mobile app in the hospital setting proved a suitable alternative patient verification method with an accuracy of 99%. Once the facial recognition system is fully linked to EMR, it will be fully accessible to clinics and hospitals for patient verification.

Acknowledgments

This research was supported by a grant of the Korea Health Technology Research and Development Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (grant number: HI15C1165), and by grant no. 03-2017-0320 and 26-2015-0030 from the Seoul National University Hospital Research Fund. This work was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (Ministry of Education; NRF-2017R1D1A1B03030163) and by the Ministry of Trade, Industry, and Energy (Korea) under industrial technology innovation program No. 10079316, “Development of hand-held type non-invasive skin cancer diagnostic device.”

Conflicts of Interest

None declared.

References

1. The Joint Commission. 2018. National patient safety goals effective January 2018 URL: https://www.jointcommission.org/assets/1/6/NPSG_Chapter_HAP_Jan2018.pdf [accessed 2018-07-04] [WebCite Cache ID 70eZskvGJ]
2. Joint Commission on Accreditation of Healthcare Organizations. The joint commission announces 2014 national patient safety goal. Jt Comm Perspect 2013 Jul;33(7):1, 3-1, 4 [FREE Full text] [Medline: 24137865]
3. Makary MA, Daniel M. Medical error-the third leading cause of death in the US. Br Med J 2016 Dec 03;353:i2139. [doi: 10.1136/bmj.i2139] [Medline: 27143499]
4. News24. 2014 Aug 20. Wrong patient gets heart surgery URL: https://www.news24.com/SouthAfrica/News/Wrong-patient-gets-heart-surgery-20140820 [accessed 2018-12-30] [WebCite Cache ID 7535OFa4K]
5. Jennifer S. The New York Times. 2000 Mar 1. Surgeon Treated Wrong Side Of Two Brains, Albany Says URL: http://www.nytimes.com/2000/03/01/nyregion/surgeon-treated-wrong-side-of-two-brains-albany-says.html [WebCite Cache ID 70ea94nMP]

6. The Joint Commission. 2018. Summary Data of Sentinel Events Reviewed by The Joint Commission URL: https://www.jointcommission.org/assets/1/18/Summary_4Q_2017.pdf [accessed 2018-07-04] [WebCite Cache ID 70eaGHFpp]
7. Deseret News. 1995 Apr 2. HOSPITAL SETTLES SUIT OVER REMOVAL OF WRONG LUNG URL: https://www.deseretnews.com/article/412903/HOSPITAL-SETTLES-SUIT-OVER-REMOVAL-OF-WRONG-LUNG.html [accessed 2018-12-30] [WebCite Cache ID 7536SjOgz]
8. Jáuregui OI, Bruchanski L, Rizzato DA, Otero CM, Luna DR. Improving patient safety through the design and development of a computerized provider order entry for parenteral nutrition linked to a barcode medication administration record. Stud Health Technol Inform 2017;245:1038-1042. [doi: 10.3233/978-1-61499-830-3-1038] [Medline: 29295259]
9. Latham T, Malomboza O, Nyirenda L, Ashford P, Emmanuel J, M'baya B, et al. Quality in practice: implementation of hospital guidelines for patient identification in Malawi. Int J Qual Health Care 2012 Dec;24(6):626-633. [doi: 10.1093/intqhc/mzs038] [Medline: 22833617]
10. Hoffmeister LV, de Moura GM. Use of identification wristbands among patients receiving inpatient treatment in a teaching hospital. Rev Lat Am Enfermagem 2015;23(1):36-43 [FREE Full text] [doi: 10.1590/0104-1169.0144.2522] [Medline: 25806629]
11. Walley SC, Berger S, Harris Y, Gallizzi G, Hayes L. Decreasing patient identification band errors by standardizing processes. Hosp Pediatr 2013 Apr;3(2):108-117 [FREE Full text] [doi: 10.1542/hpeds.2012-0075] [Medline: 24340411]
12. Hain PD, Joers B, Rush M, Slayton J, Throop P, Hoagg S, et al. An intervention to decrease patient identification band errors in a children's hospital. Qual Saf Health Care 2010 Jun;19(3):244-247. [doi: 10.1136/qshc.2008.030288] [Medline: 20364035]
13. Phillips SC, Saysana M, Worley S, Hain PD. Reduction in pediatric identification band errors: a quality collaborative. Pediatrics 2012 Jun;129(6):e1587-e1593. [doi: 10.1542/peds.2011-1911] [Medline: 22566421]
14. Pato JN. Biometric Recognition. In: Pato JN, Millett L, editors. Biometric Recognition: Challenges and Opportunities. Washington, DC: The National Academies Press; 2010:182.
15. Jonas M, Solangasenathirajan S, Hett D. Patient identification, a review of the use of biometrics in the ICU. In: Vincent JL, editor. Annual Update In Intensive Care And Emergency Medicine 2014. Cham: Springer; 2014:679-688.
16. Jacobs JA, Van Ranst M. Biometric fingerprinting for visa application: device and procedure are risk factors for infection transmission. J Travel Med 2008;15(5):335-343 [FREE Full text] [doi: 10.1111/j.1708-8305.2008.00232.x] [Medline: 19006507]
17. Du EY. Review of iris recognition: cameras, systems, and their applications. Sensor Rev 2006 Jan;26(1):66-69. [doi: 10.1108/02602280610640706]
18. Silverstein E, Snyder M. Implementation of facial recognition with Microsoft Kinect v2 sensor for patient verification. Med Phys 2017 Jun;44(6):2391-2399. [doi: 10.1002/mp.12241] [Medline: 28370061]
19. Parks CL, Monson KL. Automated facial recognition of computed tomography-derived facial images: patient privacy implications. J Digit Imaging 2017 Dec;30(2):204-214 [FREE Full text] [doi: 10.1007/s10278-016-9932-7] [Medline: 28025730]
20. Donaldson MS, Corrigan JM, Kohn LT, editors. To Err Is Human: Building a Safer Health System. Washington, DC: National Academy Press; 2000.
21. Seferian EG, Jamal S, Clark K, Cirricione M, Burnes-Bolton L, Amin M, et al. A multidisciplinary, multifaceted improvement initiative to eliminate mislabelled laboratory specimens at a large tertiary care hospital. BMJ Qual Saf 2014 Aug;23(8):690-697. [doi: 10.1136/bmjqs-2014-003005] [Medline: 24915540]
22. Tridandapani S, Ramamurthy S, Provenzale J, Obuchowski NA, Evanoff MG, Bhatti P. A multiobserver study of the effects of including point-of-care patient photographs with portable radiography: a means to detect wrong-patient errors. Acad Radiol 2014 Aug;21(8):1038-1047 [FREE Full text] [doi: 10.1016/j.acra.2014.03.006] [Medline: 25018076]
23. Henneman PL, Fisher DL, Henneman EA, Pham TA, Campbell MM, Nathanson BH. Patient identification errors are common in a simulated setting. Ann Emerg Med 2010 Jun;55(6):503-509. [doi: 10.1016/j.annemergmed.2009.11.017] [Medline: 20031263]
24. García-Betances RI, Huerta MK. A review of automatic patient identification options for public health care centers with restricted budgets. Online J Public Health Inform 2012;4(1):- [FREE Full text] [doi: 10.5210/ojphi.v4i1.4011] [Medline: 23569629]
25. Perry DC, Scott SJ. Identifying patients in hospital: are more adverse events waiting to happen? Qual Saf Health Care 2007 Apr;16(2):160 [FREE Full text] [doi: 10.1136/qshc.2006.022202] [Medline: 17403768]
26. Drahansky M, Dolezel M, Urbanek J, Brezinova E, Kim TH. Influence of skin diseases on fingerprint recognition. J Biomed Biotechnol 2012;2012:626148 [FREE Full text] [doi: 10.1155/2012/626148] [Medline: 22654483]
27. Aslam TM, Tan SZ, Dhillon B. Iris recognition in the presence of ocular disease. J R Soc Interface 2009 May 06;6(34):489-493 [FREE Full text] [doi: 10.1098/rsif.2008.0530] [Medline: 19324690]

Abbreviations

CPU: central processing unit

CT: computed tomography
EMR: electronic medical record
IRB: Institutional Review Board
NDK: Native Development Kit
PDSC: Patient Day Surgery Center
RFID: radio-frequency identification

Edited by G Eysenbach; submitted 04.07.18; peer-reviewed by V Podder, A Arevian; comments to author 05.11.18; revised version received 30.12.18; accepted 27.01.19; published 08.04.19.

Please cite as:
Jeon B, Jeong B, Jee S, Huang Y, Kim Y, Park GH, Kim J, Wufuer M, Jin X, Kim SW, Choi TH
A Facial Recognition Mobile App for Patient Safety and Biometric Identification: Design, Development, and Validation
JMIR Mhealth Uhealth 2019;7(4):e11472
URL: https://mhealth.jmir.org/2019/4/e11472/
doi:10.2196/11472
PMID:30958275

Original Paper

Usage Patterns of GlucoNote, a Self-Management Smartphone App, Based on ResearchKit for Patients With Type 2 Diabetes and Prediabetes

Satoko Yamaguchi[1,2], MD, PhD; Kayo Waki[2,3], MD, PhD; Yasuhito Nannya[4], MD, PhD; Masaomi Nangaku[5], MD, PhD; Takashi Kadowaki[1,3], MD, PhD; Kazuhiko Ohe[6], MD, PhD

[1]Department of Prevention of Diabetes and Lifestyle-Related Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[2]Department of Ubiquitous Health Informatics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[3]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[4]Department of Pathology and Tumor Biology, Kyoto University, Kyoto, Japan
[5]Division of Nephrology and Endocrinology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[6]Department of Biomedical Informatics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Kayo Waki, MD, PhD
Department of Ubiquitous Health Informatics
Graduate School of Medicine, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo, 113-8655
Japan
Phone: 81 3 5800 9129
Fax: 81 3 5800 9129
Email: kwaki-tky@umin.ac.jp

Abstract

Background: Preventing progression from prediabetes to diabetes—or slowing the progression of diabetes—is an urgent task worldwide. Previous studies have shown that mobile health (mHealth) may powerfully support self-management for patients with type 2 diabetes. Certainly, mHealth improves health care efficiency and gives patients convenient access to self-management of their own health. Many health care apps are available right now, and their use in clinical studies with large-scale real-life data is expected. However, the usage patterns of those apps—especially in the absence of intervention by medical professionals—remain unknown.

Objective: We developed GlucoNote, an app that uses Apple's ResearchKit to support self-management for patients with type 2 diabetes and prediabetes; the app does not require prescription or intervention by medical professionals. We evaluated its usage patterns via a remotely conducted study.

Methods: iPhone users across Japan who have type 2 diabetes or prediabetes were free to download GlucoNote and to participate in the study after they provided consent electronically on the app. The 522 users who enrolled in the study within 1 year of its release were analyzed. We analyzed the retention rates of 357 participants who recorded at least 1 of 4 items—body weight, blood sugar, blood pressure, or dietary information. Characteristics of participants who used GlucoNote longer than 4 weeks (*robust users*) were compared with those of participants who did not (*nonrobust users*). The changes among robust users were evaluated.

Results: The median observation and retention durations were 382 days (interquartile range [IQR] 275-423) and 8 days (IQR 1-63), respectively. The retention rates for 2 days and for 4, 8, and 12 weeks were 0.627 (95% CI 0.575-0.675), 0.353 (0.304-0.403), 0.272 (0.227-0.319), and 0.220 (0.179-0.265), respectively. Men were more likely to be robust users than women (P=.02). At week 0, robust users were more likely than nonrobust users to have a higher daily energy intake (median 1595 [IQR 1198-1788] kcal vs 1451 [IQR 769-1657] kcal; P=.04) and have higher daily step counts (median 6108 [IQR 3797-9227] vs 5171 [IQR 2885-7258]; P=.001). Among robust users, body weight decreased from weeks 0 to 4 (mean 71.3 [SD 14.1] kg to 70.8 [SD 13.9] kg; P=.002) by mean 0.6% (SD 1.6).

Conclusions: GlucoNote offered a valuable opportunity to evaluate usage patterns of apps. Future challenges include improving low retention rates and evaluating their effects.

(JMIR Mhealth Uhealth 2019;7(4):e13204) doi:10.2196/13204

KEYWORDS

telemedicine; mHealth; self-management; diabetes mellitus

Introduction

Preventing progression from prediabetes to diabetes and slowing the progression of diabetes is an urgent task. Self-management is key to preventing that progression because there are a number of modifiable risks associated with prediabetes and type 2 diabetes; these include obesity, physical inactivity, and an unhealthy diet. Interventions using mobile phone apps to support self-management have proven to be effective in improving glycemic controls for diabetes patients [1,2]; such interventions also increase physical activity and reduce weight for overweight patients [3,4]. We previously developed DialBetics, a smartphone-based self-management support system for patients with type 2 diabetes. It provides real-time advice about lifestyle modifications based on patients' data measured at home and the physical activities and diet they recorded. The system lets the medical staff remotely monitor the data and alerts them when those data reach critical values so a physician could intervene if necessary. A 3-month randomized study of 54 patients with type 2 diabetes has demonstrated that glycemic control was improved in the DialBetics group, whereas it did not improve in the non-DialBetics control group [5]. Although DialBetics proved to be effective, the number of patients with access to it was limited because only those who physically visited the outpatient clinic, received face-to-face instructions, and gave written informed consent could participate. Patients with prediabetes who do not regularly visit medical facilities could not be reached. Moreover, the system required continuous monitoring by medical staff, which can be costly.

Currently, numerous apps that support self-management of diabetes or obesity are available to the public and do not necessarily require prescriptions. Although these apps can attract a large number of users and are thus potentially a powerful tool, very few of them have undergone scientific evaluation. Rivera et al reported that of the 393 apps commercially available for weight loss, only 3 (0.8%) underwent scientific evaluation and only 1 (0.3%) reported the involvement of health care experts in app development [6].

We developed a novel app for patients with diabetes and prediabetes, taking advantage of our experience in developing DialBetics. To make an app available to a large number of users and evaluate its usage patterns, we used ResearchKit by Apple, one of the frameworks to create apps for medical research; it was released in March 2015 [7]. ResearchKit offers customizable functionality commonly used for medical research and lets investigators recruit and enroll patients entirely remotely, providing users with questionnaires to determine eligibility, obtaining electronic informed consent, and collecting biometric data, including daily step counts [8,9]. Several studies have reported apps using ResearchKit and many more are expected. Chan et al detailed the Asthma Mobile Health Study, which recruited 7593 participants from across the United States and detected increased reporting of asthma symptoms in correlation with weather conditions [10]. Bot et al reported the mPower Study, with participation by 1087 patients with Parkinson disease (PD) and 5581 without it [11]. This study administered questionnaires and structured tasks related to PD, taking advantage of ResearchKit features to provide surveys and real-time active tasks such as tapping motor activities. Crouthamel et al reported the Patient Rheumatoid Arthritis Data from the Real World app and collected patient-reported data about rheumatoid arthritis, including assessment of wrist range movement, measured by the smartphone-embedded gyroscope and accelerometer [12]. These studies using ResearchKit benefitted from a large enrollment that overcame geographical barriers.

Here, we report on the findings from GlucoNote, a self-management support app for patients with type 2 diabetes and prediabetes that we developed using ResearchKit. GlucoNote lets users self-monitor the data that they measure at home, the diet information they enter manually, and the number of steps counted by their iPhones' built-in pedometers, which are displayed as graphs. At the same time, these data are sent to the server, letting the investigators evaluate the app's usage patterns.

Methods

Design of GlucoNote

The GlucoNote app was built with Apple's ResearchKit [7]. Users enter body weight, blood sugar levels (fasting blood sugar or postprandial plasma glucose), and blood pressure levels. Changes in these parameters are displayed as graphs (Figure 1). Steps are counted by each iPhone's built-in pedometer, and after the user sends the physical activity data to the server, step counts are displayed as a graph (Figure 1). Each time the user sends the physical activity data to the server, the step counts recorded in the past 30 days are also sent to the server, including the step counts of the 30 days recorded before study enrollment. If the user does not send the physical activity data, the step counts will not be sent to the server. GlucoNote facilitates easy input of dietary information by meal photos; this function was developed for DialBetics and is described in detail elsewhere [13]. To record dietary information, users can enter photos of a meal or choose from a menu list of 2913 items based on Eat Smart, a database provided by Eat Smart, Inc. When dietary information is entered from the menu list, the app automatically calculates each meal's intake of calories, protein, fat, carbohydrate, dietary fiber, cholesterol, and salt, which are displayed as a graph. The recommended intake for 1 meal—calculated as one-third of the recommended daily intake—is also displayed for comparison (Figure 1). As it has been reported that self-monitoring is crucial in weight loss [14] and improved glycemic control [15], the GlucoNote app was designed to provide users with visual feedback of the parameters, which helps in self-monitoring of body weight,

blood sugar level, physical activity represented by step counts, and diet. Physicians, including a diabetologist, nurses, and dietitians were involved in the development of GlucoNote.

Eligibility Criteria for Participants

Before release, the study was approved by the Institutional Review Board of the University of Tokyo. All iPhone users aged ≥20 years in Japan with type 2 diabetes or prediabetes were eligible for the study after they provided electronic consent. They could download GlucoNote for free.

Participants

The GlucoNote app was made available on March 14, 2016, (but only in Japan) through the Apple App Store [16]. The GlucoNote release was announced in the homepage of the department and a press release was issued [17] (Multimedia Appendix 1). After downloading and opening GlucoNote, participants were questioned for eligibility, that is, “Are you 20 years old or older?”; “Have you been diagnosed as type 2 diabetes or prediabetes?”; “Are you able to understand and follow the consent forms?”; and “Are you living in Japan?”. Those who met the eligibility criteria proceeded to the informed consent screens. Before providing consent, participants had to read about the risks and benefits of participating in the study and their right to withdraw from it; they could withdraw at any time without giving reasons. Informed consent was by digital signature. Participants who enrolled between March 14, 2016, and March 13, 2017, were analyzed (Figure 2).

Participants were encouraged to fill in the following profile information: sex, body height, body weight, wake-up time, bedtime, smoking habits, age at diagnosis, presence of retinopathy, presence of neuropathy, and regular visit to a dentist. Target daily calorie intake could be calculated based on height and activity level.

Participants also filled in the results of medical examinations: date of examination, height, weight, waist circumference, blood pressure, blood sugar, hemoglobin A_{1c} (HbA_{1c}), total cholesterol, high-density lipoprotein cholesterol, triglycerides, serum creatinine, aspartate transaminase, alanine transaminase, gamma-glutamyl transferase, urine protein, urine sugar, and urine albumin-creatinine ratio.

As shown in Figure 2, the observation duration was defined as the time from the day of study enrollment to May 13, 2017, 2 months after the end of enrollment. Retention duration was defined as the time from the day of study enrollment to the last day during the study period on which the user recorded at least 1 of 4 items: weight, blood sugar, blood pressure, or diet. Steps are different from other data—body weight, blood sugar, blood pressure, or dietary information—that require user input each time; they were automatically counted by the iPhones, and these data of steps recorded in the past 30 days were sent to the server at one time. Therefore, steps were discounted from analyses of retention rate.

Figure 1. Sample view of GlucoNote screen.

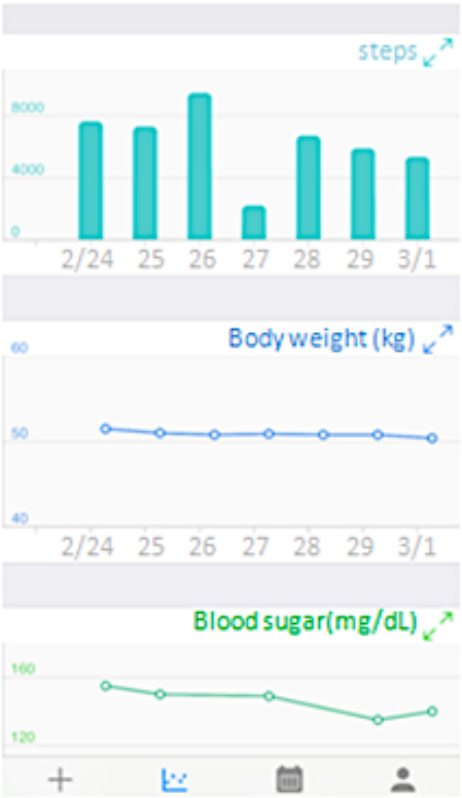
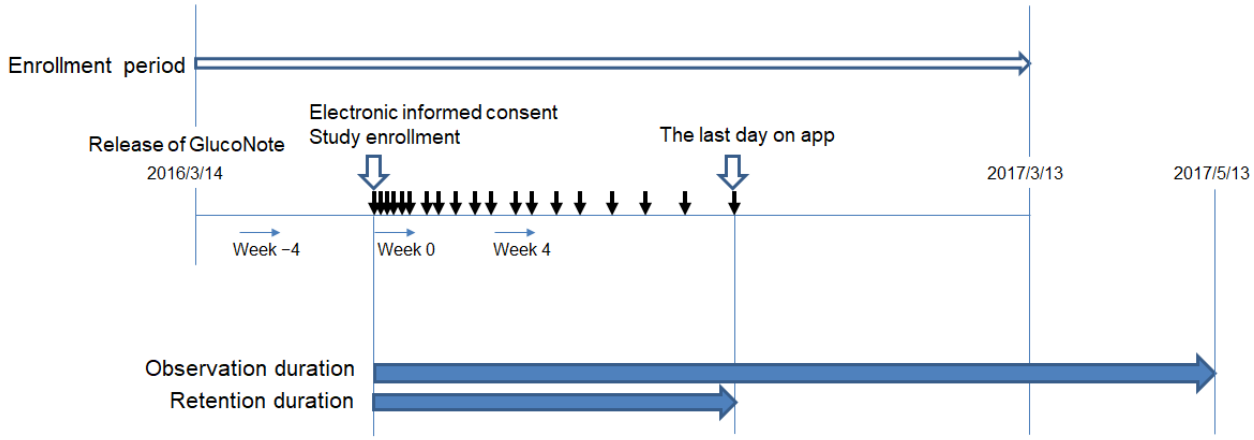


Figure 2. Definition of terms used in this study.



Data Analysis

The number of first-time app downloads was obtained from App Store Connect (Apple Inc), excluding app updates, downloads from the same Apple identification onto other devices, and redownloads [18].

For the participants' demographics in Table 1, body mass index (BMI) was calculated from their first recorded body weight.

To compare the data between weeks, each participant's mean values per week were calculated for weight, blood pressure, blood sugar, energy intake, or number of steps when the participant recorded those at least once during that week. When the user did not record that parameter for the entire week, the mean value of the corresponding week was treated as a missing value. When the user recorded the same parameter twice or more times on the same day, the data recorded first that day was used to calculate the mean value per week. The mean value recorded in the first 7 days starting from the day of study enrollment was defined as the value at week 0 (Figure 2). For the number of steps, the mean value recorded in the 7 days starting from 4 weeks before the day of enrollment was defined as the value at week −4 (Figure 2).

Daily energy intake was calculated by summing the calories of each meal when at least 3 meals were recorded per day. When only 1 or 2 meals were recorded, the daily energy intake for that day was treated as a missing value because one cannot tell whether participants did not record the meals or did not have the meals.

Retention rates were calculated by the Kaplan-Meier method [19], and the difference in retention rates between the 2 groups was evaluated by the log-rank test. The participants who continued to use the app to the end of the observation period were treated as censored cases and shown as vertical tick marks on the Kaplan-Meier curve. To compare characteristics between robust users and nonrobust users, the Fisher exact test was used for categorical variables and the Mann-Whitney *U* test was performed for continuous variables. A paired *t* test compared the data between week 0 and week 4 for parameters with a normal distribution, and a Wilcoxon signed-rank test was used for those with a non-normal distribution. Statistical analyses used R and Easy R [20].

Table 1. Demographics of participants in this study (N=522).

Characteristics	Statistics
Sex, n (%)	
Men	417 (79.9)
Women	101 (19.3)
Unanswered	4 (0.8)
Smoking status, n (%)	
Current smoker	32 (6.1)
Ex-smoker	73 (14.0)
Never smoked	95 (18.2)
Unanswered	322 (61.7)
Age at diagnosis (years), n (%)	
<40	45 (8.6)
40-59	111 (21.3)
≥60	10 (1.9)
Do not know	12 (2.3)
Unanswered	344 (65.9)
Hemoglobin A_{1c} (%), median (IQR[a]), (n=41)	6.3 (5.9-7.1)
Height (cm), median (IQR)	
Total (n=489)	169 (164-174)
Men (n=394)	171 (167-175)
Women (n=91)	159 (156-163)
Body weight (kg), median (IQR)	
Total (n=274)	70.6 (62.0-81.5)
Men (n=222)	73.5 (63.8-82.2)
Women (n=48)	61.2 (51.8-74.7)
Body mass index (kg/m^2), median (IQR)	
Total (n=270)	24.53 (21.99-27.82)
Men (n=218)	24.62 (22.12-27.45)
Women (n=48)	23.41 (20.36-28.95)

[a]IQR: interquartile range.

Results

Demographics of Participants

After GlucoNote was released on March 14, 2016, it was downloaded 1703 times during the first year, with 581 users consenting to take part in the study. The daily number of downloads during the year is shown in Multimedia Appendix 1. Of these, 59 users withdrew and the remaining 522 were analyzed (Figure 3). Their demographics are shown in Table 1. Males accounted for 79.9% (417/522) of the users, their median body weight was 70.6 (interquartile range [IQR] 62.0-81.5) kg (n=274), and BMI was 24.53 (IQR 21.99-27.82) kg/m^2 (n=270); 200 users answered the question about their smoking habit, 32 (16.0%; 32/200) of them being current smokers; 178 users answered the question about age at diagnosis, 111 (62.4%; 111/178) were diagnosed between the ages of 40 and 59 years; and 41 users recorded their HbA_{1c} level, the median value was 6.3% (IQR 5.9-7.1; Table 1).

A total of 467 participants recorded at least 1 of 5 items: body weight, blood sugar, blood pressure, diet, or steps (Figure 3). The number of participants who recorded body weight, blood sugar, blood pressure, dietary information, or step counts was 274 (58.7%; 274/467), 172 (36.8%; 172/467), 169 (36.2%; 169/467), 275 (58.9%; 275/467), or 428 (91.6%; 428/467), respectively (Figure 4). The measured data of participants at week 0 are shown in Table 2.

Figure 3. Study cohort description. BW: body weight, BP: blood pressure, BS: blood sugar.

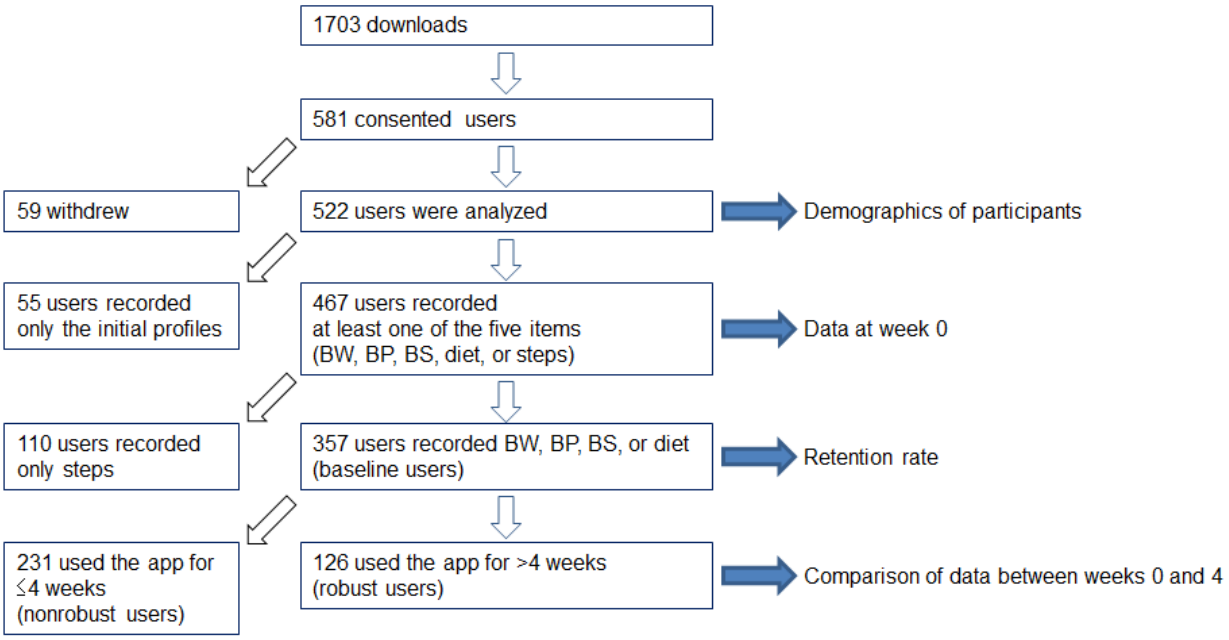


Figure 4. The proportion of users who recorded their data (n=467).

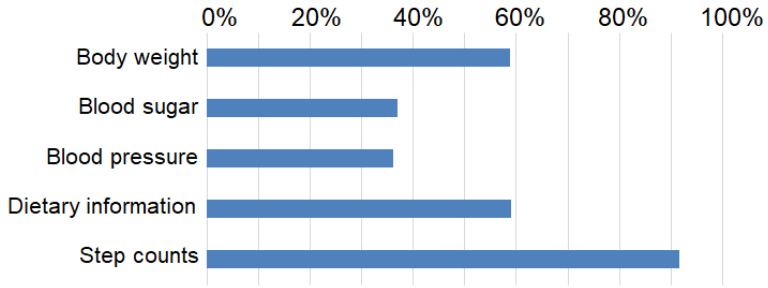


Table 2. Recorded data of participants at week 0.

Variable	Participants, n	Median (interquartile range)
Body weight (kg)	244	70.3 (61.4-81.8)
Body mass index (kg/m^2)	240	24.43 (21.91-27.83)
Fasting blood sugar (mg/dL)	124	118.5 (100.0-141.9)
Postprandial plasma glucose (mg/dL)	68	143.6 (121.8-177.9)
Systolic blood pressure (mmHg)	141	125.0 (116.8-133.0)
Diastolic blood pressure (mmHg)	141	79.4 (71.0-83.7)
Daily energy intake (kcal)	129	1510 (1003-1722)
Daily number of steps	409	5347 (3210-7716)

Retention Rate

Retention rates were analyzed for 357 participants who recorded at least 1 of 4 items: weight, blood sugar, blood pressure, or diet (*baseline users*; Figure 3). With the median observation duration of 382 days (range 63-426 days, IQR 275-423 days), the median retention duration in *baseline users* was 8 days (IQR 1-63 days; Table 3).

Retention rates for 2 days and 4, 8, and 12 weeks were 0.627 (95% CI 0.575-0.675), 0.353 (0.304-0.403), 0.272 (0.227-0.319), and 0.220 (0.179-0.265), respectively (Figure 5). Male

participants had a higher retention rate than female participants (*P*=.01; Figure 5). No significant differences were found between the groups with age at diagnosis <40 years and ≥40 years (*P*=.13) or the groups with BMI <25 kg/m^2 and BMI ≥25 kg/m^2 (*P*=.62; Figure 5).

Of the 357 participants, 126 used GlucoNote longer than 4 weeks (29 days or longer) and were analyzed as *robust users*. The median observation duration and retention duration for robust users were 375 (IQR 256-422) and 124 (IQR 60-275) days, respectively (Table 3).

Characteristics of robust users (n=126) and nonrobust users (n=231) were compared. Consistent with the result above, men were more likely to be robust users than women (*P*=.02); no difference was observed between the groups when age at diagnosis was <40 years and ≥40 years (*P*=.11; Table 3). Moreover, no differences in smoking status (*P*=.45), HbA_{1c} (*P*=.34), or body weight (*P*=.59) were observed. Those who answered the questions about smoking status and age at diagnosis were more likely to be robust users than were those who did not answer those questions (*P*<.001).

Table 3. Comparison of demographics of robust versus nonrobust users.

Variable	Total (baseline users; N=357)	Robust users (n=126)	Nonrobust users (n=231)	*P* value[a]
Observation duration (days), median (IQR[b])	382 (275-423)	375 (256-422)	382 (279-423)	.38
Retention duration (days), median (IQR)	8 (1-63)	124 (60-275)	2 (1-7)	<.001
Sex, n (%)				.02
Men	277 (77.6)	106 (84.1)	171 (74.0)	
Women	76 (21.3)	18 (14.3)	58 (25.1)	
Smoking status, n (%)				.45
Current smoker	23 (6.4)	9 (7.1)	14 (6.1)	
Ex-smoker	64 (17.9)	35 (27.8)	29 (12.6)	
Never smoked	80 (22.4)	42 (33.3)	38 (16.5)	
Age at diagnosis (years), n (%)				.11
<40	33 (9.2)	11 (8.7)	22 (9.5)	
≥40	104 (29.1)	53 (42.1)	51 (22.1)	
Hemoglobin A_{1c} (%), median (IQR), (n=40)	6.4 (5.9-7.2)	6.3 (6.0-6.8)[c]	6.7 (5.9-9.9)[d]	.34
Height (cm), median (IQR), (n=343)	169 (164-174)	170 (165-174)	169 (163-174)	.46
Body weight (kg), median (IQR), (n=274)	70.6 (62.0-81.5)	70.8 (62.0-79.9)	70.6 (62.0-82.5)	.59
Body mass index, median (IQR), (n=270)	24.53 (21.99-27.82)	24.49 (21.97-27.35)	24.57 (22.03-28.34)	.64
Smoking status question, n (%)				<.001
Answered	167 (46.8)	86 (68.3)	81 (35.1)	
Unanswered	190 (53.2)	40 (31.7)	150 (64.9)	
Age at diagnosis question , n (%)				<.001
Answered	146 (40.9)	72 (57.1)	74 (32.0)	
Unanswered	211 (59.1)	54 (42.9)	157 (68.0)	

[a]*P* value comparing robust users and nonrobust users using the Fisher exact test for categorical variables and Mann-Whitney *U* test for continuous variables.

[b]IQR: interquartile range.

[c]n=27.

[d]n=13.

Figure 5. Retention rate of GlucoNote. Retention duration was defined as duration from the day of study enrollment to the last day on the app. (A) Retention rate of overall participants and according to (B) sex (men vs women, *P*=.01), (C) age at diagnosis (<40 vs ≥40 years old [y.o], *P*=.13), and (D) body mass index (BMI <25 vs ≥25kg/m^2, *P*=.62).

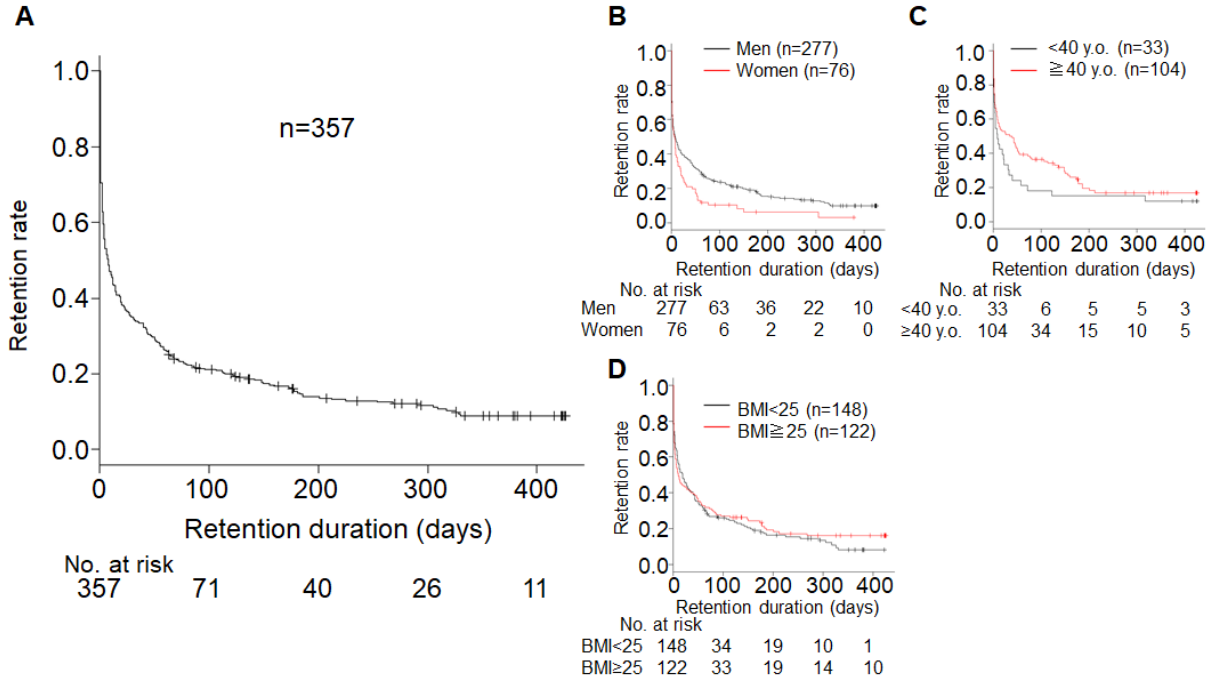


Table 4. Comparison of measured data at week 0 between robust and nonrobust users.

Variable	Robust users (n=126)		Nonrobust users (n=231)		*P* value[a]
	n	median (IQR[b])	n	median (IQR)	
Body weight (kg)	97	70.5 (60.0-79.4)	147	70.2 (61.8-83.0)	.26
Body mass index (kg/m^2)	97	24.48 (21.72-27.02)	143	24.28 (22.00-28.44)	.38
Fasting blood sugar (mg/dL)	55	122.0 (106.0-140.9)	69	112.7 (93.3-144.5)	.08
Postprandial plasma glucose (mg/dL)	31	144.2 (136.0-184.1)	37	142.5 (118.0-160.0)	.24
Systolic blood pressure (mmHg)	68	125.3 (116.4-133.0)	73	125.0 (117.4-134.0)	.85
Diastolic blood pressure (mmHg)	68	79.8 (72.4-84.4)	73	78.0 (70.7-81.5)	.12
Daily energy intake (kcal)	70	1595 (1198-1788)	59	1451 (769-1657)	.04
Daily number of steps, week 0	119	6108 (3797-9227)	194	5171 (2885-7258)	.001
Daily number of steps, week −4[c]	107	5876 (3714-7975)	185	5594 (3308-7597)	.37

[a]*P* value comparing robust users and nonrobust users using the Mann-Whitney *U* test.

[b]IQR: interquartile range.

[c]For the number of steps, the values at week −4 were also compared between 2 groups.

Data at week 0 were compared between robust users and nonrobust users (Table 4). The daily number of steps at week 0 was higher among robust users than nonrobust users (median 6108 vs 5171; *P*=.001), whereas step counts at week −4 were comparable between the 2 groups (median 5876 vs 5594; *P*=.37). In addition, daily energy intake at week 0 was higher among robust users than nonrobust users (median 1595 kcal vs 1451 kcal; *P*=.04; Table 4). No significant differences in body weight, blood sugar, or blood pressure were observed.

Comparison of Data Between Weeks 0 and 4

The data for robust users were compared between weeks 0 and 4 (Figure 2). Strikingly, body weight significantly decreased from week 0 to 4 (mean 71.3 [SD 14.1] kg to mean 70.8 [SD 13.9] kg; *P*=.002; Table 5). The mean difference in body weight between week 0 and week 4 was −0.5 kg (SD 1.2), corresponding to a mean decrease rate of 0.6% (SD 1.6). No changes in blood sugar level, blood pressure level, daily energy intake, or daily number of steps were observed between weeks 0 and 4 (Table 5).

Table 5. Comparison of robust users' data between weeks 0 and 4.

Variable	Week 0	Week 4	*P* value[a]	Users, n
Body weight (kg), mean (SD)	71.3 (14.1)	70.8 (13.9)	.002	67
Body mass index (kg/m^2), median (IQR[b])	25.18 (21.96-27.64)	25.10 (21.76-27.10)	<.001	67
Fasting blood sugar (mg/dL), median (IQR)	122.0 (108.2-142.7)	123.0 (111.0-150.4)	>0.99	35
Postprandial plasma glucose (mg/dL), mean (SD)	171.3 (30.2)	145.8 (58.0)	.26	11
Systolic blood pressure (mmHg), median (IQR)	122.5 (113.6-131.4)	121.0 (112.9-127.9)	.053	47
Diastolic blood pressure (mmHg), mean (SD)	77.2 (8.0)	77.9 (8.0)	.55	47
Daily energy intake (kcal), median (IQR)	1595 (1273-1788)	1762 (1408-2012)	.06	46
Daily number of steps, median (IQR)	6745 (4141-9883)	6376 (3468-9080)	.11	104

[a]*P* value comparing data between week 0 and 4 using a paired *t* test or Wilcoxon signed-rank test.

[b]IQR: interquartile range.

After the observation period, a questionnaire was sent to the participants asking whether they found GlucoNote useful for health management. Only 22 out of 522 participants (4.2%; 22/522) replied: 4 (18%; 4/22) people found the app *very useful*, 6 (27%; 6/22) *useful*, and 12 (55%; 12/22) *not very useful*; none of the participants pronounced it *not at all useful*.

Discussion

In this study, we evaluated a novel smartphone app, GlucoNote, which supports self-management in patients with type 2 diabetes and prediabetes and is established on the mobile health (mHealth) platform, ResearchKit. The study was conducted entirely remotely—including recruitment of participants and obtaining informed consent—without intervention needed by medical staff.

User Demographics

In our study, nearly 80% (79.9%; 417/522) of the participants were men (Table 1). This is much more biased than expected, even taking into consideration that men, in Japan, are more likely to have diabetes or prediabetes than women (28.5% vs 21.4%, respectively) [21]. There have also been reports that men use information and communication technology (ICT) more frequently, have less ICT anxiety, and have a more positive attitude toward ICT self-efficacy than women [22-25]. Consistent with those reports, our recent survey of patients with chronic conditions suggested that male patients were more likely to express willingness to use a personal health record than female patients (paper in preparation). It has further been reported that women are more likely to perceive diabetes as a stigma than men [26], and it may be that women were reluctant to use an app that was plainly meant for diabetes patients. Efforts to remove such barriers for women may contribute to broadening usage. The bias was not because of the difference in ownership of smartphones: a 2018 survey in the Tokyo area reported that 77.9% of men and 80.9% of women aged between 15 and 69 years owned a smartphone [27].

To get enough participants, we had to forego gathering information such as age because the more the input required before setting up GlucoNote, the fewer comply—meaning fewer study participants. That is why complete user demographics are unavailable. Moreover, many values are missing from the users' profiles, a problem shared by similar studies [10]. For example, only 38.3% (200/522) of the participants answered the question about smoking (Table 1). Not surprisingly, those who answered questions such as smoking status or age at diagnosis were likely to use GlucoNote longer (Table 3). Improvement in usability and security may be helpful to entice participants to enter more information. In our recent survey of patients with lifestyle-related diseases, *time and effort needed* and *concern over security* were identified as the main barriers to using such personal health records as health care apps (paper in preparation). For example, instead of relying on users to manually input the data, developing functions such as automatic recognition from a photo image of medical examination results may be useful—as may be reassuring users by improved security.

We do not have information about the use of other apps by GlucoNote users or people who visited the product page of GlucoNote, although that would have been helpful in identifying potential users.

Retention Rate

A rapid decline in retention rates has been reported in other studies using ResearchKit [12,28]. Although the ResearchKit platform offers the advantage of completely remote recruitment and enrollment, lack of human communication may mean less motivation for participants to continue compared with studies conducted face to face. In this study, the 2-day retention rate was only 0.627 (95% CI 0.575-0.675; Figure 5), meaning that more than 1 in 3 participants used the app only for 1 or 2 days. The median retention duration was as short as 8 days (Table 3). However, long-term use is essential for the app to affect users. Clearly, additional efforts to improve retention rates are necessary. Our previous study using DialBetics—conducted at a university hospital—showed a much higher retention rate of over 70% at 3 months, partly because the research team nurse contacted the participants and encouraged them when they missed measurements for 2 weeks [5]. Such intervention by medical staff certainly helps improve retention, thought it obviously means a higher cost. Alternatively, some incentives built into the app—such as reward points for long-term users—may be helpful.

The characteristics associated with the retention time of mHealth have not been well studied. In a study of asthma patients who used ResearchKit, being female and older correlated with longer retention [10]. By contrast, in our study, men had a longer retention rate than women (Figure 5 and Table 3). Age at study enrollment was not collected, but age at diagnosis did not correlate with retention rate (Figure 5 and Table 3). As mHealth is relatively new, factors predicting willingness to use it—or good adherence to mHealth—have not been well understood. To deliver mHealth effectively, it is important to identify the characteristics that mark suitable candidates for the app.

Interestingly, robust users had significantly higher step counts at week 0 than nonrobust users (median 6108 vs 5171 steps; *P*=.001), whereas step counts were comparable before using the app (median 5876 vs 5594 steps; *P*=.37; Table 4), suggesting that using the app may have prompted an increase in step count by robust users but not by nonrobust users. In contrast to steps that were counted automatically, daily energy intake data may be incomplete because users may not have recorded all their food and drinks. The higher daily energy intake among robust users at week 0 (Table 4) might reflect a likelihood that these users input meal information more completely than nonrobust users, in turn, a reflection of robust user attitudes toward using the app in contrast with those of nonrobust users.

Body Weight Decrease Among Robust Users

We observed a significant decrease in body weight (*P*=.002) after 4 weeks, with a mean decrease rate of 0.6% (SD 1.6; Table 5). Of 67 users with weight data for weeks 0 and 4, 47 (70%; 47/67) showed a decrease at week 4, suggesting that using the app might have prompted behavior modification leading to weight decrease.

It was previously reported that weight loss was the dominant predictor of reduced diabetes risk in patients with impaired glucose tolerance, with every kilogram of weight loss resulting in 16% reduction in risk [29]. A nationwide Japanese intervention program showed that for people with metabolic syndrome, a 1% to 3% weight loss after a 6-month lifestyle modification program resulted in a significant decrease in HbA_{1c} [30]. Although body weight loss in this study (mean −0.6% in 4 weeks) is only marginal and may be of little clinical significance, the long-term use of GlucoNote by improving the retention rate may lead to a reduced diabetes risk in patients with prediabetes. More importantly, it would be of interest to follow up to determine if the weight loss is maintained over time.

It is difficult to explain the exact reasons for weight loss only from the data recorded; no change in steps per day or energy intake was observed between weeks 0 and 4. However, one can speculate that some behavior changes undetectable by an app were prompted—such as increased physical activity other than walking. In addition, it is possible that the app failed to detect changes in diet because users may not have recorded all their food and drinks.

Study Limitations

There are several limitations of this study. First, as discussed above, its major limitation is the low retention rate. Second, as noted above, the participants' complete demographics were unavailable, and data collection was incomplete. Consequently, many analyses were based on different sample sizes, a common defect shared with similar studies [10-12]. Correcting this will require effort to improve response rates and collect more complete data. Moreover, only robust users were available for the analysis, increasing the risk of selecting bias. Third, the study design was free of controls and eligibility, including diagnosis of diabetes and prediabetes, and the measured data were entirely based on self-reports and had not been validated by medical professionals—a common problem shared by similar studies. Fourth, the overall number of downloads was lower than expected, with only 1703 in the first year after the release. A more effective way to inform people of the app is desirable. In addition, to widen its targets to include people with a metabolic syndrome, including overweight or even healthy people may be considered. Finally, to evaluate the effects of the app, a randomized controlled trial must be performed.

Conclusions

We developed and released GlucoNote, a novel app that uses ResearchKit to support self-management in patients with type 2 diabetes and prediabetes. This afforded a valuable opportunity to evaluate usage patterns. Analyses of the participants who enrolled in the study within 1 year of the release revealed the potential advantages and challenges of GlucoNote. Future tasks include improving retention rates and evaluating its effects.

Acknowledgments

This study was conducted at the Department of Ubiquitous Health Informatics, which is engaged in a cooperative program between the University of Tokyo and NTT DOCOMO, funder of the study. The funder had no role in the design or conduct of the study; in the collection, management, analysis, or interpretation of the data; in the preparation, review, or approval of the manuscript; or in the decision to submit the manuscript for publication.

Conflicts of Interest

KW is a member of the Department of Ubiquitous Health Informatics, the University of Tokyo. SY and TK are members of the Department of Prevention of Diabetes and Lifestyle-related Diseases, which is in a cooperative program between the University of Tokyo and Asahi Mutual Life Insurance Company; SY was a member of the Department of Ubiquitous Health Informatics when the study was conducted.

Multimedia Appendix 1

The number of daily downloads during the first year after GlucoNote release. The number of daily downloads (n=1703) and study enrollment (n=522) are shown.

[PNG File, 24KB - mhealth_v7i4e13204_app1.png]

References

1. Bonoto BC, de Araújo VE, Godói IP, de Lemos LL, Godman B, Bennie M, et al. Efficacy of mobile apps to support the care of patients with diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2017 Mar 01;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]
2. Wu Y, Yao X, Vespasiani G, Nicolucci A, Dong Y, Kwong J, et al. Mobile app-based interventions to support diabetes self-management: a systematic review of randomized controlled trials to identify functions associated with glycemic efficacy. JMIR Mhealth Uhealth 2017 Mar 14;5(3):e35 [FREE Full text] [doi: 10.2196/mhealth.6522] [Medline: 28292740]
3. Khokhar B, Jones J, Ronksley PE, Armstrong MJ, Caird J, Rabi D. Effectiveness of mobile electronic devices in weight loss among overweight and obese populations: a systematic review and meta-analysis. BMC Obes 2014;1:22 [FREE Full text] [doi: 10.1186/s40608-014-0022-4] [Medline: 26217509]
4. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
5. Waki K, Fujita H, Uchimura Y, Omae K, Aramaki E, Kato S, et al. DialBetics: a novel smartphone-based self-management support system for type 2 diabetes patients. J Diabetes Sci Technol 2014 Mar 13;8(2):209-215 [FREE Full text] [doi: 10.1177/1932296814526495] [Medline: 24876569]
6. Rivera J, McPherson A, Hamilton J, Birken C, Coons M, Iyer S, et al. Mobile apps for weight management: a scoping review. JMIR Mhealth Uhealth 2016 Jul 26;4(3):e87 [FREE Full text] [doi: 10.2196/mhealth.5115] [Medline: 27460502]
7. Apple Inc. ResearchKit URL: http://researchkit.org/ [accessed 2018-12-03] [WebCite Cache ID 74P2myxLW]
8. Jardine J, Fisher J, Carrick B. Apple's ResearchKit: smart data collection for the smartphone era? J R Soc Med 2015 Aug;108(8):294-296. [doi: 10.1177/0141076815600673] [Medline: 26268915]
9. Ritter S. Apple's research kit development framework for iPhone apps enables innovative approaches to medical research data collection. J Clin Trials 2015;5(2). [doi: 10.4172/2167-0870.1000e120]
10. Chan YY, Wang P, Rogers L, Tignor N, Zweig M, Hershman SG, et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. Nat Biotechnol 2017 Mar 13;35(4):354-362 [FREE Full text] [doi: 10.1038/nbt.3826] [Medline: 28288104]
11. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar 03;3:160011 [FREE Full text] [doi: 10.1038/sdata.2016.11] [Medline: 26938265]
12. Crouthamel M, Quattrocchi E, Watts S, Wang S, Berry P, Garcia-Gancedo L, et al. Using a ResearchKit smartphone app to collect rheumatoid arthritis symptoms from real-world participants: feasibility study. JMIR Mhealth Uhealth 2018 Sep 13;6(9):e177. [doi: 10.2196/mhealth.9656] [Medline: 30213779]
13. Waki K, Aizawa K, Kato S, Fujita H, Lee H, Kobayashi H, et al. DialBetics with a multimedia food recording tool, FoodLog: smartphone-based self-management for type 2 diabetes. J Diabetes Sci Technol 2015 May;9(3):534-540 [FREE Full text] [doi: 10.1177/1932296815579690] [Medline: 25883164]
14. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
15. McAndrew LM, Horowitz CR, Lancaster KJ, Quigley KS, Pogach LM, Mora PA, et al. Association between self-monitoring of blood glucose and diet among minority patients with diabetes. J Diabetes 2011 Jun;3(2):147-152 [FREE Full text] [doi: 10.1111/j.1753-0407.2011.00114.x] [Medline: 21599868]
16. Apple Inc. GlucoNote URL: https://itunes.apple.com/jp/app/gluconote/id1083049623?mt=8 [accessed 2018-12-04] [WebCite Cache ID 74P3BuUrX]
17. The University of Tokyo. GlucoNote URL: http://uhi.umin.jp/gluconote/index.html [accessed 2019-02-22] [WebCite Cache ID 76MeCkgPe]
18. Apple Inc. App Store Connect URL: https://help.apple.com/app-store-connect/ [accessed 2019-02-14] [WebCite Cache ID 76AxKiDZE]
19. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958 Jun;53(282):457-481. [doi: 10.2307/2281868]
20. Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. Bone Marrow Transplant 2013 Mar;48(3):452-458 [FREE Full text] [doi: 10.1038/bmt.2012.244] [Medline: 23208313]
21. Ministry of Health, Labour and Welfare. National Health and Nutrition Survey, 2016 URL: https://www.mhlw.go.jp/stf/houdou/0000177189.html [accessed 2019-03-11] [WebCite Cache ID 76mmVjtoy]
22. Broos A. Gender and information and communication technologies (ICT) anxiety: male self-assurance and female hesitation. Cyberpsychol Behav 2005 Feb;8(1):21-31. [doi: 10.1089/cpb.2005.8.21] [Medline: 15738690]

23. Colley A, Comber C. Age and gender differences in computer use and attitudes among secondary school students: what has changed? Educational Research 2003;45(2):155-165. [doi: 10.1080/0013188032000103235] [Medline: 26281194]
24. Durndell A, Haag Z. Computer self efficacy, computer anxiety, attitudes towards the internet and reported experience with the Internet, by gender, in an East European sample. Comput Hum Behav 2002 Sep;18(5):521-535. [doi: 10.1016/S0747-5632(02)00006-7]
25. Vekiri I, Chronaki A. Gender issues in technology use: perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. Comput Educ 2008 Nov;51(3):1392-1404. [doi: 10.1016/j.compedu.2008.01.003]
26. Liu NF, Brown AS, Younge MF, Guzman SJ, Close KL, Wood R. Stigma in people with type 1 or type 2 diabetes. Clin Diabetes 2017 Jan;35(1):27-34 [FREE Full text] [doi: 10.2337/cd16-0020] [Medline: 28144043]
27. Hakuhodo DY Media Partners Inc. Time Series Analysis Findings from the Annual Media Consumption Report 2018 URL: https://www.hakuhodody-media.co.jp/english/newsrelease/reports/20180528_7163.html [accessed 2019-03-05] [WebCite Cache ID 76dsESFdR]
28. Dorsey ER, Yvonne CY, 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]
29. Hamman RF, Wing RR, Edelstein SL, Lachin JM, Bray GA, Delahanty L, et al. Effect of weight loss with lifestyle intervention on risk of diabetes. Diabetes Care 2006 Sep;29(9):2102-2107 [FREE Full text] [doi: 10.2337/dc06-0560] [Medline: 16936160]
30. Muramoto A, Matsushita M, Kato A, Yamamoto N, Koike G, Nakamura M, et al. Three percent weight reduction is the minimum requirement to improve health hazards in obese and overweight people in Japan. Obes Res Clin Pract 2014;8(5):e466-e475. [doi: 10.1016/j.orcp.2013.10.003] [Medline: 25263836]

Abbreviations

BMI: body mass index
HbA $_{1c}$: hemoglobin A_{1c}
ICT: information and communication technology
IQR: interquartile range
mHealth: mobile health
PD: Parkinson disease

Edited by G Eysenbach; submitted 10.01.19; peer-reviewed by JB Park, M Lennon, J Brož, F Fatehi, S Li; comments to author 02.02.19; revised version received 11.03.19; accepted 30.03.19; published 24.04.19.

Please cite as:
Yamaguchi S, Waki K, Nannya Y, Nangaku M, Kadowaki T, Ohe K
Usage Patterns of GlucoNote, a Self-Management Smartphone App, Based on ResearchKit for Patients With Type 2 Diabetes and Prediabetes
JMIR Mhealth Uhealth 2019;7(4):e13204
URL: http://mhealth.jmir.org/2019/4/e13204/
doi:10.2196/13204
PMID:31017586

Original Paper

Assessing the Need for Mobile Health (mHealth) in Monitoring the Diabetic Lower Extremity

David Wallace[1,2*], MD; Julie Perry[3*], PhD; Janelle Yu[2], MD; Joshua Mehta[2], BSc; Paul Hunter[3], BSc; Karen Michelle Cross[1,3,4], MD, PhD

[1]Division of Plastic and Reconstructive Surgery, University of Toronto, Toronto, ON, Canada
[2]Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[3]Division of Plastic Surgery, St. Michael's Hospital, Toronto, ON, Canada
[4]Keenan Research Centre for Biomedical Science, Toronto, ON, Canada
[*]these authors contributed equally

Corresponding Author:
Karen Michelle Cross, MD, PhD
Division of Plastic Surgery
St. Michael's Hospital
30 Bond Street
Toronto, ON, M5B 1W8
Canada
Phone: 1 416 864 6060 ext 77074
Email: karen@drkarencross.ca

Abstract

Background: Complications of the diabetic lower extremity (such as diabetic foot ulcers, DFUs) occur when monitoring is infrequent, and often result in serious sequelae like amputation or even death.

Objective: To evaluate the potential application of mobile health (mHealth) to diabetic foot monitoring. We surveyed the self-management routines of a group of diabetic patients, as well as patient and clinician opinions on the use of mHealth in this context.

Methods: Patients with DFUs in Toronto, Ontario, Canada completed a 25-item questionnaire addressing their foot care practices, mobile phone use, and views on mHealth. Wound care clinicians across Canada were also surveyed using a 9-item questionnaire.

Results: Of the patients surveyed, 59/115 (51.3%) spend less than a minute checking their feet, and 17/115 (15%) of patients find it difficult to see their doctor or get to the hospital regularly. Mobile phone use was widespread in our patient cohort (93/115, 80.9%). Of mobile phone users, 68/93 (73.1%) would use a device on their mobile phone to help them check their feet. Of the clinicians who completed the questionnaire, only 7/202 (3.5%) were familiar with mHealth; however, 181/202 (92%) of clinicians expressed interest in using mHealth to monitor their patients between visits.

Conclusions: Patient education or motivation and clinician training were identified as the major barriers to mHealth use in the diabetic lower extremity, which may be a viable mechanism to improve DFU monitoring practices.

(JMIR Mhealth Uhealth 2019;7(4):e11879) doi:10.2196/11879

KEYWORDS

mHealth; diabetes; diabetic foot ulcers

Introduction

The burden of treatment in diabetes is high: patients must monitor their diet and blood glucose levels, take medication, refill prescriptions, travel to medical appointments, seek information, and keep records [1]. This “illness work” can become time- and identity-consuming, and can impose so much on a patient’s everyday life that treatment compliance rates drop and outcomes decline [1]. One common complication that can result from patient burn-out are diabetic foot ulcers (DFUs) due to a combination of vascular problems of the lower extremity and the lack of sensation that is common in diabetes (diabetic neuropathy). An ill-fitting sock or shoe can rub enough to cause a foot ulcer in a diabetic person, which may go unnoticed until it is large or infected. When caught early, DFUs are highly treatable; however, most ulcers are not treated until they become

more advanced, such that one third of ulcers never heal and result in foot or lower extremity amputation [2,3]. The 3-year overall mortality rate following lower extremity amputation due to DFUs is as high as 70% [4]—higher than cancers of the breast and colon. These devastating outcomes are preventable, but consistent monitoring of foot health and early reporting of problems is critical.

The concept of *minimally disruptive medicine* (MDM) recognizes the workload associated with a chronic illness, and aims to simplify, consolidate, and synchronize healthcare activities to help patients manage their conditions efficiently [5]. A growing area of research with application to MDM is the delivery of healthcare remotely via apps on mobile devices (called mobile health or mHealth). Mobile phones are ubiquitous in society, and several software platforms have been developed for the self-management of diabetes. Much of mHealth's application to diabetes has been focused on apps that track patient blood glucose levels at home. Fewer hyperglycemic events and lower average glucose levels were reported in patients using mobile health platforms than non-mobile health-using controls [6]. There is currently no mHealth tool for diabetic foot health monitoring, although there is evidence that monitoring via telemedicine is as effective as standard outpatient monitoring in regards to healing and amputation rate [7]. Of note, however, Rasmussen et al observed a significantly higher mortality rate in the telemedicine group in their study, with no easily ascribable cause. Therefore, while highly promising, adoption of mHealth monitoring practices should proceed with caution.

To determine if an MDM-based mHealth intervention is needed for the diabetic lower extremity, we sought to characterize the daily self-management routines of diabetics at St. Michael's Hospital (SMH) in Toronto, Ontario, Canada, and the attitudes towards mHealth among Canadian foot care clinicians. Our goal was to understand the barriers to implementing such a strategy from both sides of the patient-clinician continuum.

Methods

Patient Questionnaire

This was a descriptive study of patients with DFUs presenting to plastic surgery clinics at SMH in Toronto over a 6-month period in 2017. SMH is a large, tertiary, academic level-1 trauma centre located in the downtown core. This study was approved by the SMH Research Ethics Board (REB 17-023). The questionnaire was designed by clinicians who treat this patient population, but was not validated in any way.

Patients at SMH were approached by study investigators following clinic visits and informed about the study. Patients were given the opportunity to ask questions, and if they chose to participate, they signed informed consent.

The study investigators designed a questionnaire addressing multiple themes associated with patients with DFUs, including characteristics of patient health, mobile phone use, and views on mHealth (see Multimedia Appendix 1). The majority of questions were designed to have "yes" or "no" answers or asked participants to "rate your experience from 1 to 10" in order to facilitate a shorter clinic experience for the patient. Patients were not given the option to add comments. The questionnaire was comprised of 25 questions and was designed to take patients approximately 5 minutes to complete. All responses from the patients surveyed were inputted into an electronic database and coded anonymously using a unique patient identifier. All continuous data were reported using means, medians, modes, and ranges. All categorical data were reported using frequencies and proportions. All analysis was carried out using Microsoft Excel.

Healthcare Provider Questionnaire

A 9-question survey (see Multimedia Appendix 1) was designed in Survey Monkey and distributed via email to the membership of Wounds Canada following Research Ethics Board approval at SMH. Responses were collected for 4 weeks, after which the survey was terminated. No personal identifiers were collected to link respondents to their responses.

Results

Patient Survey

Of the 117 patients who were approached in plastic surgery and diabetes clinics at SMH, 115 agreed to be asked a series of qualitative questions describing their foot checking practices and comfort with mobile technology. The average age of participants was 54.8 years, and 60/115 (52.2%) of the participants were men (Table 1). Of the participants, 68/115 (59.1%) were Type 2 diabetics, 91/115 (79.1%) were insulin dependent, and 108/115 (93.9%) used a glucometer. The average BMI of patients was 28.2 kg/m^2. Of the participants, 100/115 (87.0%) were non-smokers, 91/115 (79.1%) stated that they visit their physician a few times a year, and most report that being in control of their own health is very important to them (mean rating of 8.3 out of 10, where 10 is the highest importance; Table 1).

When asked about their current foot checking practices, 89/115 (77.4%) of patients reported checking their feet regularly, although 59/115 (51.3%) reported spending less than a minute checking, and only 16/115 (13.9%) use a mirror to check the bottoms of their feet (Table 1). Most patients (103/115, 89.6%) reported being comfortable touching their toes (suggesting that they are flexible enough to bend and check their feet), but 83/115 (72.2%) of respondents reported wearing corrective lenses, which may affect their ability to see their feet clearly. Importantly, only 11/115 (9.6%) of participants reported a previous DFU, and only 3/115 (2.6%) of patients had a prior amputation involving their toe or leg (Table 1).

Table 1. Summary of patient demographic data and survey (n=115).

Patient demographic or survey response	Responses
Age, mean (range)	54.8 (18-84)
Gender, n (%)	
Male	60 (52.2)
Female	60 (52.2)
Unknown	5 (4.4)
Body mass index (kg/m^2), mean (SD)	28.17 (7.91)
Occupation, n (%)	
Administration or management	8 (7.0)
Education	7 (6.1)
Engineering or technology	6 (5.2)
Finance or business	21 (18.3)
Healthcare	4 (3.5)
Labor	4 (3.5)
Other	17 (14.8)
Retired	30 (26.1)
Student	3 (2.6)
Unemployed	13 (11.3)
No answer	2 (1.7)
Diabetes type, n (%)	
1	42 (36.5)
2	68 (59.1)
Not sure	6 (5.2)
Uses insulin, n (%)	
Yes	91 (79.1)
No	23 (20.0)
No answer	1 (0.8)
Smoking status, n (%)	
Smoker	12 (10.4)
Never smoker	100 (87.0)
No answer	3 (2.6)
Wears corrective lenses, n (%)	
Yes	83 (72.2)
No	32 (27.8)
Comfortable touching their toes, n (%)	
Yes	103 (89.6)
No	8 (7.0)
No answer	4 (3.5)
Has had a prior diabetic foot ulcer, n (%)	
Yes	11 (9.6)
No	101 (87.8)
No answer	3 (2.6)
Has had a toe or leg amputation, n (%)	

Patient demographic or survey response	Responses
Yes	3 (2.6)
No	112 (97.4)
No answer	0 (0)
Importance of being in control of their own health[a], mean (SD)	8.28 (2.37)
Has difficulty getting to the hospital or seeing their doctor, n (%)	
Yes	17 (14.8)
No	95 (82.6)
No answer	3 (2.6)
Method of transportation to the hospital, n (%)	
Ambulance	1 (0.9)
Car	25 (21.7)
Electric scooter	1 (0.9)
Family member or friend	3 (2.6)
Public transportation	47 (40.9)
Taxi	4 (3.5)
Walk	9 (7.8)
Other	2 (1.7)
Multiple methods	21 (18.3)
No answer	2 (1.7)
Length of time spent checking feet, n (%)	
Never	23 (20)
Less than 1 minute	59 (51.3)
More than 1 minute	30 (26.1)
No answer	3 (3)
Uses a mirror to check the bottom of their feet, n (%)	
Yes	16 (13.9)
No	99 (86)
Frequency of doctors visits about feet, n (%)	
Every week	1 (0.8)
Every month	5 (4.3)
A few times a year	61 (53)
Only when I am sick	32 (27.8)
Never	13 (11.3)
No answer	3 (2.6)
Uses Glucometer, n (%)	
Yes	108 (93.9)
No	7 (6.1)
Owns Cellphone, n (%)	
Yes	94 (81.7)
No	21 (18.2)
Would use mHealth to check their feet, n (%)	
Yes	86 (74.7)

Patient demographic or survey response	Responses
No	29 (25.2)

[a]Measured on a scale of 1 to 10, where 10 is considered very important.

Although most of the study respondents (95/115, 82.6%) visit their doctor a few times a year, 17/115 (14.8%) of patients reported that it was difficult to get to the hospital or to see their doctor. Interestingly, 43/115 (37.4%) respondents said that they were retired or unemployed. We found that the occupations of our survey respondents were primarily in "white collar" sectors—46/115 (40.1%) of respondents listed their career as one of "Finance or business," "Administration or management," "Education," "Engineering or technology" or "Healthcare"—and only 4/155 (3.5%) of respondents considered their occupation to be "General labor." Mobile phone ownership was widespread (93/115, 80.4%) as was glucometer usage (108/115, 94%). Of the patients surveyed, 68/115 (73.1%) would use a device on their phone to help them check their feet (Table 1).

Clinician Survey

Responses to the clinician survey were mostly from wound care nurses or nurse practitioners (149/202; 73.8%) and chiropodist or podiatrists (20/202, 10.0%) who have been in practice for more than 5 years (Table 2). Most clinicians see their patients regularly—121/202 (60.8%) see their patients weekly or monthly, but 62/202 (31.2%) of clinicians see their patients only when they have a problem (Table 2). This lack of routine care is perceived to be due primarily to patient barriers (41/202, 20.3%), provider barriers (22/202, 10.9%) or a combination of both (115/202, 57.0%), and many clinicians left survey comments identifying patient financial constraints and education about the importance of foot checking as barriers to more frequent care. The concept of mHealth was unfamiliar to the vast majority of wound care clinicians (145/202, 71.8%), and only 7/202 (3.5%) are currently using mHealth in their practices. However, 161/202 (81.7%) respondents thought that more frequent monitoring through mHealth could improve patient outcomes, and 181/202 (91.4%) would consider using an mHealth approach to monitor or supplement patient monitoring between visits (Table 2). The most frequently identified concerns about mHealth were the reliability of patient-generated data (98/202, 49.3%) and the reliability or accuracy of the technology itself (82/202, 41.2%). Several respondents left comments suggesting that their elderly patients would have trouble managing a new mobile phone–based technology, or have trouble affording such a device.

Table 2. Summary of responses by clinicians surveyed (N=202).

Demographic	Responses, n (%)
Clinician role	
Nurse or nurse practitioner	149 (73.8)
Chiropodist or podiatrist	20 (9.9)
Family physician	5 (2.5)
Internal medicine	2 (1.0)
Surgeon (plastic, orthopaedic, vascular)	0 (0)
Other	26 (12.9)
Percentage of practice is devoted to diabetic foot ulcers?	
Less than 10%	43 (21.3)
10-50%	111 (55.0)
Greater than 50%	48 (23.8)
Years in practice	
Less than 1 year	5 (2.5)
1-5 years	33 (16.3)
5+ years	164 (81.2)
Familiarity with mHealth	
Using mHealth within practice	7 (3.5)
Familiar, but not using mHealth	50 (24.8)
Not familiar with mHealth	145 (71.8)
Likelihood to use mHealth in practice	
Would use mHealth to supplement care/monitor between visits	88 (43.6)
Would not change current practice, but would use mHealth as a supplement	93 (46)
Have concerns about using mHealth	17 (8.4)
Barriers to seeing patients at ideal frequency	
Patient barriers	41 (20.3)
Provider barriers	22 (10.9)
Combination of barriers	115 (56.9)
Other	24 (11.9)

Discussion

Diabetics carry a heavy burden of illness that requires significant healthcare "work," including diet and lifestyle planning and tracking, medication adherence, doctor's appointments, and self-monitoring. In a large cohort study recently completed in Alberta, only 14% of respondents reported checking their feet 6 days a week or more, and only 41% and 34% had their feet checked regularly by a clinician for ulcers or sensory loss, respectively [8]. Given that frequent monitoring (and resulting early detection) of foot ulcers is critical to their effective treatment, we are actively seeking ways to increase the frequency of monitoring patients while balancing their quality of life. The present study sought to understand the practices and health "work" done by a group of patients presenting to a plastic surgery wound clinic at SMH in Toronto for foot monitoring and treatment of DFUs, with the goal of developing an effective mHealth strategy for the diabetic lower extremity following the principles of MDM.

While the average lifetime incidence of DFUs among Canadian diabetics is 15% [9], only 11/115 (9.6%) of our study participants reported a prior DFU and only 3/115 (2.6%) had a prior amputation. Our study participants reported seeing their doctor regularly for both their feet and other concerns, and also reported checking their feet regularly, which has been shown to result in lower rates of DFUs and amputation. In fact, our sample population ranks being in control of their own health as very important, which is in itself a predictor of positive outcomes. Our results may also be reflective of our urban location: Al Sayah et al (2015) found that predictors of clinical monitoring included residing in an urban locations, and there are clear regional differences in amputation rates in the United States, with underserviced and rural communities having a +51.3% higher odds of major amputation, +14.9% higher odds

of minor amputation, and +41.4% higher odds of inpatient death (P<0.05) than their urban diabetic counterparts [10]. We are clearly missing a vulnerable population who are at high risk for DFUs (and amputations) who cannot get to a clinic in person. It is time to start thinking outside the box to reach them.

Diabetics are a technology-oriented patient population, as 108/115 (94%) of our study population reported using a glucometer to monitor their disease. The Canadian population as a whole is becoming increasingly oriented towards mobile phones: 85.6% of the population owned a mobile phone in 2016, up from 62.9% in 2006 [11]. Of the patients who own a mobile phone, 84/115 (73%) responded that they would be interested in using an app on their mobile phone to help them check their feet. Currently available mHealth tools for diabetes management have been shown to significantly improve the frequency of blood glucose monitoring resulting in fewer hyperglycemic events than control groups [6], suggesting that mHealth can positively affect patient-monitoring frequency. Specific tools for the diabetic lower extremity are lacking, although a search of the Apple and Google Play stores returned several apps that are directed at providing education for diabetic foot screening. Unfortunately, although these apps are available, they are grossly underutilized. Despite estimates that 422 million people worldwide have diabetes (and should be checking their feet), the most highly downloaded app from Google Play had only 500-1000 installs. Part of the reason why so few DFU apps exist is that there are limited objective outcome measures available for the diabetic lower extremity. Patients may also fail to appreciate the health implications of a DFU, and neglect checking their feet in favor of checking their blood glucose regularly. Educational campaigns emphasizing the importance of foot health and efforts to develop strong science for the diagnosis of early stage DFUs are necessary components of any future mHealth strategies.

Our patient survey suggests mHealth may be useful from a patient's perspective, but we were also interested in the perspective of wound care clinicians. The results of our clinician survey suggest that approximately 1 in 3 patients are only being seen by a clinician once they have already developed a problem. Reasons cited by clinicians for this lack of care are not due to systemic barriers like wait times, but are largely a matter of patient education and engagement. Clinicians also cited financial barriers to mobile phone ownership and lack of comfort with technology among their patient populations, which are barriers that must be considered when developing future mHealth strategies. It could be argued, however, that the cost of ulcer prevention versus amputation should be considered from a public policy making perspective, and that the economics of mHealth are attractive in a publicly-funded healthcare system [12]. Baring the barriers stated, most clinicians would be willing to use an mHealth tool as part of their clinical monitoring.

Despite their willingness however, clinicians did express concern about the reliability of using patient-generated data and relying on pictures alone for wound care (eg, no information on parameters such as wound smell). Furthermore, clinical adoption of mHealth would require the development of fee structures for billing, and mechanisms to ensure patients could be called into the clinic quickly if their condition deteriorated. Regulating technology that is used in healthcare would also require substantial oversight at the national level and on a hospital-by-hospital basis, as there are already many companies vying for space in this potentially lucrative market, some of which have a better understanding of wound care than others. Although work remains to be done before mHealth is ready for widespread use in wound care, in our opinion, these challenges are not insurmountable.

To the authors' knowledge, this is the largest study evaluating the opinion of patients with DFU on mHealth, albeit in a single hospital population. We also surveyed both sides of the patient-care continuum; both patients and healthcare providers are willing to use mHealth for monitoring of the diabetic lower extremity. Weaknesses of our study include the fact that our patient population was drawn from a single plastic surgery wound clinic. It is unknown what proportion of diabetic foot patients ever present to clinics like this, and thus the external validity of the study may only be applicable to a small subset of patients.

In conclusion, we must find ways to increase foot monitoring frequency and effectiveness in diabetic patients. Using unconventional strategies like mHealth may be feasible but should incorporate educational campaigns to motivate patients and clinicians alike, and should move beyond simply taking a picture of a wound and instead build upon evidence-based outcome measures for foot health like tissue oxygenation, perfusion, and free-radical accumulation.

Acknowledgments

This work was funded by a grant from the Translational Innovation Fund at St. Michael's Hospital to KC.

Conflicts of Interest

KC has equity interest in a small start-up company developing mHealth solutions for the diabetic lower extremity.

Multimedia Appendix 1

Supplementary figures.

[PDF File (Adobe PDF File), 352KB - mhealth_v7i4e11879_app1.pdf]

References

1. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf M, Wei E. "You Get Reminded You're a Sick Person": Personal Data Tracking and Patients With Multiple Chronic Conditions. J Med Internet Res 2015 Aug 19;17(8):e202 [FREE Full text] [doi: 10.2196/jmir.4209] [Medline: 26290186]
2. Prompers L, Schaper N, Apelqvist J, Edmonds M, Jude E, Mauricio D, et al. Prediction of outcome in individuals with diabetic foot ulcers: focus on the differences between individuals with and without peripheral arterial disease. The EURODIALE Study. Diabetologia 2008 May;51(5):747-755 [FREE Full text] [doi: 10.1007/s00125-008-0940-0] [Medline: 18297261]
3. Smith-Strøm H, Iversen MM, Igland J, Østbye T, Graue M, Skeie S, et al. Severity and duration of diabetic foot ulcer (DFU) before seeking care as predictors of healing time: A retrospective cohort study. PLoS One 2017 May 12;12(5):e0177176 [FREE Full text] [doi: 10.1371/journal.pone.0177176] [Medline: 28498862]
4. Stern J, Wong CK, Yerovinkina M, Spindler SJ, See AS, Panjaki S, et al. A Meta-analysis of Long-term Mortality and Associated Risk Factors following Lower Extremity Amputation. Ann Vasc Surg 2017 Jul;42:322-327. [doi: 10.1016/j.avsg.2016.12.015] [Medline: 28389295]
5. Serrano V, Spencer-Bonilla G, Boehmer KR, Montori VM. Minimally Disruptive Medicine for Patients with Diabetes. Curr Diab Rep 2017 Sep 23;17(11):104. [doi: 10.1007/s11892-017-0935-7] [Medline: 28942581]
6. Offringa R, Sheng T, Parks L, Clements M, Kerr D, Greenfield MS. Digital Diabetes Management Application Improves Glycemic Outcomes in People With Type 1 and Type 2 Diabetes. J Diabetes Sci Technol 2018 May;12(3):701-708 [FREE Full text] [doi: 10.1177/1932296817747291] [Medline: 29277103]
7. Rasmussen BS, Froekjaer J, Bjerregaard MR, Lauritsen J, Hangaard J, Henriksen CW, et al. A Randomized Controlled Trial Comparing Telemedical and Standard Outpatient Monitoring of Diabetic Foot Ulcers. Diabetes Care 2015 Sep 26;38(9):1723-1729. [doi: 10.2337/dc15-0332] [Medline: 26116717]
8. Al Sayah F, Soprovich A, Qiu W, Edwards AL, Johnson JA. Diabetic Foot Disease, Self-Care and Clinical Monitoring in Adults with Type 2 Diabetes: The Alberta's Caring for Diabetes (ABCD) Cohort Study. Can J Diabetes 2015 Oct;39 Suppl 3:S120-S126. [doi: 10.1016/j.jcjd.2015.05.006] [Medline: 26243464]
9. Hopkins R, Burke N, Harlock J, Jegathisawaran J, Goeree R. Economic burden of illness associated with diabetic foot ulcers in Canada. BMC Health Serv Res 2015 Jan 22;15:13 [FREE Full text] [doi: 10.1186/s12913-015-0687-5] [Medline: 25608648]
10. Skrepnek GH, Mills JL, Armstrong DG. A Diabetic Emergency One Million Feet Long: Disparities and Burdens of Illness among Diabetic Foot Ulcer Cases within Emergency Departments in the United States, 2006-2010. PLoS One 2015 Aug 6;10(8):e0134914 [FREE Full text] [doi: 10.1371/journal.pone.0134914] [Medline: 26248037]
11. Canadian Radio-television and Telecommunications Commission. Communications Monitoring Report 2016: Canada's Communication System: An Overview for Canadians. 2016. URL: https://crtc.gc.ca/eng/publications/reports/policymonitoring/2016/cmr2.htm [accessed 2019-01-30] [WebCite Cache ID 75oZOWE0h]
12. Boodoo CP, Perry JA, Leung G, Cross KM, Isaranuwatchai W. Cost-effectiveness of telemonitoring screening for diabetic foot ulcer: a mathematical model. CMAJ Open 2018;6(4):E486-E494 [FREE Full text] [doi: 10.9778/cmajo.20180088] [Medline: 30337474]

Abbreviations

DFU: diabetic foot ulcer
MDM: minimally disruptive medicine
SMH: St. Michael's Hospital

Edited by G Eysenbach; submitted 08.08.18; peer-reviewed by MS Liew, B Hafeez, I Fasterholdt, RK B; comments to author 21.09.18; revised version received 10.10.18; accepted 03.11.18; published 16.04.19.

Please cite as:
Wallace D, Perry J, Yu J, Mehta J, Hunter P, Cross KM
Assessing the Need for Mobile Health (mHealth) in Monitoring the Diabetic Lower Extremity
JMIR Mhealth Uhealth 2019;7(4):e11879
URL: https://mhealth.jmir.org/2019/4/e11879/
doi:10.2196/11879
PMID:30990455

Original Paper

The Potential of Self-Management mHealth for Pediatric Cystic Fibrosis: Mixed-Methods Study for Health Care and App Assessment

Antonio Martinez-Millana[1,2], PhD; Annabel Zettl[3], BA; Jacqueline Floch[4], Dr Ing; Joaquim Calvo-Lerma[5], PhD; Jose Luis Sevillano[6], PhD; Carmen Ribes-Koninckx[5], MD, PhD; Vicente Traver[1,2], PhD

[1]Instituto ITACA, Universitat Politècnica de València, Valencia, Spain

[2]Unidad Mixta de Reingeniería de Procesos Sociosanitarios, Instituto de Investigación Sanitaria del Hospital Universitario y Politecnico La Fe, Valencia, Spain

[3]Youse GmbH, München, Germany

[4]SINTEF, Trondheim, Norway

[5]Instituto de Investigación Sanitaria La Fe, Valencia, Spain

[6]Department of Architecture and Technology of Computers, University of Seville, Seville, Spain

Corresponding Author:
Antonio Martinez-Millana, PhD
Instituto ITACA
Universitat Politècnica de València
Camino de vera sn
Valencia,
Spain
Phone: 34 963 877 606
Email: anmarmil@itaca.upv.es

Abstract

Background: Remote care services and patient empowerment have boosted mobile health (mHealth). A study of user needs related to mHealth for pediatric cystic fibrosis (PCF) identified the set of preferred features mobile apps should support; however, the potential use of PCF apps and their suitability to fit into PCF clinical management remains unexplored.

Objective: We examine whether PCF holds potential for the implementation of mHealth care.

Methods: The study is based on a literature review and qualitative analysis of content and was conducted in two parts: (1) we reviewed scientific and gray literature to explore how European countries manage PCF and conducted a qualitative study of 6 PCF units and (2) we performed a systematic review of apps available in the myhealthapps.net repository searching for cystic fibrosis (CF) management and nutrition apps, which we analyzed for characteristics, business models, number of downloads, and usability.

Results: European CF routine care guidelines are acknowledged in most European countries, and treatments are fully covered in almost all countries. The majority of teams in CF units are interdisciplinary. With respect to the systematic review of apps, we reviewed 12 apps for CF management and 9 for general nutrition management in the myhealthapps.net directory. All analyzed apps provided functionalities for recording aspects related to the disease and nutrition such as medication, meals, measurements, reminders, and educational material. None of the apps reviewed in this study supported pancreatic enzyme replacement therapy. CF apps proved to be less appealing and usable than nutrition apps (2.66 [SD 1.15] vs 4.01 [SD 0.90]; $P<.001$, z-value: –2.6). User needs detected in previous research are partially matched by current apps for CF management.

Conclusions: The health care context for PCF is a unique opportunity for the adoption of mHealth. Well-established clinical guidelines, heterogeneous clinical teams, and coverage by national health care systems provide a suitable scenario for the use of mHealth solutions. However, available apps for CF self-management do not cover essential aspects such as nutrition and education. To increase the adoption of mHealth for CF self-management, new apps should include these features.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2016-014931

(JMIR Mhealth Uhealth 2019;7(4):e13362) doi:10.2196/13362

KEYWORDS

mHealth; cystic fibrosis; health care systems; user requirements; apps; market analysis

Introduction

Pediatric cystic fibrosis (PCF) is a childhood-onset inherited disease characterized by progressive lung and pancreas deterioration, leading to premature death. Estimations in the European Union are about 30,000 cystic fibrosis (CF) patients, with one new case per 1800 to 25,000 newborns depending on the region [1]. CF inception is based in the mutation of a gene encoding the CF transmembrane conductance regulator. Despite impressive advances relating the molecular basis of CF to organ-level disease, life expectancy and quality of life are still limited in CF [2]. The combination of early diagnosis and early intervention promises to slow down the progression and improve the clinical course of the disease [3].

The primary pillars of proper CF management include a continuum of care and treatment in which the following goals are highlighted: (1) maintaining lung functions as near to normal as possible by controlling respiratory infection; (2) providing nutritional support (ie, pancreatic enzyme supplements and an adapted and balanced diet) to achieve adequate nutritional status and growth [4]; and (3) complication prevention, detection, and management [5]. To achieve these goals, a multidisciplinary team of specialized and expert health professionals is required [6,7].

National public health systems and health insurance systems are in charge of supporting and covering PCF patient treatments. While the ways countries approach these costs differ significantly [8], care services provided are similar in many high-income countries—mainly through specialized units in university hospitals and specialized centers in which follow-up outpatient visits are scheduled at 2- to 3-month intervals for physical examination and specific laboratory tests [5,8].

CF management is time-consuming and requires a multidisciplinary approach [4]. The number and variety of treatments patients are prescribed complicate both follow-up and adherence [9]. Mobile technologies (eg, mobile phones, wearable devices) provide highly scalable new approaches to disease management [10] and have shown evidence on benefits for health maintenance, disease management, and prevention of complications [11]. Mobile phone use has been rapidly increasing worldwide [12], and more than 50% of mobile phone owners use their devices to obtain health information [13].

Research in mobile health (mHealth) has demonstrated intrinsic motivations to facilitate mHealth adoption and perceived risks to inhibit it [14]. In 2013, Rai et al [15] presented a framework for evaluating drivers of mHealth implementation and found that users (patients and doctors) were more favorable to adopting mHealth as a complement to (not a substitute for) face-to-face visits. This framework has been assessed more recently [16-18], showing beneficial results when patients are supported from the very early stages after diagnosis and treatment initiation and when the technology achieves medical credibility. Nevertheless, authors highlight the need for qualitative research to assess institutional factors that would improve the acceptance of mHealth [16-18].

With regard to CF, Hilliard et al [19] asked 16 patients to report experiences using mobile apps and identify potential features to support CF self-management. Participants expressed excitement and interest in potential future mHealth apps designed to facilitate CF management, but they also raised concerns about usability, usefulness, and security. In our research within the MyCyFAPP project [20] and in the context of a user-centered design approach [21], we interviewed 74 people including persons with CF and caregivers and health professionals with CF care expertise. This study indicated the readiness for self-management in the CF community in countries that provide well-functioning health services for CF care and identified a large diversity of user requirements and a need for personalization. In addition, we derived a set of potentially useful features in an mHealth app for CF: educational tools, enzyme dosage calculation, nutritional management, treatment organization and follow-up, health diary, practical guidelines for treatment, communication with doctors, and communication with peers.

Several studies analyze the features of apps and their impact on the self-management of diseases (eg, diabetes [22], multiple sclerosis [23], cardiovascular diseases [24], and asthma [25]). These studies identify strengths and weaknesses of current apps and the requirements for improving the support of chronic diseases in the mHealth paradigm. Mobile apps have the potential to improve clinical outcomes and increase attendance and adherence rates [11].

Therefore, given that digital tools may be helpful in the self-management of CF [11] and considering the latest results on the types of features that would be particularly helpful for PCF mobile health management [21], we examined the way forward for the implementation of mHealth care by analyzing how European health care systems approach PCF management and identifying the characteristics of currently available apps for CF management. The rationale was to contribute to the adoption of mHealth in the management of PCF by the identification of strengths and opportunities. Our approach addressed the self-management of PCF using mobile technologies and targeted teenagers, young adults, and parents of young children.

The research presented in this paper was conducted within the context of the MyCyFAPP European Project (grant agreement number 643806) [20]. MyCyFAPP is developing a self-management tool for PCF patients that allows for personalized and accurate control and monitoring of the disease.

Methods

Design of Study

Health Care and App Assessment

The approach for analyzing the context of CF management consisted of two parts connected by the research objective. In part 1, analysis of CF care and health care systems was oriented to discover how different European Union countries provide overall care related to PCF. This analysis included health care and pharmacological support to patients and families through national health systems and health insurances. In addition, we analyzed the organization and coverage of health care systems in the specific domain of CF and in particular their ability to involve mobile apps in the process of managing the disease. The study did not focus on individual organizations but rather took a national perspective. In part 2, analysis of available apps in the specific context of CF in the official market stores was conducted to understand which types of functionalities are already provided and whether they are being used in a clinical context.

Part 1: National Health Care Cystic Fibrosis Coverage Analysis

The evaluation of health care services has been considered and used in the medical literature for improving the quality of care delivery and its efficiency [26,27]. Kerem et al [28] proposed a set of key performance indicators based on staff, health services, and facilities so CF centers could be compared on annual basis.

In the proposed country-by-country analysis of health care systems, we aimed to discover and compare how countries are organized to provide care and support to PCF patients based on a subset of the indicators proposed by Kerem et al [28]. We included medical aspects such as the follow-up protocol, newborn screening for CF, and composition of health care teams in CF units as well as economical aspects such as the coverage of medication costs and type of services provided to patients.

The main sources used for analysis were the European Cystic Fibrosis Society (ECFS) Patient Registry Annual Data Report from 2016 [29] and reports from national health care systems. The ECFS report contains demographic and clinical data of 44,719 consenting CF patients from 89 CF units and centers in 31 countries. Criteria for selecting countries were based on the number of registered patients and the geographical distribution and proportion of children (patients aged younger than 18 years) as collected in the ECFS 2016 report. This selection included Spain, Italy, Portugal, the Netherlands, Belgium, France, Norway, Germany, and the United Kingdom (Table 1). The health care organization and management of PCF in these countries were analyzed using official reports and literature. Data were extracted using a structured script: prevalence of PCF, protocols and guidelines for PCF management and control, insurance model, and availability of newborn screening programs.

Analysis of specific PCF units relied on semistructured interviews with professionals involved in the management of PCF who signed the informed consent from the study that was approved by the ethical committees of the participant hospitals. Beyond the context analysis, we aimed at discovering the composition of the clinical teams involved in the care and follow-up of PCF patients, the number and type of professional roles, and the size of the population covered by each unit. Interviews and one group discussion included pediatricians from 6 PCF units participating in the MyCyFAPP project who agreed to participate and gave consent to the analysis and publishing of their responses.

Table 1. Extraction of the estimated coverage, registered patients, and proportion of pediatric cystic fibrosis for the countries included in the study [29].

Country	Estimated coverage[a], %	Registered patients, n	Proportion of pediatric cystic fibrosis, %
Belgium	>90	1282	39
France	90	6713	44
Germany	>80	5738	41
Italy	95	5384	41
The Netherlands	98	1449	40
Portugal	>95	339	56
Norway	72	230	34
Spain	70	1995	54
United Kingdom	99	10,465	43

[a]Percentage of the total estimated cystic fibrosis population covered by the study.

Textbox 1. Observed characteristics and descriptions of information collected for each app.

- **App promoters:** entities that designed, developed, funded, and launched app (full name, website, reference contact information)
- **Promoter types:** hospitals, cystic fibrosis units, enterprises and the type (pharma, information and communications technology-based, etc), patient associations, charities, research projects, and other types of entities or consortia
- **Promoter goals:** short description of the main benefits for the user and intention of the promoting entities offering the app (eg, for dissemination purposes, to make a business out of the features offered, to help cystic fibrosis unit patients achieve self-management goals)
- **Geographical reach:** available languages and geographical scope in which the app would work as expected
- **Business models:** how promoters make the app sustainable in economic terms. This field indicated if there is any revenue model or payment for the use of the app or if the app was financed by the entity for promotion or as a kind of free outcome for the cystic fibrosis community. If possible, a description of the type of funding of the app was provided (eg, research funds, donations, private investment, public investment)
- **Usability score:** assessment of the user experience. Three human-computer interaction experts evaluated the usability of the apps. They used an ad hoc questionnaire composed of 5 questions focusing on design, interaction, attractiveness, structure, and understandability, each of them graded using a 5-point Likert scale in which "1" meant a low score and "5" meant a high score
- **Type of language/vocabulary:** identification of the context of use of each app (eg, if app was oriented to medical aspects and therefore written in a very technical language, if app was focused on providing guidelines and used basic or plain knowledge). Target user: adult, child, or both
- **Available platforms and facts:** marketplace availability (Google Play for Android and App Store for iOS), number of downloads per platform (as an estimation of use, although downloading does not represent use of the app), number of reviews (indicating the number of people who were interested enough to make their own assessment about app), publication date of the app on each platform and last review date (to establish app age and how often it is updated), app version currently available (to estimate evolution and releases)

Part 2: Rapid Review of Cystic Fibrosis and Nutrition Apps

The analysis of available mobile apps had the objective of reviewing solutions in the field of CF oriented to pediatric patients, families, and health professionals in the management of CF or nutrition. The rapid review considered predefined specific criteria, and the search was done on myhealthapps.net (MHA), the directory of health apps supported and reviewed by patient associations and key collaborators [26]. MHA is an indexing and searching engine for apps endorsed by patients and consumer health forums that includes apps from trustworthy developers. The identification of apps was done using the search engine at the MHA portal using the following keywords: "cystic fibrosis" and "health diary" for CF apps and "nutrition" or "food" for nutrition apps.

The inclusion criteria for apps were defined as having a (1) title or description referring to CF or (2) description referring to nutrition management linked to chronic diseases. Exclusion criteria were defined as (1) no intended use for chronic condition management and (2) not available in official app stores or repositories. Collected information included the (1) app creator, (2) app purposes, (3) identified business models, and (4) other relevant information to be considered as a reference for app evolution, sustainability, and fit into a health care system. A description of the characteristics collected in the study can be found in Textbox 1.

Statistical Analysis

Data extracted from the systematic review of apps were compared between CF and nutrition apps. The independence of categorical variables was assessed by means of chi-square tests (alpha=.05), and continuous variables were assessed using the Wilcoxon Test (alpha=.05). Statistical calculations were done using built-in functions in Matlab R2018a (The MathWorks, Inc).

Results

Part 1: National Health Care Cystic Fibrosis Coverage Analysis

General Findings

Routine care was homogeneous in the 9 countries included in this review. Regular follow-ups met international guidelines, scheduling a full assessment of PCF patients each 2 or 3 months. Newborn screening was implemented in 8 of the 9 studied countries. Treatments were fully covered in almost all countries. Only Belgium and the Netherlands only partially funded expenses related to treatment. Generally, the type of staff involved in CF units and reference centers for CF included pediatricians, dietitians, psychologists, and nurses, but some units also included social workers, physical rehabilitators, and pharmacists.

Belgium

The Belgian health care system is based on national and private health facilities funded by a combination of social security contributions and health insurance funds. The national health system reimburses a large portion of medical expenses. For CF, a full assessment every 2 to 3 months with multidisciplinary medical care and treatment will be reimbursed. The budget is managed by the national health system. For CF, some medications (eg, Creon and fat-soluble vitamins) are fully reimbursed and some medications (eg, antibiotics) are partially reimbursed. Newborn screening for CF has not yet been implemented in Belgium [30].

France

The French health care system is financed by the national health insurance and national budget. Fees and insurance levels are determined by several factors (income, consumed goods and services). CF patients are considered disabled and are eligible for 100% reimbursement from the national health insurance

system. Also, parents of children with CF are granted social rights related to the long-term care and support of the disease in the children. Newborn screening of CF is established in France [31].

Germany

Germany has a universal health care system funded by a combination of statutory health insurance and private health insurance. The system is decentralized with private practice physicians providing ambulatory care and independent, mostly nonprofit hospitals providing the majority of inpatient care. Hospitals in Germany are reimbursed by health insurance companies according to diagnosis-related groups, a flat rate system that defines reimbursement rates depending on the diagnosis. Newborn screening is implemented countrywide. Follow-up of CF patients is scheduled every 3 months in outpatient clinics with a more in-depth checkup recommended annually [32].

Italy

The national health service in Italy is universal and free of charge. The system is decentralized and divided into 14 local health agencies (*Azienda Sanitaria Locale*). Basic medical care (eg, general medicine, pediatric care, hospital admissions) is fully covered by the national system but for other types of specialized services (eg, outpatient consultations, laboratory tests, prostheses) patients usually must pay a low-cost fee determined by their health care region, income level, and subscription to a private insurance.

Patients with CF are exempt from paying for follow-ups according to a specific law (law 548/93, GU 30/12/1993, n°305). According to this law, patients with CF must be followed up at specialized CF centers, which provide multidisciplinary care (including gastroenterology and pulmonology services) and drugs at no cost for patients. Follow-up is generally scheduled every 2 to 3 months, but the frequency of visits depends on patient health status and can range from daily to every 6 months [33]. A national newborn screening program is implemented; therefore, in the great majority of the patients, first referral to a CF center occurs very early in life, soon after diagnosis [33].

The Netherlands

The Netherlands has a dual-level health care system in which primary and curative care (eg, general medicine, hospital admissions) are financed by a private mandatory insurance. These companies are obliged to provide a package with a defined set of insured treatments. This insurance covers 41% of all health care expenses. Secondary care (eg, long-term care, end-of-life care) is covered by social insurance funded by a national budget. Follow-up of CF patients is done at least every 3 months by a pulmonologist and at least once a year by the pediatric gastroenterologist [34]. Treatments and medical services for CF are fully covered. CF is included in the national newborn screening program [35].

Norway

Health care in Norway is divided into 4 regional health authorities under the responsibility of the Ministry of Health and Care. Health care services are funded by a national budget and partially covered by fees. There is a unique competence center for CF—the Norwegian Resource Center for Cystic Fibrosis (NSCF)—located in Oslo. When CF is diagnosed and treatment initiated, all patients are treated by the NSCF. The parents get training about nutrition, physiotherapy, and other important topics relevant to the patient. Later on, patients are treated by local hospitals and local physiotherapists (NSCF if they live in the southeast region). NSCF provides training to the persons responsible for providing health services locally. CF patients are followed up regularly (every 1 to 3 months depending on health status) by doctors and nutritionists at a local hospital in the region where the patient lives. All patients are normally followed up at NSCF once a year. Shared cost is set for medical supervision and medication with some services being fully covered (eg, physiotherapy). All costs related to screening, confirmation of diagnosis, further treatment, and travel costs to NSCF are covered by the public health insurance. CF newborn screening is included in the national health care system [36].

Portugal

The health care system in Portugal consists of 3 different coexisting systems: national, professional, and private insurance. The national system provides universal coverage and is split into 5 regional administrations for which the national Ministry of Health defines health policies. General patients have to pay a fee for the medical consultation; however, CF patients are exempt from these fees. CF patients are scheduled every 1 to 2 months depending on patient health status (it could range from daily to every 6 months) in the CF unit, which includes gastroenterology and pulmonology services. Follow-up includes dietitians (with a nutritional evaluation) and nurses (for adherence to medical therapy) [37]. All the costs related to CF medications are supported by the national health system. Newborn screening is implemented countrywide and includes CF [38].

Spain

The national health system in Spain is universal and free of charge. The national Ministry of Health coordinates the 17 regional health systems, but the regional systems are autonomous and define their own health policies. Spain has 22 regional CF centers, which are mainly linked to university hospitals [39]. PCF patients are followed up every 2 to 3 months with complete assessments by pediatric pneumologists and gastroenterologists [39]. Additional and more frequent visits are scheduled depending on the patient's clinical condition. An annual in-depth evaluation is also the rule in most CF units. In the case of PCF, the health care system covers all medical and pharmacy costs. Medications are delivered to the patients through the Pharmacy unit of the hospital so patients with CF do not have any associated medication costs. Additionally, oral nutritional supplements are provided for free at the hospitals. In 2012, a newborn screening program was implemented in almost all maternity wards in the country.

United Kingdom

The National Health System (NHS) in the United Kingdom is universal and free of charge. It consists of a decentralized system

divided in 4 geopolitical regions. The NHS has established a national guideline for the diagnosis, treatment, and follow-up of CF [40] for health care professionals in primary care, health care professionals in secondary care, providers of CF services, and practitioners in CF. CF prescription charges have been abolished entirely for people in Scotland, Northern Ireland, and Wales, but in England many people with CF still have to pay for their prescriptions. The Cystic Fibrosis Trust offers financial support for emergency care. Finally, the United Kingdom has deployed newborn screening countrywide [41].

Center-Level Analysis

In Multimedia Appendix 1, we show a comparison of the specific managerial and clinical organization of 6 European CF units that are participants in the MyCyFAPP project [20] and gave us access to information about their institutions. Among these centers, 4 of the 6 have a separate unit for managing PCF patients. The median number of pediatric patients (patients aged younger than 18 years) in CF units is 127. The center with the highest number of children is the University Hospital of Milan (399); the lowest is the *Associação Para Investigação, Desenvolvimento Da Faculdade De Medicina* of Lisbon (60).

The main difference among centers is the composition of clinical teams involved in PCF care and follow-up. All the interviewed units have roles for covering gastrointestinal, pulmonary, and nutritional aspects of the disease (adhering to current guidelines [4]), but we observe that one (*Ramon y Cajal Hospital* in Madrid) includes two endocrinologists. Two other differences are the inclusion of social workers and psychologists at the University Hospital of Milan, Erasmus University Medical Center, and KU Leuven Hospital and the inclusion of a pharmacologist at the KU Leuven Hospital.

Involved clinical staff is heterogeneous among and within countries; differences are not only country-specific but depend on specific constraints at a hospital/unit level, as we observe by comparing the two units in Spain (*La Fe* vs *Ramon y Cajal Hospital*).

Part 2: Systematic Review of Cystic Fibrosis and Nutrition Apps

Searches in the MHA directory [42] were performed at the beginning of the MyCyFAPP project (February 2016) and provided more than 500 apps for CF or nutrition management. Of the total, 450 were general purpose apps for managing chronic conditions from which we selected the apps targeted at nutrition and CF-specific apps based on the inclusion and exclusion criteria. Out of them, 9 were specifically targeted at nutrition, 4 were CF-specific games, and 12 were CF-specific apps (Multimedia Appendix 2).

Relevant findings that summarize related aspects in CF-specific mobile apps on the market were classified according to the data extraction criteria (Multimedia Appendix 3). The majority of the apps (19/21, 90%) were provided by individual developers (small to medium enterprises, individuals). Few apps (2/21, 9%) were promoted by the pharma industry (*Cf MedCare* and *MedSched*). One of them, oriented to patients, has very limited features and poor design, and one, oriented to health care professionals, has a focused content in genes and mutations. Only two apps (2/21, 9%) were promoted by hospitals (*Tools4U* and *Genia*), one app (1/21, 5%) by a patient association (*Flower Breath*), and one implemented by the IT department of a hospital from the NHS Tayside region (*Cystic Fibrosis: A Pocket Guide*).

Many of the analyzed apps focused on scientific information and content for education. In these cases, the content was tailored not to patients but to researchers. None of the apps supported pancreatic enzyme replacement therapy (PERT). Most of the analyzed CF-specific apps did not have a concrete business model to ensure sustainability. In Table 2, we summarize the relevant variables by comparing specific and nonspecific CF apps. Download payment ranged from 0 to 2 Euro (US $2.26), and 2/21 apps (9%) included advertisements as a revenue channel.

For non-CF specific apps, business models were much more consolidated, offering coaching services related to calorie intake and healthy nutritional habits. CF-specific apps were focused on the management of the disease (medication and pulmonary symptoms). All of the analyzed apps provided functionalities for recording aspects related to the disease such as medication, meals, and body measurements (eg, weight, forced expiration volume). In addition, some apps provided reminders and educational material. Only two apps provided a feature for communication with doctors and a virtual social community to get in touch with other patients. Nutrition apps were characterized by providing a wide range of features tailored to the user. Features for coaching, personalizing goals, and aspects of food management were predominant in these apps. Only one app was endorsed by international organizations. Most apps are available for both Android and iOS (except for the United States and Sweden, where only iOS versions are available).

Overall, CF apps showed different results in the observed metrics. Reviews and downloads were lower for CF apps than for nutrition apps, as nutrition apps were used by a wider range of users including healthy people and patients with chronic conditions. CF apps were less appealing and usable than nutrition apps (2.66 [SD 1.15] versus 4.01 [SD 0.90], $P<.001$, z-value: –2.6). Updates of apps did not include features for the specific management of CF patient needs, such as nutritional counseling, therapy advice, and health literacy.

Table 2. Comparative analysis of available apps for cystic fibrosis management and nutrition.

Characteristic	CF[a]-specific apps (n=11), n (%)	Nutrition apps (n=9), n (%)	*P* value	*z*-value
Licensing			.19	6.0
Free	8 (72)	0 (0)		
Payment	3 (27)	2 (22)		
Freemium[b]	1 (1)	7 (78)		
Targeted user			.20	6.0
Children	2 (18)	1 (11)		
Adults	6 (54)	8 (89)		
All	3 (27)	0 (0)		
Number of reviews, range	5-2160	0-55,000	<.001	–1.6
Number of downloads			.26	11.2
1-100	0	0		
100-1000	4	1		
1000-100,000	3	4		
100,000-1,000,000	1	3		
1,000,000-5,000,000	0	1		
Last update, aggregate	2017	2018	—[c]	—
Usability score, mean (SD)	2.66 (1.15)	4.01 (0.90)	<.001	–2.6

[a]CF: cystic fibrosis.

[b]Freemium model combines free basic services and advanced payment services.

[c]Not applicable.

Discussion

Principal Finding 1: Homogeneity of Health Care Systems in Pediatric Cystic Fibrosis

Routine care was quite homogeneous in Western Europe, based on regular follow-up every 2 or 3 months with full assessment of patients with CF. Treatments were fully covered in almost all countries. Only Belgium and the Netherlands were partially funding the expenses related to treatments. Only Belgium had not established a program for newborn screening for CF. The fact that the visits are scheduled at similar intervals does not mean a similar care delivery, as some are more focused on gastroenterology and nutrition and others more on pulmonology.

Principal Finding 2: Heterogeneity of Clinical Teams in Pediatric Cystic Fibrosis

Generally, the type of staff involved in CF units includes pediatricians for gastroenterology and pulmonology. Most of the units also included dietitians/nutritionists and nurses. A few units included social workers, psychologists, physical rehabilitators, and pharmacists.

Principal Finding 3: Apps for Cystic Fibrosis Partially Meet User Needs

Apps for CF were focused on the management of medication, meal intake, pulmonary symptom recording, and education. Our systematic review concludes that current CF apps only partially meet the user needs identified by Floch et al [21], PERT management was not featured in any of the discovered apps, and communication with peers and doctors was implemented only in two apps. From the promoter point of view, the presence of the pharma industry in the app market was limited to one company with a relatively weak presence. Many tools were promoted by individuals and small companies without links to health care systems or CF units. Associations of patients were another group of entities that promoted apps, but once more, their presence was limited and normally their operation range was in the scope of a country or a region. Finally, the most interesting promoting entity was a health care provider, namely a hospital that used the app as a tool for improving care processes for its patients, although support was still limited in terms of available features. CF apps yielded a higher representation in the free licensing business model, whereas nutrition apps showed higher distribution in pay-per-use and in-app purchase business models, although this different behavior was not statistically significant. We evaluated the usability of apps by means of an ad hoc questionnaire and the analysis of three human-computer interaction experts. App usefulness, meaning the ability of the app to cope with actual user needs, is an important attribute but difficult to measure without involving users. We therefore included as surrogate parameters the number of reviews and number of downloads as recorded in MHA at the time of the review.

mHealth in Pediatric Cystic Fibrosis Management: The Time Is Now

Care models in Western European countries differ slightly in the management of CF, and some of them provide a favorable context for support of self-management. Nutrition is a key factor in the progression of any disease [43], and CF standards of care guidelines recommend nutrition counseling and support at any stage of the disease by expert dietitians [44]. The interviewed units included nutritionist or dietitians on the clinical team, while other disciplines (eg, social workers, psychologists, endocrinologists) weren't represented in all units. Mobile apps could fill this gap by delivering psychosocial advice and follow-up and, in parallel, offering services for multidisciplinary control by means of distributed technology (external services or services from other countries).

Complete patient empowerment supporting shared decision making (ie, patient is involved in health care decision-making processes) and self-determination (ie, patient has the power to choose own goals) is unrealistic in the case of CF [21]. We do not see mHealth as a unique tool for managing PCF but as a combination of self-management and patient empowerment with telemedicine tools for accessing medical, nutritional, and psychological services irrespective of patient location. This model has been already tested in other chronic diseases in children (eg, type 1 diabetes mellitus), in which a combination of self-reporting outcomes, communication and tools for education, and gamification has shown positive health impact [22]. In a recent systematic review of 23 mHealth interventions, Marcolino et al [11] discovered that the most popular mHealth intervention was based on strategies to promote behavior change. In combination with our results, this shows that the potential of mHealth in PCF needs further efforts, and apps must go beyond offering management tools to support, educate, and motivate patients to adopt and maintain healthier habits.

CF is a disease for which tremendous progress has been made in the field of medical care [45] but for which costs of care are very high and in the case of nonhospital costs can account for up to 50% of total lifetime expenditures of the disease [46]. Children with CF are usually diagnosed in the first year of their life [47] and subsequently need intensive support from family and health care services. In a study on the effect of deprivation, Taylor-Robinson et al [48] discovered that children with CF from more disadvantaged areas had worse growth and lung function compared with children from more affluent areas. Disadvantaged families had a higher burden of treatment, leading to a disruption of school and family life. In this context, telemedicine and mHealth can ameliorate the access to health care services and follow-up by means of remote consultation, access to health education, and patient empowerment [49,50]. Equality in access to health care services has been recognized as a key factor to guarantee equal medical assistance and well-being for all European citizens [51].

The current approach for PCF management is focused on reacting to the effect of health deterioration, focusing on the acute phase of the disease. Democratization of technology is a good opportunity to change the health care management model with the essential objective of improving the quality of life by shifting the hospital-centered approach to a patient-centered paradigm. A better understanding of disease mechanisms and treatment effects leads to an improvement in health outcomes [52]. This suggests that small changes in children's environment can have large effects on behavior and can be used in self-management of PCF.

The introduction of telemedicine will require changes in health care systems and CF units, but before organizational and cost-effectiveness studies are conducted, further research is needed on the implementation of systems that cover specific patient needs and engage patients and relatives to use the technology for managing the disease.

CF apps including the features of nutrition and PERT self-management have great potential to improve the prognosis of the disease as nutrition is an action that occurs every day, several times a day. If we achieve an adequate nutrient intake and PERT dose, the nutritional status will very likely improve. The importance of this is that nutritional status is the parameter that most correlates with disease prognosis and survival.

Limitations

Our study has some limitations. Part 1 covered a wide proportion of European countries, although not all of them have been included. Data were collected with the help of the ECFS Patient Registry [29], but data coverage is relatively low in some countries, meaning that not everyone with CF is represented in the data. Moreover, the study has been limited to treatment, follow-up, and general organizational information, excluding other interesting indicators such as costs, resources, and macroeconomic indicators.

Our evaluation of the heterogeneity of clinical teams in PCF has a high risk of selection bias as we only used data from 6 PCF units participating in the MyCyFAPP project. Although the information from these units is reliable, these may not be a representative sample of all units in Europe. In the part 2, the systematic review consisted of queries to the MHA directory search engine, and some apps for CF management may have been missed because the search terms didn't match the description or the metadata. The unique source for the systematic review of apps was the MHA repository [26] because it contains apps endorsed by patient groups and trustworthy developers. A wider search of apps may be conducted in official app stores such as Google Play (Android) and App Store (iOS) to increase the sample size. Finally, for the evaluation of apps we defined a simple score to discriminate design and appeal, although a standard quality assessment tool should be used to perform an in-depth analysis.

Comparison With Prior Work

mHealth is a big opportunity to relieve the burden of health care systems [12]. Apps are increasingly being used to follow treatments, improve patient adherence [53], and provide remote access to health services [54]. Prior research has focused on the potential of mHealth in the self-management of diseases [23], but the potential of mHealth to meet the existing needs of patients using existing app functionalities should also be considered.

A mobile app represents a unique channel for the direct communication between patients and doctors beyond classical mechanisms like phone calls or emails. Health care professionals will have to adapt their clinical protocols and internal procedures to this new paradigm [24]. More evidence is needed on the impact and effectiveness of integrating such tools into CF routine care [55]. Wearable sensors and tracking systems [56] are opportunities for the automatic collection of data by CF-related apps, a factor that could increase their effectiveness in follow-up, gamification, and motivation.

An explosion of health-related apps has appeared in the different app markets in the last 5 to 7 years [11]. New apps are landing not in an incipient research and market scenario but in a consolidated research field and business. Apps in health should demonstrate their effectiveness in meeting clinical and engagement end points in the long term [57]. Multidisciplinary research should aim at delivering tools for remote and self-management from several perspectives, including usability aspects, motivation factors, perceived usefulness, and economic sustainability [58].

Conclusions

The health care context for PCF is a good opportunity for the development of mHealth solutions. The heterogeneity of clinical teams, coverage of national health systems, and geographical distribution of CF units build a suitable scenario for the use of mHealth solutions. CF apps included in our study are mainly based on symptom recording and management of reminders for medications and meals. mHealth features for specific nutritional management, social support, communication with health professionals, and decision support should be explored to increase the adoption of apps in the management of CF.

Acknowledgments

The authors appreciate the commitment of the European Commission MyCyFAPP (H2020–643806) project consortium during concept development and execution of this research. The authors also wish to acknowledge the ITACA Institute (Universitat Politècnica de València) for making the publication of this paper possible through the Excellence Support program for the publication in high-impact international journals.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of the characteristics of each referenced pediatric cystic fibrosis center.

[PDF File (Adobe PDF File), 217KB - mhealth_v7i4e13362_app1.pdf]

Multimedia Appendix 2

Diagram of searches and selected apps.

[PDF File (Adobe PDF File), 206KB - mhealth_v7i4e13362_app2.pdf]

Multimedia Appendix 3

List of analyzed apps and their features.

[PDF File (Adobe PDF File), 124KB - mhealth_v7i4e13362_app3.pdf]

References

1. Mehta G, Macek M, Mehta A, European Registry Working Group. Cystic fibrosis across Europe: EuroCareCF analysis of demographic data from 35 countries. J Cyst Fibros 2010 Dec;9 Suppl 2:S5-S21 [FREE Full text] [doi: 10.1016/j.jcf.2010.08.002] [Medline: 21041121]
2. Sharma N, Singh M, Kaur G, Thapa BR, Prasad R. Identification and characterization of CFTR gene mutations in Indian CF patients. Ann Hum Genet 2009 Jan;73(1):26-33 [FREE Full text] [doi: 10.1111/j.1469-1809.2008.00477.x] [Medline: 18782298]
3. VanDevanter DR, Wagener JS, Pasta DJ, Elkin E, Jacobs JR, Morgan WJ, et al. Pulmonary outcome prediction (POP) tools for cystic fibrosis patients. Pediatr Pulmonol 2010 Dec;45(12):1156-1166 [FREE Full text] [doi: 10.1002/ppul.21311] [Medline: 20717915]
4. Castellani C, Duff AJA, Bell SC, Heijerman HGM, Munck A, Ratjen F, et al. ECFS best practice guidelines: the 2018 revision. J Cyst Fibros 2018 Mar;17(2):153-178 [FREE Full text] [doi: 10.1016/j.jcf.2018.02.006] [Medline: 29506920]
5. Colombo C, Littlewood J. The implementation of standards of care in Europe: state of the art. J Cyst Fibros 2011 Jun;10 Suppl 2:S7-S15 [FREE Full text] [doi: 10.1016/S1569-1993(11)60003-9] [Medline: 21658645]

6. Goss C, Burns J. Exacerbations in cystic fibrosis. 1: Epidemiology and pathogenesis. Thorax 2007 Apr;62(4):360-367 [FREE Full text] [doi: 10.1136/thx.2006.060889] [Medline: 17387214]
7. Conway S, Balfour-Lynn IM, De Rijcke K, Drevinek P, Foweraker J, Havermans T, et al. European Cystic Fibrosis Society standards of care: framework for the Cystic Fibrosis Centre. J Cyst Fibros 2014 May;13 Suppl 1:S3-S22 [FREE Full text] [doi: 10.1016/j.jcf.2014.03.009] [Medline: 24856776]
8. Sheppard DN. EuroCareCF: working together to improve patient care and therapy development. J Cyst Fibros 2011 Jun;10 Suppl 2:S1-S4 [FREE Full text] [doi: 10.1016/S1569-1993(11)00089-0] [Medline: 21658628]
9. Michl RK, Tabori H, Hentschel J, Beck JF, Mainz JG. Clinical approach to the diagnosis and treatment of cystic fibrosis and CFTR-related disorders. Expert Rev Respir Med 2016 Oct 12:1-10. [doi: 10.1080/17476348.2016.1240032] [Medline: 27653767]
10. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
11. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, 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]
12. World Health Organization. 2011. New horizons for health through mobile technologies. Global observatory for eHealth series, volume 3 URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2019-04-09] [WebCite Cache ID 77W0oByJi]
13. Dehling T, Gao F, Schneider S, Sunyaev A. Exploring the far side of mobile health: information security and privacy of mobile health apps on iOS and Android. JMIR Mhealth Uhealth 2015;3(1):e8 [FREE Full text] [doi: 10.2196/mhealth.3672] [Medline: 25599627]
14. Cocosila M, Archer N. BLED 2009 Proceedings. 2009. An empirical investigation of mobile health adoption in preventive interventions URL: https://pdfs.semanticscholar.org/75d3/8902104741e225414955c06409f6eae7e846.pdf [accessed 2019-04-09] [WebCite Cache ID 77W12Lpha]
15. Rai A, Chen L, Pye J, Baird A. Understanding determinants of consumer mobile health usage intentions, assimilation, and channel preferences. J Med Internet Res 2013 Aug 02;15(8):e149 [FREE Full text] [doi: 10.2196/jmir.2635] [Medline: 23912839]
16. Wang SH. Web-based medical service: technology attractiveness, medical creditability, information source, and behavior intention. J Med Internet Res 2017 Aug 02;19(8):e285 [FREE Full text] [doi: 10.2196/jmir.8114] [Medline: 28768608]
17. Langius-Eklöf A, Crafoord M, Christiansen M, Fjell M, Sundberg K. Effects of an interactive mHealth innovation for early detection of patient-reported symptom distress with focus on participatory care: protocol for a study based on prospective, randomised, controlled trials in patients with prostate and breast cancer. BMC Cancer 2017 Jul 04;17(1):466 [FREE Full text] [doi: 10.1186/s12885-017-3450-y] [Medline: 28676102]
18. Blumenthal J, Wilkinson A, Chignell M. Physiotherapists' and physiotherapy students' perspectives on the use of mobile or wearable technology in their practice. Physiother Can 2018;70(3):251-261. [doi: 10.3138/ptc.2016-100.e] [Medline: 30275650]
19. Hilliard ME, Hahn A, Ridge AK, Eakin MN, Riekert KA. User preferences and design recommendations for an mhealth app to promote cystic fibrosis self-management. JMIR Mhealth Uhealth 2014;2(4):e44 [FREE Full text] [doi: 10.2196/mhealth.3599] [Medline: 25344616]
20. MyCyFAPP. URL: http://www.mycyfapp.eu/index.php/en/ [accessed 2018-09-27] [WebCite Cache ID 72k7M3jQ3]
21. Floch J, Zettl A, Fricke L, Weisser T, Grut L, Vilarinho T, et al. User needs in the development of a health app ecosystem for self-management of cystic fibrosis: user-centered development approach. JMIR Mhealth Uhealth 2018 May 08;6(5):e113 [FREE Full text] [doi: 10.2196/mhealth.8236] [Medline: 29739742]
22. Martinez-Millana A, Jarones E, Fernandez-Llatas C, Hartvigsen G, Traver V. App features for type 1 diabetes support and patient empowerment: systematic literature review and benchmark comparison. JMIR Mhealth Uhealth 2018 Nov 21;6(11):e12237 [FREE Full text] [doi: 10.2196/12237] [Medline: 30463839]
23. Giunti G, Guisado FE, Dorronzoro ZE, Rivera RO. Supply and demand in mHealth apps for persons with multiple sclerosis: systematic search in app stores and scoping literature review. JMIR Mhealth Uhealth 2018 May 23;6(5):e10512 [FREE Full text] [doi: 10.2196/10512] [Medline: 29792295]
24. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
25. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. Telemed J E Health 2014 Jan;20(1):75-82 [FREE Full text] [doi: 10.1089/tmj.2012.0328] [Medline: 24205809]
26. Swanson RC, Bongiovanni A, Bradley E, Murugan V, Sundewall J, Betigeri A, et al. Toward a consensus on guiding principles for health systems strengthening. PLoS Med 2010 Dec 21;7(12):e1000385 [FREE Full text] [doi: 10.1371/journal.pmed.1000385] [Medline: 21203584]

27. Bosch-Capblanch X, Lavis JN, Lewin S, Atun R, Røttingen J, Dröschel D, et al. Guidance for evidence-informed policies about health systems: rationale for and challenges of guidance development. PLoS Med 2012;9(3):e1001185 [FREE Full text] [doi: 10.1371/journal.pmed.1001185] [Medline: 22412356]
28. Kerem E, Conway S, Elborn S, Heijerman H, Consensus Committee. Standards of care for patients with cystic fibrosis: a European consensus. J Cyst Fibros 2005 Mar;4(1):7-26 [FREE Full text] [doi: 10.1016/j.jcf.2004.12.002] [Medline: 15752677]
29. European Cystic Fibrosis Society. ECFS Patient Registry Annual Data Report 2016 URL: https://www.ecfs.eu/sites/default/files/general-content-images/working-groups/ecfs-patient-registry/ECFSPR_Report2016_06062018.pdf [accessed 2019-04-10] [WebCite Cache ID 77XKgy53v]
30. De Boeck K, Munck A, de Monestrol I, Gulmans V, Lemonnier L, Middleton PG, et al. Does newborn screening influence the young cystic fibrosis cohort included in national registries? Eur Respir J 2017 Dec;49(1) [FREE Full text] [doi: 10.1183/13993003.00686-2016] [Medline: 28077474]
31. De Boeck K, Bulteel V, Tiddens H, Wagner T, Fajac I, Conway S, ECFS-CTN network partners. Guideline on the design and conduct of cystic fibrosis clinical trials: the European Cystic Fibrosis Society-Clinical Trials Network (ECFS-CTN). J Cyst Fibros 2011 Jun;10 Suppl 2:S67-S74 [FREE Full text] [doi: 10.1016/S1569-1993(11)60010-6] [Medline: 21658644]
32. Naehrig S, Chao C, Naehrlich L. Cystic Fibrosis. Dtsch Arztebl Int 2017 Dec 21;114(33-34):564-574 [FREE Full text] [doi: 10.3238/arztebl.2017.0564] [Medline: 28855057]
33. Fuiano S, Zoccali C. [Linee Guida per la diagnosi e terapia della nefropatia diabetica]. G Ital Di Nefrol 2003;20(S-24):S96-S108 [FREE Full text]
34. Dutch Cystic Fibrosis Registy Report 2017. URL: https://www.ncfs.nl/bestanden/rapport-cf-registratie-2017.pdf [accessed 2019-04-14] [WebCite Cache ID 77dAbNqn8]
35. Dankert-Roelse JE, Bouva MJ, Jakobs BS, Janssens HM, de Winter-de Groot KM, Schönbeck Y, et al. Newborn blood spot screening for cystic fibrosis with a four-step screening strategy in the Netherlands. J Cyst Fibros 2018 Aug 23. [doi: 10.1016/j.jcf.2018.07.008] [Medline: 30146269]
36. Lundman E, Gaup HJ, Bakkeheim E, Olafsdottir EJ, Rootwelt T, Storrøsten OT, et al. Implementation of newborn screening for cystic fibrosis in Norway: results from the first three years. J Cyst Fibros 2016 Dec;15(3):318-324 [FREE Full text] [doi: 10.1016/j.jcf.2015.12.017] [Medline: 26795017]
37. Silva A, Amorim A, Azevedo P, Lopes C, Gamboa F. Cystic fibrosis: characterization of the adult population in Portugal. Rev Port Pneumol (2006) 2016;22(3):141-145. [doi: 10.1016/j.rppnen.2015.12.010] [Medline: 26898888]
38. Marcão A, Barreto C, Pereira L, Vaz L, Cavaco J, Casimiro A, et al. Cystic fibrosis newborn screening in Portugal: PAP value in populations with stringent rules for genetic studies. Int J Neonatal Screen 2018 Jun 29;4(3):22. [doi: 10.3390/ijns4030022]
39. Castro HE, Aguirre AS, Ortega JMN. Fibrosis quística. Protocolos diagnóstico-terapéuticos de Gastroenterología, Hepatología y Nutrición Pediátrica 2013:77-84 [FREE Full text]
40. Sermet-Gaudelus I, Mayell SJ, Southern KW, European Cystic Fibrosis Society (ECFS), Neonatal Screening Working Group. Guidelines on the early management of infants diagnosed with cystic fibrosis following newborn screening. J Cyst Fibros 2010 Sep;9(5):323-329 [FREE Full text] [doi: 10.1016/j.jcf.2010.04.008] [Medline: 20605539]
41. Public Health England. 2017. NHS Newborn Blood Spot Screening Programme: managing positive results from cystic fibrosis screening URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/642329/Managing_positive_results_from_cystic_fibrosis_screening_2017.pdf [accessed 2019-04-10] [WebCite Cache ID 77XDQdgoo]
42. MyHealthApps Directory. URL: http://myhealthapps.net/ [accessed 2019-04-10] [WebCite Cache ID 72kA2oRsw]
43. Cederholm T, Barazzoni R, Austin P, Ballmer P, Biolo G, Bischoff SC, et al. ESPEN guidelines on definitions and terminology of clinical nutrition. Clin Nutr 2017 Dec;36(1):49-64. [doi: 10.1016/j.clnu.2016.09.004] [Medline: 27642056]
44. Smyth AR, Bell SC, Bojcin S, Bryon M, Duff A, Flume P, European Cystic Fibrosis Society. European Cystic Fibrosis Society standards of care: best practice guidelines. J Cyst Fibros 2014 May;13 Suppl 1:S23-S42. [doi: 10.1016/j.jcf.2014.03.010] [Medline: 24856775]
45. Burgel P, Bellis G, Olesen HV, Viviani L, Zolin A, Blasi F, ERS/ECFS Task Force on Provision of Care for Adults with Cystic Fibrosis in Europe. Future trends in cystic fibrosis demography in 34 European countries. Eur Respir J 2015 Jul;46(1):133-141. [doi: 10.1183/09031936.00196314] [Medline: 25792639]
46. Wildhagen MF, Verheij JB, Verzijl JG, Gerritsen J, Bakker W, Hilderink HB, et al. The nonhospital costs of care of patients with CF in The Netherlands: results of a questionnaire. Eur Respir J 1996 Nov;9(11):2215-2219 [FREE Full text] [Medline: 8947062]
47. Stern M, Bertrand DP, Bignamini E, Corey M, Dembski B, Goss CH, et al. European Cystic Fibrosis Society standards of care: quality management in cystic fibrosis. J Cyst Fibros 2014 May;13 Suppl 1:S43-S59 [FREE Full text] [doi: 10.1016/j.jcf.2014.03.011] [Medline: 24856777]
48. Taylor-Robinson DC, Smyth RL, Diggle PJ, Whitehead M. The effect of social deprivation on clinical outcomes and the use of treatments in the UK cystic fibrosis population: a longitudinal study. Lancet Respir Med 2013 Apr;1(2):121-128 [FREE Full text] [doi: 10.1016/S2213-2600(13)70002-X] [Medline: 24429092]

49. Ketchell RI. Telemedicine is the way forward for the management of cystic fibrosis—the case in favour. Paediatr Respir Rev 2018 Mar;26:19-21. [doi: 10.1016/j.prrv.2017.03.004] [Medline: 28416301]
50. Lenney W. Telemedicine is the way forward for the management of cystic fibrosis—the case against. Paediatr Respir Rev 2018 Mar;26:22-23. [doi: 10.1016/j.prrv.2017.03.005] [Medline: 28366680]
51. European Commission. 2013 Sep. Commission staff working document: report on health inequalities in the European Union URL: https://ec.europa.eu/health/sites/health/files/social_determinants/docs/report_healthinequalities_swd_2013_328_en.pdf [accessed 2019-04-10] [WebCite Cache ID 77XEGgIv1]
52. Adams RJ. Improving health outcomes with better patient understanding and education. Risk Manag Healthc Policy 2010;3:61-72 [FREE Full text] [doi: 10.2147/RMHP.S7500] [Medline: 22312219]
53. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. J Health Commun 2012;17 Suppl 1:82-104. [doi: 10.1080/10810730.2011.649160] [Medline: 22548603]
54. Devi BR, Syed-Abdul S, Kumar A, Iqbal U, Nguyen P, Li YJ, et al. mHealth: an updated systematic review with a focus on HIV/AIDS and tuberculosis long term management using mobile phones. Comput Methods Programs Biomed 2015 Nov;122(2):257-265. [doi: 10.1016/j.cmpb.2015.08.003] [Medline: 26304621]
55. Cummings E, Hauser J, Cameron-Tucker H, Fitzpatrick P, Jessup M, Walters EH, et al. Enhancing self-efficacy for self-management in people with cystic fibrosis. Stud Health Technol Inform 2011;169:33-37. [Medline: 21893709]
56. Bayo-Monton J, Martinez-Millana A, Han W, Fernandez-Llatas C, Sun Y, Traver V. Wearable sensors integrated with Internet of Things for advancing eHealth care. Sensors (Basel) 2018 Jun 06;18(6) [FREE Full text] [doi: 10.3390/s18061851] [Medline: 29882790]
57. Sun C, Malcolm JC, Wong B, Shorr R, Doyle M. Improving glycemic control in adults and children with type 1 diabetes with the use of smartphone-based mobile applications: a systematic review. Can J Diabetes 2018 Mar 27:1-8. [doi: 10.1016/j.jcjd.2018.03.010] [Medline: 30026048]
58. Kujala S. User involvement: a review of the benefits and challenges. Behav Inf Technol 2003 Jan;22(1):1-16. [doi: 10.1080/01449290301782]

Abbreviations

CF: cystic fibrosis
ECFS: European Cystic Fibrosis Society
MHA: myhealthapps.net directory
mHealth: mobile health
NSCF: Norwegian Resource Center for Cystic Fibrosis
NHS: National Health System
PCF: pediatric cystic fibrosis
PERT: pancreatic enzyme replacement therapy

Edited by G Eysenbach; submitted 11.01.19; peer-reviewed by J del Hoyo Francisco, T Vagg, L Collins; comments to author 07.02.19; revised version received 05.03.19; accepted 29.03.19; published 18.04.19.

Please cite as:
Martinez-Millana A, Zettl A, Floch J, Calvo-Lerma J, Sevillano JL, Ribes-Koninckx C, Traver V
The Potential of Self-Management mHealth for Pediatric Cystic Fibrosis: Mixed-Methods Study for Health Care and App Assessment
JMIR Mhealth Uhealth 2019;7(4):e13362
URL: http://mhealth.jmir.org/2019/4/e13362/
doi:10.2196/13362
PMID:30998222

Original Paper

A Smartphone App to Assist Smoking Cessation Among Aboriginal Australians: Findings From a Pilot Randomized Controlled Trial

David Peiris[1], PhD, MBBS, MPH; Lachlan Wright[1]; Madeline News[1], MSci; Kris Rogers[1], PhD; Julie Redfern[2], PhD; Clara Chow[2], PhD; David Thomas[3], PhD

[1]The George Institute for Global Health, UNSW Sydney, Newtown, Australia

[2]Westmead Applied Research Centre, University of Sydney, Sydney, Australia

[3]Menzies School of Health Research, Darwin, Australia

Corresponding Author:
David Peiris, PhD, MBBS, MPH
The George Institute for Global Health, UNSW Sydney
1 King St
Newtown,
Australia
Phone: 61 280524513
Fax: 61 280524500
Email: dpeiris@georgeinstitute.org

Abstract

Background: Mobile health (mHealth) apps have the potential to increase smoking cessation, but little research has been conducted with Aboriginal communities in Australia.

Objective: We conducted a pilot study to assess the feasibility and acceptability and explore the effectiveness of a novel mHealth app to assist Aboriginal people to quit smoking.

Methods: A pilot randomized controlled trial (RCT) and process evaluation comprising usage analytics data and in-depth interviews was conducted. Current Aboriginal smokers (>16 years old), who were willing to make a quit attempt in the next month, were recruited from Aboriginal Community Controlled Health Services and a government telephone coaching service. The intervention was a multifaceted Android or iOS app comprising a personalized profile and quit plan, text and in-app motivational messages, and a challenge feature allowing users to compete with others. The comparator was usual cessation support services. Outcome data collection and analysis were conducted blinded to treatment allocation. The primary outcome was self-reported continuous smoking abstinence verified by carbon monoxide breath testing at 6 months. Secondary outcomes included point prevalence of abstinence and use of smoking cessation therapies and services.

Results: A total of 49 participants were recruited. Competing service delivery priorities, the lack of resources for research, and lack of support for randomization to a control group were the major recruitment barriers. At baseline, 23/49 (47%) of participants had tried to quit in recent weeks. At 6-month follow-up, only 1 participant (intervention arm) was abstinent. The process evaluation highlighted low to moderate app usage (3-10 new users per month and 4-8 returning users per month), an average of 2.9 sessions per user per month and 6.3 min per session. Key themes from interviews with intervention participants (n=15) included the following: (1) the powerful influence of prevailing social norms around acceptability of smoking; (2) high usage of mobile devices for phone, text, and social media but very low use of other smartphone apps; (3) the role of family and social group support in supporting quit attempts; and (4) low awareness and utilization of smoking cessation support services. Despite the broad acceptability of the app, participants also recommended technical improvements to improve functionality, greater customization of text messages, integration with existing social media platforms, and gamification features.

Conclusions: Smoking cessation apps need to be integrated with commonly used functions of mobile phones and draw on social networks to support their use. Although they have the potential to increase utilization of cessation support services and treatments, more research is needed to identify optimal implementation models. Robust evaluation is critical to determine their impact; however, an RCT design may not be feasible in this setting.

Trial Registration: Australian and New Zealand Clinical Trials Registry ACTRN12616001550493; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=371792 (Archived by WebCite at http://www.webcitation.org/76TiV7HA6).

(JMIR Mhealth Uhealth 2019;7(4):e12745) doi:10.2196/12745

KEYWORDS

smoking cessation; oceanic ancestry group; mobile apps

Introduction

Smoking Rates for Aboriginal and Torres Strait Islander People

Aboriginal and Torres Strait Islander people comprise approximately 3% of the total Australian population [1]. Although there have been recent encouraging declines in smoking overall in Australia, in 2014-15, 39% of Aboriginal and Torres Strait Islander people over the age of 15 years smoked cigarettes daily (36% of females and 42% of males), which is 2.8 times the nonindigenous prevalence [1,2]. Although a smaller proportion of Aboriginal and Torres Strait Islander smokers compared with the general Australian population have ever tried to quit (69% vs 81.4%, respectively) or have ever sustained a quit attempt for more than or equal to 1 month (47% vs 60%), Aboriginal and Torres Strait Islander smokers are just as likely to have attempted quitting in the previous year [3]. There is less social disapproval of smoking among Aboriginal and Torres Strait Islander people when compared with the general population (62% vs 79%), and daily Aboriginal and Torres Strait Islander smokers are less likely to use nicotine replacement therapy (NRT) and other smoking cessation therapies than the general population (37% vs 59%) [3].

Mobile Health Interventions for Smoking Cessation

Mobile health (mHealth) interventions are emerging as potential strategies to increase smoking cessation. A recent systematic review of 12 studies and 11,885 participants found that smokers who received mHealth support were around 1.7 times more likely to abstain from smoking than smokers who did not receive the programs [4]. Another systematic review examined text messaging interventions for smoking cessation (22 interventions, 10 countries, and 15,593 smokers) [5]. Smokers who received a text messaging intervention were 37% more likely to abstain from smoking relative to controls across several smoking abstinence measures (point prevalence, continuous abstinence, prolonged abstinence, and repeated point prevalence).

There has been only 1 randomized controlled trial (RCT) of an mHealth intervention for smoking cessation in Australia. QuitCoach is a personalized tailored internet-delivered advice program, and onQ (now named QuiTxt) is an interactive automated text messaging program [6]. In a randomized trial comprising 3530 smokers or recent quitters, only 42.5% of those offered 1 of the interventions engaged with those interventions. There were no significant differences between intervention and control groups combined in abstinence rates; however, among those who actually had used an intervention, there was a significant overall increase in abstinence (OR 1.95, 95% CI 1.04 -3.67). In a follow-up study conducted by this group of a smartphone app versus simple text messaging, authors found that the app-based intervention was similar to text messaging in terms of utilization; however, the text messaging system was associated with higher abstinence rates [7].

There have been 2 randomized trials of intensive smoking cessation support for Aboriginal and Torres Strait Islander people in program in remote communities (163 participants) and among pregnant women (263 participants) [8,9]. Although both trials did not demonstrate significant improvement in quit rates, a meta-analysis of data from both trials did show a 2.4-fold increase in quit rates. A systematic review of clinical trials internationally to reduce smoking rates in indigenous populations, conducted in 2013, analyzed 5 studies—3 testing Quitline protocols (structured steps for what happens when a person calls the Quitline including triage, counseling, and follow-up processes) combined with cessation products compared with Quitline alone and 2 using culturally adapted cessation counseling using mobile phones [10]. Outcomes were mixed; however, a New Zealand trial of a text message intervention involving 1705 participants including 355 Maori (the indigenous people of New Zealand) found a greater than 2-fold increase in smoking abstinence at 6 weeks (28% vs 13%), which appeared to be sustained at 6 months [11]. Abstinence rates were equally successful among Maori as non-Maori in the intervention arm [12]. This program is now being administered via New Zealand Quitline.

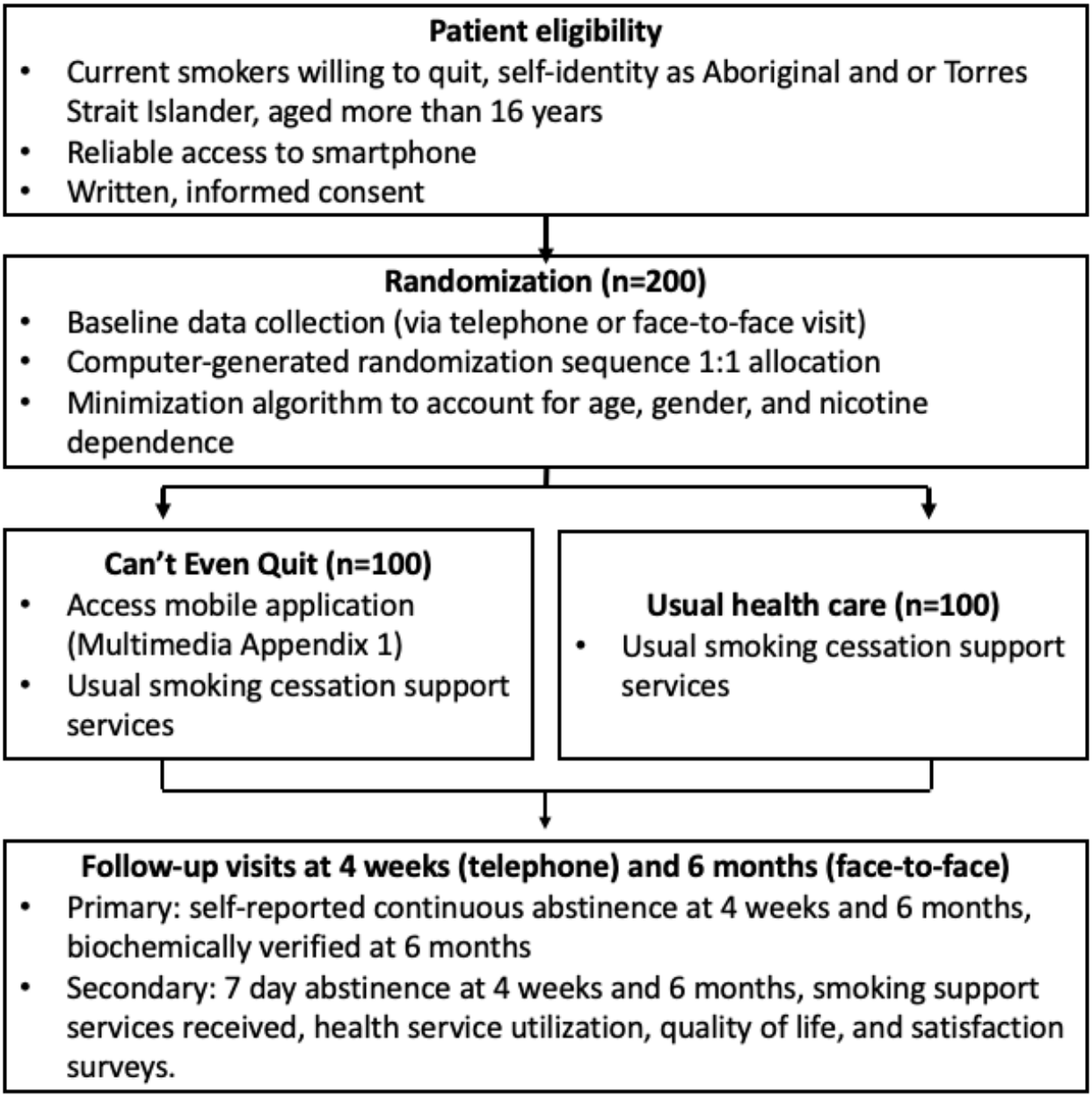
Contemporary data on mobile phone access for Aboriginal and Torres Strait Islander people were unclear; however, in 2008, 67% of nonremote and 61% of remote households had access to a prepaid mobile phone and 41% and 19%, respectively, had mobile access via a contract [13]. These numbers are almost certainly much higher now. Despite high rates of mobile phone usage, we are not aware of any trials of smoking cessation apps specifically built for indigenous people. Therefore, the aim of this study was to assess the feasibility, acceptability, and preliminary effectiveness of a novel mHealth app to assist Aboriginal people to quit smoking.

Methods

Design

The pilot study was an RCT (Figure 1). Outcome data collection and analysis were conducted blinded to treatment allocation. The Consolidated Standards of Reporting Trials-EHEALTH statement is available in Multimedia Appendix 1 and the study protocol in Multimedia Appendix 2.

Figure 1. Trial schema.



Participant Eligibility Criteria

Participants were eligible if they could provide informed consent and met all of the following criteria: (1) current smokers aged 16 years or older, (2) self-identification as an Aboriginal and/or Torres Strait Islander person, (3) willing to make an attempt to quit smoking in the next month, and (4) had access to an iPhone or Android smartphone. Only 1 person per household was invited to participate in the study.

Recruitment

Recruitment was intended to take place via up to 4 Aboriginal Community Controlled Health Services (ACCHSs), a New South Wales (NSW) government telephone coaching service, and Quitline. Participants were recruited via community events, ad hoc referrals from health professionals, and direct phone calls by research staff based on recommendations made by the ACCHS.

Allocation

Randomization was conducted via a central computer-based randomization service. Allocation was 1:1 intervention versus control using a minimization algorithm to balance for sex, age (<30 years vs ≥30 years), and heaviness of smoking index score (low [score≤2] vs moderate or high addiction [score>2]) for nicotine dependence [14]. Outcome analysis and data collection were conducted blinded to treatment allocation.

Intervention

A detailed description of the intervention development, the codesign process, and core features is available in Multimedia Appendix 3. In brief, the intervention was delivered via a multifaceted Android or iOS app comprising a personalized profile and quit plan, text and in-app motivational messages, and a challenge feature allowing users to "compete" with others. Screenshots are shown in Figures 2-5.

All participants were free to use any other smoking cessation service or support and were offered Quitline and local ACCHS contact numbers. Participants allocated to the intervention group were given the opportunity to download and start using the app immediately. There was no cost associated with using the app. A support worker facilitated the registration profile, password set-up and the creation of a tailored quit plan. Participants were then given a tutorial of the app features including how to set up a challenge, motivational support, and tracking progress on their plan. App set up was also done without the use of a support worker as desired.

Figure 2. The challenge function screenshots.

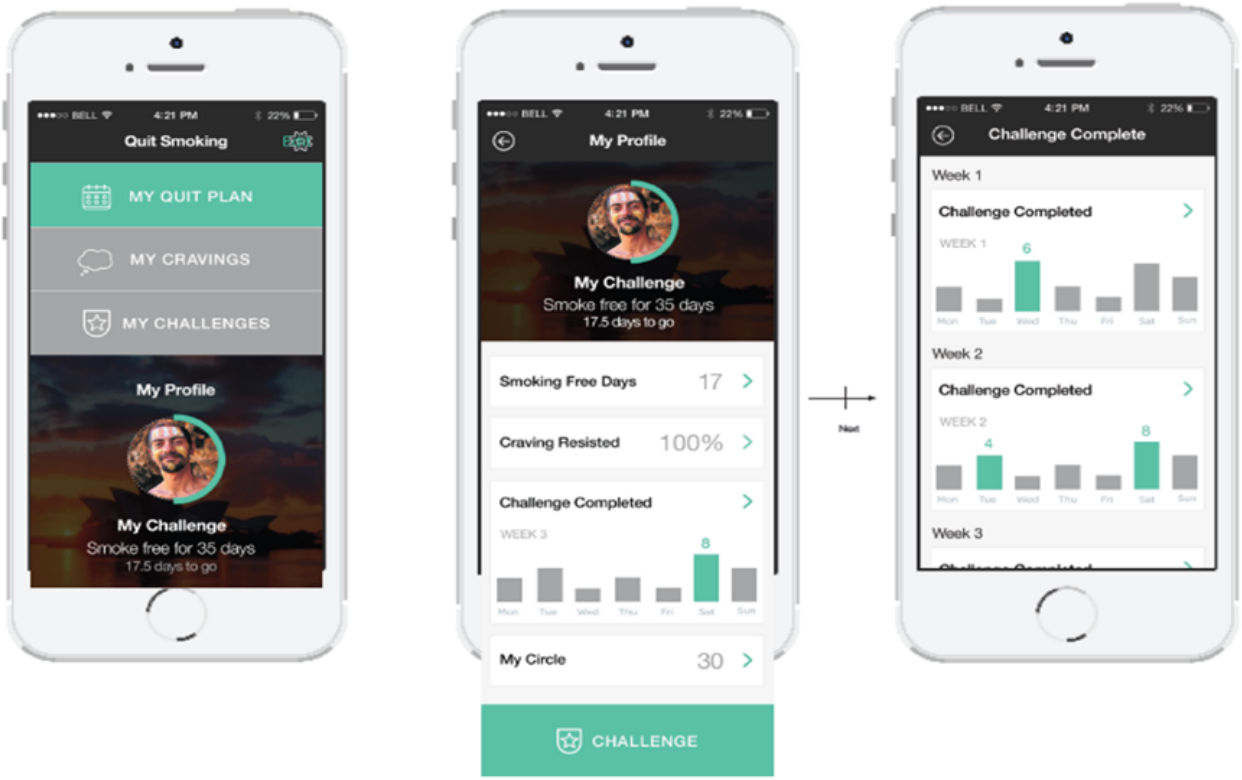


Figure 3. My quit plan screenshots.

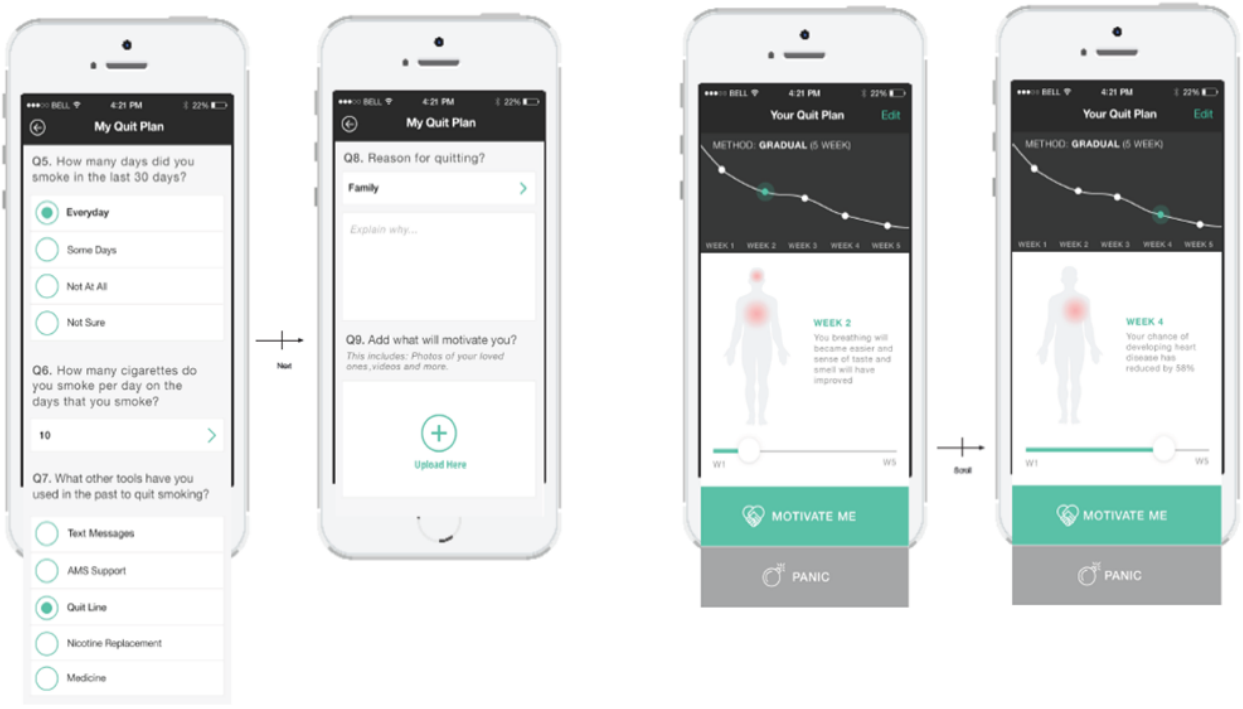


Figure 4. "Motivate me" screenshots.

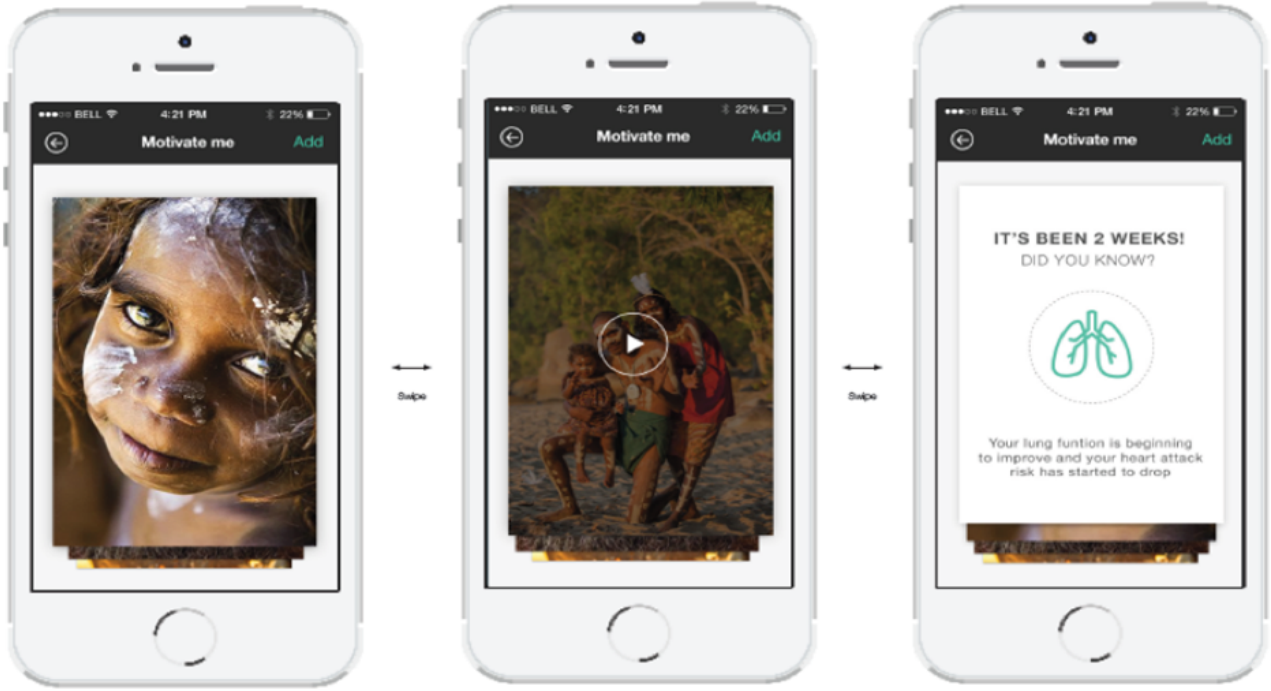
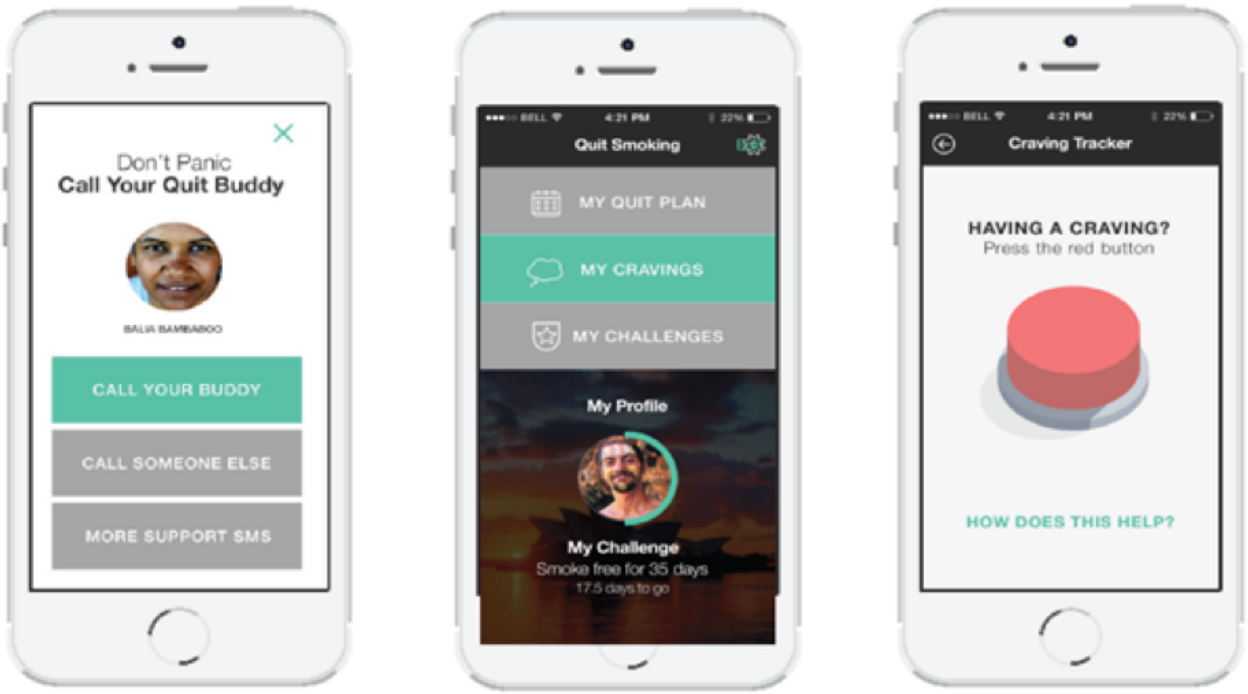


Figure 5. Manage cravings screenshots.



Control

Participants in the control arm and their supporting health services were encouraged to make use of all smoking cessation support services available to them. They were barred from accessing the app if their mobile number matched our database of control arm participants. People not involved in the trial could also download the app if they were invited by an intervention arm participant to use the "challenge" function.

Data Collection Methods

All study instruments were administered by a questionnaire and an interview. Baseline data collection, consent, and randomization were conducted by telephone or via an in-person visit with a trained project officer. Participants allocated to the intervention were telephoned at 1 week to determine if they needed any support with using the app. There was also a "contact us" number in the app itself to call for help. At 4 weeks, an "evaluation" project officer conducted an outcome assessment

by telephone or in-person for all participants in the trial. At 6 months' post randomization, this officer conducted the final outcome assessment and biochemical verification of smoking status (if needed) as part of a face-to-face visit (Figure 1). In addition to the objective outcome assessment visits, an analytic tool built in the app assessed usage patterns and changes in usage patterns over time.

Quantitative Outcomes

The primary outcome was self-reported continuous smoking abstinence, objectively verified at 6 months. Self-reported continuous abstinence is defined as the participant reporting they had quit at both 4 weeks and 6 months of follow-up. Self-reported smoking cessation was confirmed with an Airmet Scientific Micro Plus Smokerlyzer (carbon monoxide meter breath test where a reading of less than or equal to 6 represents no recent tobacco smoking) [15]. Secondary outcomes included point prevalence of abstinence and use of smoking cessation therapies and services.

Statistical Considerations

On the basis of the Text2Quit and Text2Stop trials, we estimated the control arm abstinence rate at 6 months to be 5% [11,16]. Assuming that the abstinence rate is double that of the control arm, we calculated that 1000 participants would be needed to detect a 5% absolute difference (ie, 10.0% vs 5.0%; relative risk 2.00). This assumes a 15% loss to follow-up, 2-sided alpha=.05 and 80% probability of detecting a significant difference with *P* values less than .05 judged as significant. This study was planned as a pilot RCT in which we aimed to recruit 200 participants to assess feasibility, resource considerations, acceptability, and preliminary effectiveness data to inform a future adequately powered trial.

Using the secondary outcome measure of 7-day abstinence rates at 4 weeks and assuming the control arm abstinence rate is 12% and the intervention arm abstinence rate is 2.4 times greater (as observed in the Text2Stop trial), we assumed there would be 80% probability of detecting a significant difference with 200 participants from the pilot study under the same assumptions as above.

Data were analyzed on the basis of intention-to-treat. Characteristics were compared between the groups at 4 weeks and 6-month follow-up using exact Chi-square tests for categorical variables with 2-sided *P* values, where the assumptions of the statistical test could be met. Results were reported in terms of relative risks with exact 95% CI. Analyses were undertaken using SAS 9.4 (SAS Institute, Cary, NC). Given this was a pilot study, no subgroup analyses were prespecified. No imputation techniques were undertaken for missing data.

Qualitative Evaluation

Semistructured, in-depth interviews with intervention arm participants were conducted to better understand barriers and enablers to smoking cessation in general and to use of the app. Participants were sampled to include interviewees with a range of app engagement levels and a balance of gender and ages. Insights from the Talking About the Smokes (TATS) national survey of attitudes to smoking, quit patterns, and management received were drawn on to guide the interview schedule (Multimedia Appendix 4) [17]. Interviews were conducted both over the phone and in-person and were digitally recorded and professionally transcribed. The first 5 interviews were analyzed by 3 team members, and a preliminary coding framework was derived by consensus. The remaining interviews were then analyzed according to this framework by 2 team members. The thematic framework was iteratively refined as new data and themes became apparent. Interviews continued until no new major themes emerged.

Governance and Ethics

An executive committee comprising the NSW Ministry of Health (Centre for Population Health and the Centre for Aboriginal Health), Aboriginal Health and Medical Research Council, and The George Institute for Global Health oversaw the project delivery. The Cancer Institute provided ad hoc advice and reviewed progress throughout the project. An expert user group was convened early in the project and regularly consulted during the intervention development phase (Multimedia Appendix 2). Ethical approval was granted by the Aboriginal Health and Medical Research Council Human Research Ethics Committee. Formal approvals from each of the participating sites were also granted. Due to an administrative oversight, the trial was inadvertently registered retrospectively after the first participant was enrolled in the trial.

Results

Recruitment

We were only able to collaborate with 1 ACCHS, 1 regional community event, and the NSW government coaching service to assist in recruitment. Furthermore, 3 ACCHSs decided against participation. The major barriers to participation included the following: (1) competing service delivery priorities, (2) lack of resources for research, and (3) preference that the intervention be made available to all interested clients of the service rather than random allocation to intervention and control arms. Due to a change in the contracted service provider for Quitline during the trial, we were also unable to negotiate an agreement with the new provider within the project timelines.

Due to these recruitment challenges, only 49 of the target 200 participants were recruited between March 2016 and January 2017, with the majority recruited from 1 ACCHS (Figure 6). There were no changes made to the trial outcomes after the trial commenced.

Table 1 shows the baseline characteristics of the trial participants by randomization arm. Broadly, both groups were similar. The majority of participants in both arms had made at least one quit attempt in the previous 12 months and almost half had tried to quit in recent weeks (23/49, 47%). The majority of participants (39/49, 80%) smoked their first cigarette of the day within 30 min of waking. Most participants (44/49, 90%) commenced smoking before the age of 20 years. Most people lived in households where at least one other person smoked (33/49, 67%) A minority had made use of support services, medication, or NRT. Most participants had seen a doctor in the previous 12

months. In terms of use of mobile technology, most participants used their mobile phone for more than 30 min a day for voice calls (29/49, 60%), text messages (31/49, 64%), smartphone apps (35/49, 72%), internet (37/49, 75%), and social media (41/49, 83%).

Figure 6. Study flow.

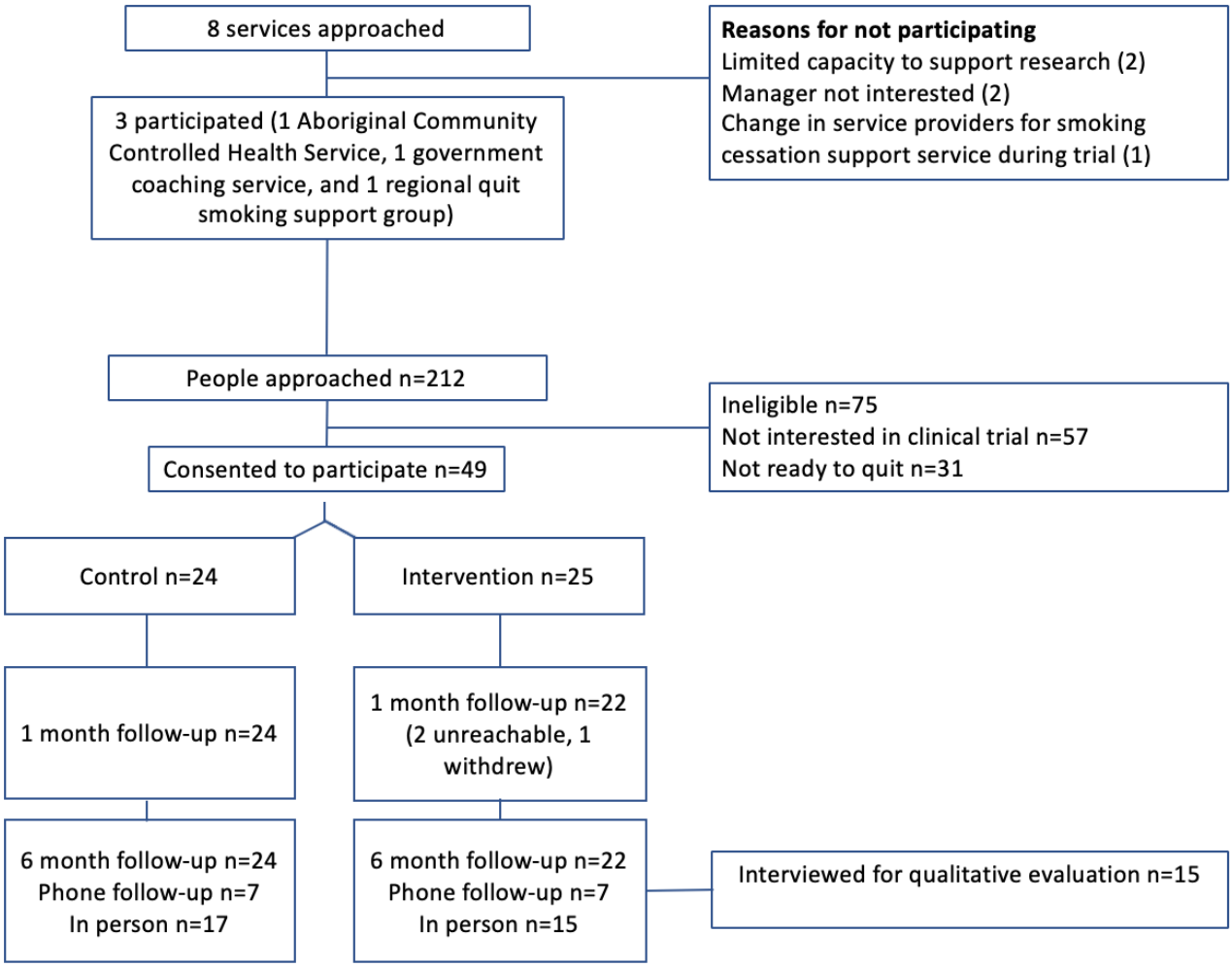


Table 1. Participant baseline characteristics.

Characteristic	Intervention (n=25)	Control (n=24)	Total (n=49)
Female, n (%)	19 (76)	19 (79)	38 (78)
Male, n (%)	6 (24)	5 (21)	11 (22)
Aboriginal, n (%)	25 (100)	24 (100)	49 (100)
Age (years), mean (SD)	42 (14)	42 (14)	42 (14)
Smoke every day (or nearly every day), n (%)	25 (100)	23 (96)	48 (98)
Number of smokes per day, n (%)			
10 or less	6 (24)	8 (33)	14 (29)
11-20	15 (60)	9 (38)	24 (49)
21-30	3 (12)	6 (25)	9 (18)
31 or more	0 (0)	1 (4)	1 (2)
Level of addiction[a], n (%)	0 (0)	1 (4)	1 (2)
Low	8 (32)	8 (33)	16 (33)
Moderate	15 (60)	15 (63)	27 (55)
High	1 (4)	1 (4)	5 (10)
Highest level of formal education or employment, n (%)			
Some high school (no certificate)	8 (32)	6 (25)	14 (29)
Completed high school	2 (8)	5 (21)	7 (15)
Technical education certificate	11 (44)	8 (33)	19 (40)
Some university (no degree)	1 (4)	4 (17)	5 (10)
Completed university degree	1 (4)	0 (0)	1 (2)
Postgraduate degree	1 (4)	1 (4)	2 (4)
Paid employment	13 (52)	10 (42)	23 (48)
Current mobile phone used, n (%)			
iPhone	8 (32)	9 (38)	17 (35)
Android	16 (64)	15 (63)	31 (65)
Mobile phone subscription plan, n (%)			
Prepaid	14 (56)	15 (63)	29 (60)
Regular plan	10 (40)	9 (38)	19 (40)
Quit patterns, n (%)			
Number of people who had a previous quit attempt in the last 12 months	21 (84)	21 (88)	42 (88)
Use of any type of nicotine replacement therapy or other stop-smoking medications in the last year	5 (20)	10 (42)	15 (31)
Sought any information or quit services in the last 12 months	1 (4)	5 (21)	6 (13)
Previously used Quitline	0 (0)	0 (0)	0 (0)
Local quit program	1 (4)	0 (0)	1 (2)
Seen a health worker, doctor, nurse, or other health professional in the last year, n (%)			
Doctor	18 (75)	21 (88)	39 (82)
Nurse	5 (20)	6 (25)	11 (23)
Aboriginal health worker	3 (12)	3 (13)	6 (13)
Tobacco worker	1 (4)	1 (4)	2 (4)

[a]Using the heaviness of smoking index where a score of 0-2 indicates low addiction, 3-4 moderate addiction, and 5-6 high addiction.

Trial Outcomes

Table 2 shows the primary outcomes of the trial for the 46 people who completed the study. Only 2 participants (intervention arm) reported abstinence at either the 4-week or 6-month interview, and nobody had continuous abstinence at both time points. There were generally higher numbers of intervention arm participants reporting the use of supportive cessation services, but these were not statistically significant. There were no harms or unintended consequences observed in this study.

App Usage Analytics

The usage data for the app from March 2016 to April 2017 are shown in Figure 7. These data include both intervention arm participants and nontrial participants who were invited to use the app as part of the challenge function (81 new users in total for the trial period). From June to December 2016, there was a relatively consistent rate of usage in terms of new users (3-10 users per month) and returning users (4-8 users per month). Session frequency average was 2.9 times per user per month, average session time was 6.3 min, and the average number of unique screens viewed per session was 7.6. These statistics indicate a low to moderate level of usage through the trial period.

Table 2. Trial outcomes.

Outcome[a]	Intervention (n=22)	Control (n=24)
Self-reported smoking abstinence, n (%)		
At 4-week visit	1 (4.5)	0 (0)
At 6-month visit (verified with breath test)	1 (4.5)	0 (0)
At both 4-week and 6-month visit	0 (0)	0 (0)
Self-reported quit attempt since starting the trial	5 (23)	6 (25)
Use of other smoking cessation services during the trial, n (%)		
Any nicotine replacement therapy use during study	5 (23)	5 (21)
Any smoking-related medication use	4 (18)	2 (8)
Quitline or any other quit program	2 (9)	0 (0)
Any combination of the above cessation aids	9 (41)	7 (29)

[a]The total number of participants who achieved 1 of the main or secondary outcomes was small, so we did not perform statistical testing or calculate a relative risk ratio as we could not meet the assumptions required.

Figure 7. App usage by month (March 2016-April 2017).

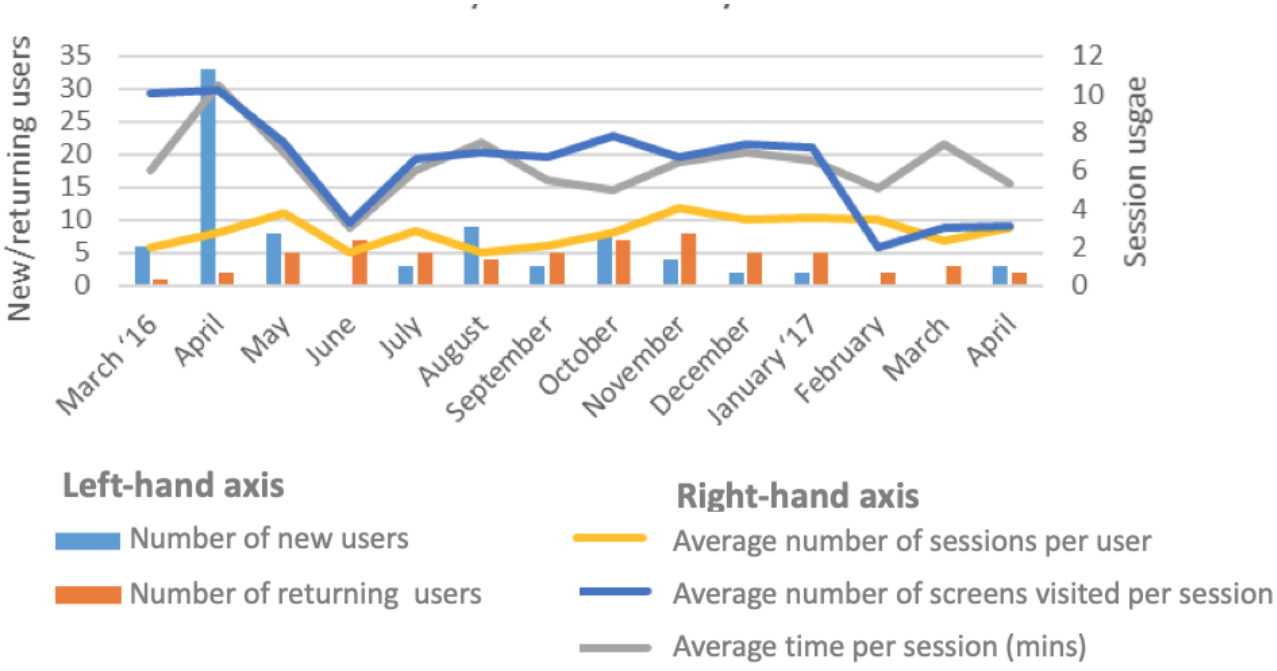
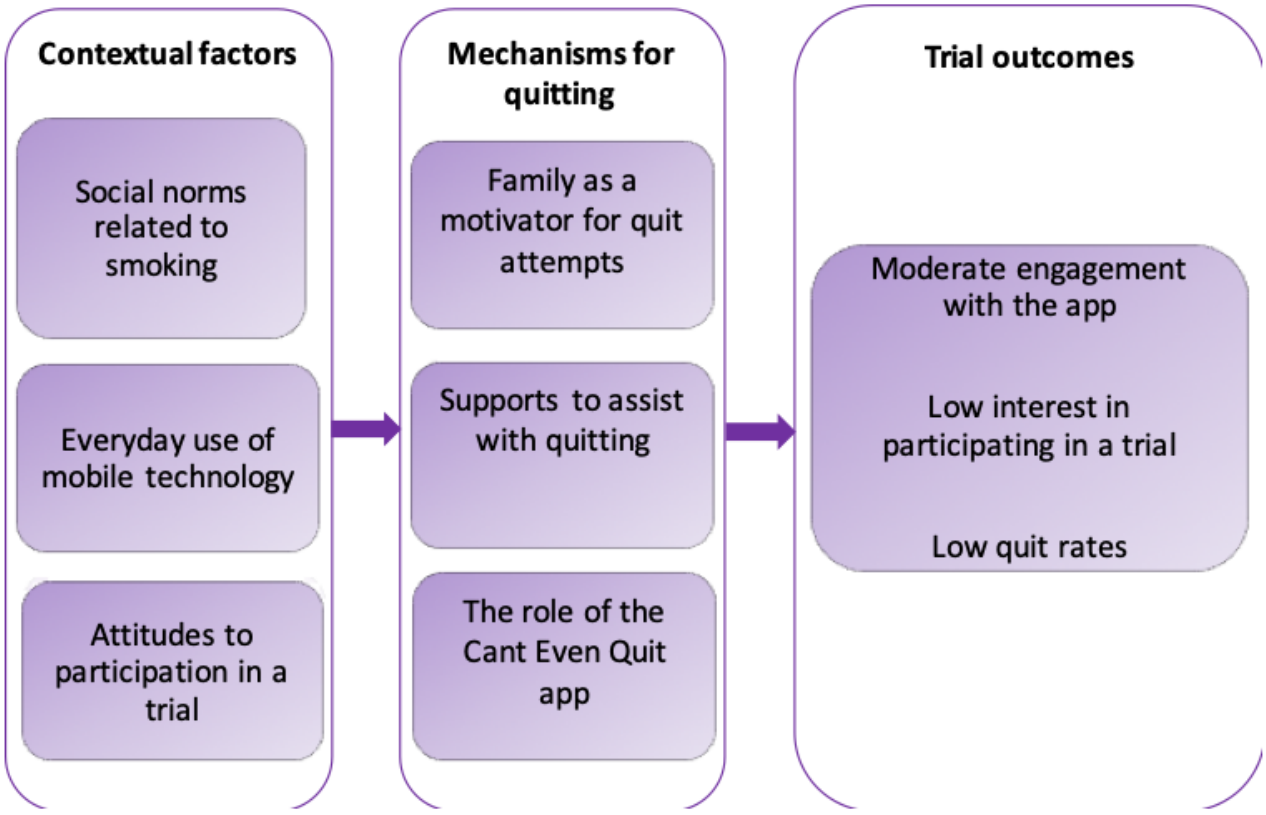


Figure 8. Qualitative themes and their influence on the trial outcomes.



Interview Findings

A total of 15 interviews with intervention arm participants were conducted, ranging from 20 to 45 min duration. Furthermore, 6 major themes arose from the analyses of these interviews. These are illustrated in Figure 8 using a context, mechanism, and outcome configuration to demonstrate the potential influence of these themes on smoking cessation and to assist with explaining the outcomes observed in the trial. More descriptive detail on the context- and mechanism-related themes is provided below.

Contextual Factors Influencing the Trial and Outcomes

Theme 1: Social Norms Related to Smoking

All trial participants had commenced smoking before the age of 18 years. Interviewees described living in households where smoking was the norm among adults and how these norms strongly influenced children:

> *You'd see the plumes of smoke in the kitchen...There were ashtrays all over the place...we used to play with the butts and pretend to smoke them ourselves* [female, 61 years]

Another participant described the strong influence of girls' peer groups in driving the commencement of smoking describing how she:

> *Would go home and my friend would give me one out of her mum's packet and it went on from there.* [female, 44 years]

Another participant described being paid in cigarettes for babysitting for his neighbor. The consequence of such entrenched normalizing behaviors was that smoking was viewed as a natural aspect of progression to adulthood with 1 participant commenting:

> *I got my first pay check at 14 and I bought a packet of cigarettes.* [female, 39 years]

Theme 2: Everyday Use of Mobile Technology

Most participants cited phone usage of some kind. Facebook was one of the most popular features with 1 participant describing it:

> *Like the fridge, you're always opening it.* [female, 44 years]

Only 2 participants had used a health-related app (unrelated to the study app). For these participants, there was an initial curiosity to explore the features and its potential to assist with one's health, but this tended to be short-lived:

> *I downloaded it, but then, yeah, it was a lot of rubbish so I took it off.* [female, 48 years]

Most older interviewees did not appear to have any major issues with knowledge on how to access phone features. However, most people rarely made use of smartphone apps apart from those used for playing games. Moreover, 1 participant commented:

> *...I play them (game apps) on the way to work...maybe three hours a day.* [male, 24 years]

Furthermore, 1 interviewee suggested that game apps could provide a stronger motivation for engaging in health apps.

Theme 3: Attitudes to Participation in a Trial

A study design involving a control arm that did not receive the app was not viewed favorably, given this was a cohort of current smokers who had expressed the intention to quit smoking within the next month:

> *I seen it [the trial] promoted on Facebook...and a couple of posters in the waiting area, and I was actually encouraging a lot of patients to come down for it, but I didn't know it... [depended on]... whether you got randomized. I was glad I was in the trial [intervention arm], because I would have been frustrated if I wasn't.* [female, 44 years]

This sentiment was also expressed by several ACCHS managers. Although the decision to participate was influenced by a range of factors, some managers were uncomfortable with both the resource burden of participating in a trial and the lack of ability to provide the intervention as a part of routine service delivery. Many expressed interest in participating in a posttrial phase once the RCT was completed.

Mechanisms for Quitting

Theme 4: Family as a Major Motivator to Quit Attempts

Although social norms were described as strong drivers in initiating and maintaining smoking, similarly many participants cited family as being the major motivator for quitting:

> *So what I want...is when I cuddle my girls they're not saying, "Oh Nanny, you smell like cigarettes"...that breaks my heart. And it makes me more determined...I want them to remember me as their nanny who loved them very much...* [female, 61 years]

Although many participants had used smoking cessation aids, key events such as the death of a relative appeared to be critical moments in supporting quit attempts:

> *Losing my mum two years ago has really, really made me want to quit.* [female, 44 years]

The benefits of being smoke-free were also frequently framed in the context of improving family well-being:

> *I could use that (money from cigarettes) for something better, for another cruise where I can take the girls...if I save $25 every time I buy a smoke, I might have enough to go on a world cruise. It could happen.* [female, 44 years]

Theme 5: Supports to Assist With Quitting

Around one-half of the participants had previously attempted a range of smoking cessation methods to help them quit, in particular, NRT and mFedication. In addition to these support strategies, several participants mentioned the value of a group atmosphere to support smoking cessation:

> *...a support group, like AA...sharing ideas, so you encourage each other a bit more.* [female, 44 years]

Another participant suggested:

> *You'd have each other to lean on and to express what you're feeling.* [female, 39 years]

Some commented on how an app could extend this group support further:

> *...then you could send this person a message, "Mate, I'm having a downer."...And if they're on the other end they can give you some positive feedback.* [female, 48 years]

Although all participants had heard of Quitline and some had been referred to the program, no one reported actually using the telephone service. Some participants expressed problems with the referral process and this led to a loss of confidence in the service:

> *I don't think they rang me...I don't know if they're just tokenistic.* [female, 44 years]

The same participant suggested that the fact that the referral was coming from a third party may have signaled a lack of motivation:

> *Maybe if you don't do it yourself (self-refer), they don't take it as seriously.*

Theme 6: The Role of the Can't Even Quit App

Technical Difficulties

Although most participants found the app intuitive and easy to use, some participants would have benefited from more extensive training. For example, 1 participant was not aware of the ability to alter the quit date in the app settings. Given the app was designed to provide tailored content according to the user's quit status, this caused problems:

> *I stuffed up my start date...to start giving up smoking. And then, it flashed something up, and I'm thinking, "I haven't quit yet, what are you sending me these for?" ...and that messed everything up. And probably just another excuse not to give up.* [female, 44 years]

Several participants indicated interest in the challenge function both for oneself and to challenge others; however, there were a range of technical issues that limited its usage such as halting a current challenge and initiating a new one and inviting new users to engage in a challenge. Other technical barriers related to difficulties downloading the app over a cellular network and an automatic log out requiring the user to sign in again after 3 days of inactivity. For a minority of people, these issues substantially limited use of the app.

An Intrusive Versus Supportive Service

A positive feature of the app was the support provided by trial staff in using the intervention. Several participants appreciated the phone call follow-ups:

> *I'd say keep the ringing up. As much as us smokers dodge people that's because it is a hard subject and we do need that push and encouragement.* [female, 44 years]

Although this level of "intrusiveness" from the trial staff was generally viewed favorably, some interviewees warned of the risk of "overkill." Moreover, 1 participant highlighted the need to strike a balance:

> *I was part of that NSW Health service...Get Healthy. Oh and you'd dodge them like a bullet! They were painful...So once I see your number, I go, "Oh, THEM again!" I think you've got to find the right balance.* [female, 44 years]

This was particularly an issue with the motivational text messages. For some participants, the high message frequency (Multimedia Appendix 2) was viewed positively, and when the frequency reduced, they started to disengage:

> *I was getting texts all the time, reading all of them, trying to take in all the information...They were good...motivating, especially when you're having a hard day and you get texts all the time, and it's like, "Yeah, I can do it." ...In the beginning I was getting about five or six [text messages], and then it just petered off...* [female, 46 years]

Others found the text message content to be initially useful, but that they would start to lose their appeal over time:

> *I started ignoring them. They went into a dark place.* [female, 46 years]

Some suggested the messages needed to be more engaging over time:

> *It was exciting at first, but...I want fun stuff...* [female, 48 years]

Others found the messages came too frequently and this detracted from the personalized nature of these messages:

> *One time I got five in the space of three and a half minutes, and that's when it felt to me it was just automated.* [female, 44 years]

Other Suggestions for Improvement

Many commented on the value of incorporating games into the app:

> *There were no games that you could play...like my jigsaw puzzle on my iPad, there was nothing ...[where] you could go "Oh yeah, I might have to do a little quiz.* [female, 48 years]

Several participants were regular Facebook users and most of these users felt the app could have been more social, interactive, and inclusive of user-generated content:

> *...Bring more people to it...maybe put funny little videos or...more stories where you can link in with people...Have it open.* [female, 48 years]

Addition of an enhanced social element and linkage to support services such as Quitline was seen as essential to harnessing greater interest in the app:

> *...if they (Quitline) can see that you are having a hard time, then maybe you can get a support call from them...Because...there wasn't enough people for me to join to become friends with. So you're sort of isolated on there.* [female, 48 years]

Discussion

Principal Findings

This study outlines the development, implementation, and pilot evaluation of an mHealth strategy to support Aboriginal people with smoking cessation. The evaluation shed valuable information on the challenges of implementing an RCT in this setting and clearly raises serious questions about the feasibility of conducting a larger-scale trial. Despite these challenges, it also shed light on the future potential for mobile apps in this area. We discuss the findings and implications across 3 broad areas: (1) the population, (2) the setting and social context, and (3) this particular app and lessons for other apps in this setting.

The population which participated may represent a "hard-to-reach" group in terms of cessation support services and consequently may differ from the types of populations that have participated in previous mHealth quit smoking trials. A large number of female participants were enrolled in the trial, and analyses from systematic review data suggest that mHealth smoking cessation interventions may be less effective for women than men [5]. Consistent with other studies of quitting among Aboriginal and Torres Strait Islander people, we found that few participants successfully sustained abstinence, used NRT, smoking cessation medications, or cessation services [3,18,19]. These factors suggest that a "one app fits all" approach is unlikely to be successful and a better understanding of the needs and opportunities for groups that have difficulty sustaining quit attempts is needed. A recent international critical review of the literature relating to Indigenous people's use of health technologies found that such technologies required meaningful user involvement, community-based processes for development, and creative adaptation to local needs [20].

The importance of the setting and social context in which people smoke was a prominent finding from the interviews. Several interviewees commented that social norms around smoking are key drivers of smoking rates in the community. It has been suggested that changes in these norms have occurred more slowly in Aboriginal and Torres Strait Islander communities, and this may be a key factor in the differential decline in smoking rates compared with the general population. Our interviews highlighted many examples of social influences where participants were exposed to smoking behavior early in life and the role this may have played in initiation and maintenance of smoking [21]. A social network analysis study of Aboriginal people in the Australian Capital Territory found separate clustering of smokers from nonsmokers in the social network and that the proportion of adults in a person's household who smoked was associated with being a smoker [22].

Despite historically high levels of social acceptability for smoking in Aboriginal communities, this is clearly changing. Some have suggested that social contagion of behaviors such as quitting rely on changing norms [23]. The TATS study found that attitudes and beliefs positively associated with wanting to quit included regretting ever starting to smoke, perceiving that local Aboriginal and Torres Strait Islander community leaders disapprove of smoking, believing nonsmokers set a good example for children, worrying about future smoking-related

health effects, and believing quitting to be beneficial [3]. We also found in the interviews that although families can influence initiating smoking, they can also be a highly motivating influence to quit smoking. Several participants interviewed commented on how key family members viewed smoking as undesirable, and this was a strong driver for making quit attempts. Others said they would like group support and a group atmosphere to support them in their smoking cessation journeys.

The above 2 findings related to population and social context have implications for future app development considerations. Although most participants cited phone usage of some kind (mainly texting, phone calls, and Facebook), smartphone apps were not widely used, and only 2 participants had used a health-related app previously. The usage analytics showed modest but sustained usage over the life of the study. Many participants, particularly female, recommended that the Can't Even Quit app needed to be socially "opened up" like Facebook so that more stories could be shared. Aboriginal and Torres Strait Islander people may use Facebook at higher rates than the general Australian population, and a recent qualitative study highlighted potential opportunities for leveraging Aboriginal and Torres Strait Islander people's use of social media departing from traditional one-way information models of health promotion to assets-based, self-empowerment models [24]. If smoking cessation apps could access a person's social network, they may be able to facilitate support from network members and spread an intervention through these networks [25]. Currently, most health apps that offer Facebook integration are related to exercise tracking and calorie counting. A 2014 review found only 9 Facebook-integrated smoking cessation apps and these were of variable quality [25]. Pediatric researchers have started to explore Facebook communities for children with chronic diseases and proposed 3 considerations for how this platform could be better leveraged for health issues [26]. These include the following: (1) "back end" content management systems mediated by trusted health professionals that can be integrated with Facebook, (2) tools that can extract data from Facebook feeds and providing structured threads and forums to promote topic-specific information exchange between users, shifting dialogue from "one to many" to "many to many"; and (3) using the platform for consultation and codesign processes when developing health integrated apps.

Clearly, social media apps are only 1 strategy for strengthening social networks related to health issues. A review of the Australian government's Tackling Indigenous Smoking and Health Lifestyle program highlighted the need for multilevel approaches [27]. In addition to social media and community-based campaigns, health service strategies that are locally tailored, maximize community participation, and strengthen workforce capacity are needed. In this study, the broad appreciation for regular support from research staff suggests that the model of "prescribing" the app as part of a coordinated smoking support service package may be attractive. Given the self-reported rates of seeing health professionals were high (88%) in this study and consistent with previous literature, strengthening health service–mediated interventions should be pursued further [18]. Although we had intended to better understand the role of the app as part of the Quitline service, our inability to recruit participants via this means precluded being able to address this. Questions remain regarding the optimal level of support that is needed with such services. A key finding from the interviews was the need to balance intrusiveness versus supportiveness, particularly for text messages. The evidence on the optimal dose of message frequency is unclear, and further research to test different intensities are needed to help answer this question.

Limitations

Although we were not expecting to show statistically significant changes in the primary outcome of this pilot, we did intend to generate sufficient preliminary data to inform a large-scale, well-powered trial. Unfortunately, the study fell short of the target numbers for recruitment. On the basis of previous trials, we calculated that over 1000 participants would be needed for an adequately powered study. The overall abstinence rates at 4 weeks and 6 months in this study were lower than the hypothesized 5% quit rate estimated from the literature. This indicates that quit rates may be lower than assumed, and therefore, an even greater sample size may be needed for a fully powered RCT. Our difficulties in achieving recruitment targets were multifactorial, and barriers occurred at both institutional and individual participant levels. More effort will also be needed to attract higher numbers of men into future research of this nature, as many of the findings presented here may reflect gender-specific views. Although we had success implementing the trial at 1 ACCHS, this was not true for several other sites that we approached. For ACCHSs, competing interests, the lack of dedicated resources to support a research study, and a perceived challenge in "selling" a trial where only half of the participants would be given the intervention were the major barriers [28]. Indeed, many ACCHS managers indicated a strong interest in accessing a program if it were made available in a nontrial setting. Alternative evaluation and alternative designs that maintain methodological rigor, enhance institutional and participant acceptability, and place a limited resource burden on health services are needed. Potential options include testing multiple interventions using a factorial RCT design as was done in the Victorian RCT [6], embedding periodic outcome assessments into the app itself and constructing matched control groups from routinely collected datasets.

Conclusions

This pilot of a smoking cessation app for Aboriginal people was difficult to implement, and few conclusions can be made regarding its effectiveness. Despite this, many insights were gained from the evaluation that bear consideration for future research in this space. In relation to health apps, the findings underscore the importance of a socio-technical approach to their development, comprehensively understanding and embracing the complex interaction between people, their social contexts, and rapidly changing technologies. Although maximal benefit would come from working with people that experience the greatest difficulty with sustaining a quit attempt, it is important to recognize that this is a heterogeneous group that may have particular needs that are not being addressed through existing interventions. Finally, to enhance adoption of such apps, we suggest that a 2-pronged approach is needed that focusses on

both service delivery models that promote app "prescription" as part of an overall comprehensive smoking cessation support service and health promotion strategies that support community empowerment.

Acknowledgments

This research is a project of the George Institute for Global Health and was commissioned by NSW Health. The information and opinions contained in it do not necessarily reflect the views or policy of NSW Health. The authors would like to thank all the communities and organizations participating in the Can't Even Quit project. In particular, the authors extend gratitude to those individuals that participated in the expert user groups, the pilot trial, and those participants that generously gave time for interviews. The authors also acknowledge the support provided by staff from the Centre for Population Health, Centre for Aboriginal Health, Tharawal Aboriginal Corporation, Bullinah Aboriginal Health Service, Get Healthy NSW, Cancer Institute NSW, and Cancer Council of Victoria.

Authors' Contributions

DP and LW conceptualized the project. DT provided advice on study design throughout the project. LW and MN were responsible for participant recruitment, data collection, and analysis. DP and KR performed the statistical analyses. KN contributed to qualitative data analysis and preparation of the report. All authors critically reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 519KB - mhealth_v7i4e12745_app1.pdf]

Multimedia Appendix 2

Study protocol.

[PDF File (Adobe PDF File), 1MB - mhealth_v7i4e12745_app2.pdf]

Multimedia Appendix 3

Development of the intervention.

[PDF File (Adobe PDF File), 237KB - mhealth_v7i4e12745_app3.pdf]

Multimedia Appendix 4

Participant interview guide.

[PDF File (Adobe PDF File), 55KB - mhealth_v7i4e12745_app4.pdf]

References

1. Australian Bureau of Statistics. 2016. 4714.0 - National Aboriginal and Torres Strait Islander Social Survey, 2014-15 URL: https://www.abs.gov.au/ausstats/abs@.nsf/mf/4714.0 [accessed 2019-03-01] [WebCite Cache ID 76XqiZabv]
2. Theo V, Bridget B, Lucy S, Alan D L. The Burden of Disease and Injury in Aboriginal and Torres Strait Islander Peoples. Brisbane: School of Population Health, The University of Queensland; 2007:2003.
3. Nicholson AK, Borland R, Davey ME, Stevens M, Thomas DP. Past quit attempts in a national sample of Aboriginal and Torres Strait Islander smokers. Med J Aust 2015 Jun 01;202(10):S20-S25. [doi: 10.5694/mja15.00202] [Medline: 26017251]
4. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2012;11:CD006611. [doi: 10.1002/14651858.CD006611.pub3] [Medline: 23152238]
5. Scott-Sheldon LA, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text messaging-based interventions for smoking cessation: a systematic review and meta-analysis. JMIR Mhealth Uhealth 2016;4(2):e49 [FREE Full text] [doi: 10.2196/mhealth.5436] [Medline: 27207211]
6. Borland R, Balmford J, Benda P. Population-level effects of automated smoking cessation help programs: a randomized controlled trial. Addiction 2013 Mar;108(3):618-628. [doi: 10.1111/j.1360-0443.2012.04091.x] [Medline: 22994457]

7. Buller DB, Borland R, Bettinghaus EP, Shane JH, Zimmerman DE. Randomized trial of a smartphone mobile application compared to text messaging to support smoking cessation. Telemed J E Health 2014 Mar;20(3):206-214 [FREE Full text] [doi: 10.1089/tmj.2013.0169] [Medline: 24350804]
8. Marley JV, Atkinson D, Kitaura T, Nelson C, Gray D, Metcalf S, et al. The Be Our Ally Beat Smoking (BOABS) study, a randomised controlled trial of an intensive smoking cessation intervention in a remote aboriginal Australian health care setting. BMC Public Health 2014 Jan 13;14:32 [FREE Full text] [doi: 10.1186/1471-2458-14-32] [Medline: 24418597]
9. Eades SJ, Sanson-Fisher RW, Panaretto K. An intensive smoking intervention for pregnant Aboriginal and Torres Strait Islander women: a randomised controlled trial. Med J Aust 2013 Jan 21;198(1):23-24. [Medline: 23330759]
10. Johnston V, Westphal DW, Glover M, Thomas DP, Segan C, Walker N. Reducing smoking among indigenous populations: new evidence from a review of trials. Nicotine Tob Res 2013 Aug;15(8):1329-1338. [doi: 10.1093/ntr/ntt022] [Medline: 23519776]
11. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin R, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]
12. Bramley D, Riddell T, Whittaker R, Corbett T, Lin RB, Wills M, et al. Smoking cessation using mobile phone text messaging is as effective in Maori as non-Maori. N Z Med J 2005 Jun 03;118(1216):U1494. [Medline: 15937529]
13. Australian Bureau of Statistics. Canberra: ABS; 2010. National Aboriginal Torres Strait Islander Social Survey 2008 URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4714.0Main+Features12008 [accessed 2019-03-01] [WebCite Cache ID 76XshiApO]
14. Borland R, Yong H, O'Connor RJ, Hyland A, Thompson ME. The reliability and predictive validity of the Heaviness of Smoking Index and its two components: findings from the International Tobacco Control Four Country study. Nicotine Tob Res 2010 Oct;12 Suppl:S45-S50 [FREE Full text] [doi: 10.1093/ntr/ntq038] [Medline: 20889480]
15. Maclaren DJ, Conigrave KM, Robertson JA, Ivers RG, Eades S, Clough AR. Using breath carbon monoxide to validate self-reported tobacco smoking in remote Australian Indigenous communities. Popul Health Metr 2010 Feb 20;8(1):2 [FREE Full text] [doi: 10.1186/1478-7954-8-2] [Medline: 20170528]
16. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]
17. Couzos S, Nicholson AK, Hunt JM, Davey ME, May JK, Bennet PT, et al. Talking About The Smokes: a large-scale, community-based participatory research project. Med J Aust 2015 Jun 01;202(10):S13-S19. [Medline: 26017250]
18. Thomas DP, Briggs VL, Couzos S, Panaretto KS, van der Sterren AE, Stevens M, et al. Use of nicotine replacement therapy and stop-smoking medicines in a national sample of Aboriginal and Torres Strait Islander smokers and ex-smokers. Med J Aust 2015 Jun 01;202(10):S78-S84. [doi: 10.5694/mja15.00205] [Medline: 26017263]
19. Thomas DP, Bennet PT, Briggs VL, Couzos S, Hunt JM, Panaretto KS, et al. Smoking cessation advice and non-pharmacological support in a national sample of Aboriginal and Torres Strait Islander smokers and ex-smokers. Med J Aust 2015 Jun 01;202(10):S73-S77. [doi: 10.5694/mja15.00293] [Medline: 26017262]
20. Jones L, Jacklin K, O'Connell ME. Development and use of health-related technologies in indigenous communities: critical review. J Med Internet Res 2017 Dec 20;19(7):e256 [FREE Full text] [doi: 10.2196/jmir.7520] [Medline: 28729237]
21. van der Sterren A, Greenhalgh E, Knoche D, Winstanley MH. The Cancer Council. Melbourne: Cancer Council Victoria; 2016. 8.1 Aboriginal peoples and Torres Strait Islanders: social disadvantage, health and smoking—an overview URL: http://www.tobaccoinaustralia.org.au/chapter-8-aptsi/8-1-overview [accessed 2019-01-24] [WebCite Cache ID 75eaymwhH]
22. van der Sterren A, Greenhalgh E, Knoche D, Winstanley MH. The Cancer Council. Melbourne: Cancer Council Victoria; 2016. 8.9 Attitudes to and beliefs about smoking among Aboriginal peoples and Torres Strait Islanders URL: http://www.tobaccoinaustralia.org.au/chapter-8-aptsi/8-9-attitudes-to-and-beliefs-about-smoking [accessed 2019-01-24] [WebCite Cache ID 75eb66gcb]
23. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med 2008 May 22;358(21):2249-2258 [FREE Full text] [doi: 10.1056/NEJMsa0706154] [Medline: 18499567]
24. Hefler M, Kerrigan V, Henryks J, Freeman B, Thomas DP. Social media and health information sharing among Australian Indigenous people. Health Promot Int 2018 Apr 17:- Epub ahead of print(forthcoming). [doi: 10.1093/heapro/day018] [Medline: 29672684]
25. Jacobs MA, Cobb CO, Abroms L, Graham AL. Facebook apps for smoking cessation: a review of content and adherence to evidence-based guidelines. J Med Internet Res 2014;16(9):e205 [FREE Full text] [doi: 10.2196/jmir.3491] [Medline: 25205129]
26. Mogul DB, Nagy PG, Bridges JFP. Building stronger online communities through the creation of Facebook-integrated health applications. JAMA Pediatr 2017 Oct 01;171(10):933-934. [doi: 10.1001/jamapediatrics.2017.2300] [Medline: 28846765]
27. Upton P, Davey R, Evans M, Mikhailovich K, Simpson L, Hacklin D. The Department of Health. Tackling Indigenous Smoking and Healthy Lifestyle Programme Review: Executive Summary URL: http://www.health.gov.au/internet/main/

publishing.nsf/Content/904B8752C99678A1CA257EA00026976F/$File/TIS-and-Healthy-lifestyle-Programme_Executive%20Summary.pdf [accessed 2019-02-28] [WebCite Cache ID 73rMyUTGn]
28. Liu H, Laba T, Massi L, Jan S, Usherwood T, Patel A, et al. Facilitators and barriers to implementation of a pragmatic clinical trial in Aboriginal health services. Med J Aust 2015 Jul 06;203(1):24-27. [doi: 10.5694/mja14.00581] [Medline: 26126563]

Abbreviations

ACCHS: Aboriginal Community Controlled Health Service
mHealth: mobile health
NRT: nicotine replacement therapy
NSW: New South Wales
RCT: randomized controlled trial
TATS: Talking About the Smokes

Edited by G Eysenbach; submitted 06.12.18; peer-reviewed by N Walker, M Dyson, E Isaura; comments to author 16.01.19; revised version received 30.01.19; accepted 30.01.19; published 02.04.19.

Please cite as:
Peiris D, Wright L, News M, Rogers K, Redfern J, Chow C, Thomas D
A Smartphone App to Assist Smoking Cessation Among Aboriginal Australians: Findings From a Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(4):e12745
URL: https://mhealth.jmir.org/2019/4/e12745/
doi:10.2196/12745
PMID:30938691

Original Paper

Evaluating the Impact of the HeartHab App on Motivation, Physical Activity, Quality of Life, and Risk Factors of Coronary Artery Disease Patients: Multidisciplinary Crossover Study

Supraja Sankaran[1], MSc, PhD; Paul Dendale[2,3], MD, PhD; Karin Coninx[1], MSc, PhD

[1]Expertise Center for Digital Media, Hasselt University, Diepenbeek, Belgium

[2]Department of Cardiology, Heart Center, Jessa Hospital, Hasselt, Belgium

[3]Faculty of Medicine and Life Sciences, Hasselt University, Diepenbeek, Belgium

Corresponding Author:
Supraja Sankaran, MSc, PhD
Expertise Center for Digital Media
Hasselt University
Wetenschapspark 2
Diepenbeek, 3590
Belgium
Phone: 32 011 26 84 11
Email: supraja.sankaran@uhasselt.be

Abstract

Background: Telerehabilitation approaches have been successful in supporting coronary artery disease (CAD) patients to rehabilitate at home after hospital-based rehabilitation. However, on completing a telerehabilitation program, the effects are not sustained beyond the intervention period because of the lack of lifestyle adaptations. Furthermore, decline in patients' motivation lead to recurrence of disease and increased rehospitalization rates. We developed HeartHab, using persuasive design principles and personalization, to enable sustenance of rehabilitation effects beyond the intervention period. HeartHab promotes patients' understanding, motivates them to reach personalized rehabilitation goals, and helps to maintain positive lifestyle adaptations during telerehabilitation.

Objective: This study aimed to investigate the impact of the HeartHab app on patients' overall motivation, increasing physical activities, reaching exercise targets, quality of life, and modifiable risk factors in patients with CAD during telerehabilitation. The study also investigated carryover effects to determine the maintenance of effects after the conclusion of the intervention.

Methods: A total of 32 CAD patients were randomized on a 1:1 ratio to telerehabilitation or usual care. We conducted a 4-month crossover study with a crossover point at 2 months using a mixed-methods approach for evaluation. We collected qualitative data on users' motivation, user experience, and quality of life using questionnaires, semistructured interviews and context-based sentiment analysis. Quantitative data on health parameters, exercise capacity, and risk factors were gathered from blood tests and ergo-spirometry tests. Data procured during the app usage phase were compared against baseline values to assess the impact of the app on parameters such as motivation, physical activity, quality of life, and risk factors. Carryover effects were used to gather insights on the maintenance of effects.

Results: The qualitative data showed that 75% (21/28) of patients found the HeartHab app motivating and felt encouraged to achieve their rehabilitation targets. 84% (21/25) of patients either reached or exceeded their prescribed physical activity targets. We found positive significant effects on glycated hemoglobin (P=.01; d=1.03; 95% CI 0.24-1.82) with a mean decrease of 1.5 mg/dL and high-density lipoprotein (HDL) cholesterol (P=.04; d=0.78; 95% CI 0.02-1.55) with a mean increase of 0.61 mg/dL after patients used the HeartHab app. We observed significant carryover effects on weight, HDL cholesterol, and maximal oxygen consumption (VO_2 max), indicating the maintenance of effects.

Conclusions: Persuasive design techniques integrated in HeartHab and tailoring of exercise targets were effective in motivating patients to reach their telerehabilitation targets. This study demonstrated significant effects on glucose and HDL cholesterol and positive carryover effects on weight, HDL cholesterol, and VO_2 max. There was also a perceived improvement in quality of life. A longer-term evaluation with more patients could possibly reveal effectiveness on other risk factors and maintenance of the positive health behavior change.

Trial Registration: ClinicalTrials.gov NCT03102671; https://clinicaltrials.gov/ct2/show/NCT03102671 (Archived by WebCite at http://www.webcitation.org/76gzI9Pvd)

(JMIR Mhealth Uhealth 2019;7(4):e10874) doi:10.2196/10874

KEYWORDS

heart diseases; cardiac rehabilitation; human factors engineering; evaluation studies; telerehabilitation; mobile app; multidisciplinary research

Introduction

Cardiac Telerehabilitation

A global increase in multiple risk factors such as inactive lifestyles, obesity, diabetes, hypertension, high cholesterol, and smoking is contributing to the higher occurrence of coronary artery disease (CAD). Secondary prevention through cardiac rehabilitation (CR) programs has proven to be effective in reducing these risk factors and thereby reducing the risk of recurrence of disease and rehospitalization [1-3]. Despite these established benefits, the uptake of and adherence to these programs remain low. Distance to rehabilitation centers, time, cost, and psychological barriers are some of the main reasons for reduced uptake of such rehabilitation programs [4,5]. Telerehabilitation includes the use of technology-based approaches to monitor patients remotely [6], alongside facilitating the delivery of other rehabilitation components [7], while they rehabilitate independently after their hospital-based rehabilitation phase. Telerehabilitation approaches enable us to overcome issues in conventional rehabilitation such as distance to rehabilitation center, being restricted to fixed time and appointments, and cost per session [8,9]. However, in both conventional and telerehabilitation approaches, it has been demonstrated that effects of rehabilitation programs are not sustained after stopping the intervention [3], although the rate of decline in effects is lower in telerehabilitation when compared with conventional rehabilitation [10]. In addition, patients do not maintain positive lifestyle adaptations enforced during rehabilitation after the end of the program, leading to increased risk of recurrence.

Previous Work

Multiple studies have proven the benefits of telerehabilitation approaches from different perspectives. Reviews and studies have shown that technology-supported CR approaches are as effective as conventional hospital-based rehabilitation approaches [11-13]. Some studies also prove the cost-effectiveness of engaging in such interventions [9,14]. However, as highlighted earlier, the rate of decline of effects upon completion of intervention is lower in telerehabilitation as compared with conventional rehabilitation [14]. Psychological studies have shown that effectively implementing behavior change interventions supports habit formation and facilitates in sustaining motivation over a longer period [15-17]. Furthermore, when it comes to technology-supported interventions, it is important to design systems that are user-friendly and accessible. To enhance user experience by considering user (patient) needs and perspectives, human-computer interaction (HCI) studies mainly focus on the usability and interaction techniques when designing technology-supported rehabilitation systems. The consideration of patient-specific needs and perspectives can facilitate in tailoring the delivery of these technology-based interventions to cater precisely to the target patients, thereby minimizing attrition.

Goal of the Crossover Study

This study evaluates the effectiveness of an app-based multidisciplinary telerehabilitation program on 4 core factors associated with CR: F1) patients' perspectives, experience, and impact on motivation; F2) impact on physical activity; F3) impact on quality of life; and F4) impact on risk factors.

Methods

App Design

We developed HeartHab, a comprehensive patient-tailored app to support cardiac telerehabilitation. HeartHab covers various modules related to CR such as monitoring risk factors (blood pressure, weight, glucose, cholesterol etc), medication management, physical activity training, e-coaching via specially designed videos, and symptoms monitoring. The design of HeartHab's user interface is grounded on persuasive design techniques drawn from both psychology and HCI literature. The key principle of designing persuasive systems is drawn from the Fogg Behavior Model [18], Persuasive Systems Design model [19], and the Behavior Wizard [20]. On the design level, persuasive design patterns, which are suitable in the context of CR, were derived from the above theories and integrated into visual elements presented to patients such as their risk factor thresholds, activity targets and progress, and so on [21]. On an application level, persuasion is achieved through intelligibility, personalization, and tailoring of various rehabilitation components. The app remotely connects to a caregiver's dashboard used by cardiologists, nurses, and physiotherapists to remotely monitor patients' risk factors physical activity progress, prescribe medication, tailor physical activity targets, and receive timely alerts and notifications of each patient. A video walkthrough of various components of HeartHab is presented in Multimedia Appendix 1.

In our previous work, we studied the usability of HeartHab and patient perceptions of the persuasive techniques used in multiple field studies [10,22,23]. Having validated the usability of the app and having gained promising insights on the influence of intelligibility, tailoring, and other persuasive design elements toward motivating patients, we wanted to assess the actual impact on the aforementioned CR factors (F1-F4). For this, we used a multidisciplinary crossover approach by combining clinical and HCI evaluation methods. By following such an approach, we go beyond conventional randomized controlled trials and usability studies to do a mixed-methods evaluation

and present both qualitative and quantitative outcomes. Furthermore, the crossover approach facilitates in getting detailed insights in a shorter period with considerably lower number of participants. We did not allow for a wash-out period in this specific context as one of our key objectives was to study if there were any crossover effects at the end of an intervention to observe maintenance of effects after the termination of the intervention. The crossover impact also presents indications of maintenance and longer-term outcomes. In this paper, we describe the approach and results of the evaluation in detail including a detailed statistical analysis of crossover effects that were specifically observed.

The medical ethical committees of Hasselt University and Jessa Hospital approved the study. The patients signed an informed consent form before startup and were allowed to withdraw from the study at any point.

Recruitment

A total of 50 cardiac patients who met the inclusion criteria (Textboxes 1 and 2) were screened on the cardiology database of a regional hospital to participate in the study. They were informed of the study details and, if agreeable to participate in the study, asked to sign the consent form. Overall, 32 patients consented to participate in the study and were randomly assigned in a 1:1 ratio to 1 of the 2 treatment strategies (usual care or using the HeartHab app; Figure 1) between April and June 2017. Patients randomized to group 1 used HeartHab in the first phase and received usual care in the second phase, whereas patients randomized to group 2 received usual care in the first phase and HeartHab in the second phase. Each phase of the study was for a period of 2 months (Figure 1). In the usual care phase, patients did not receive any form of rehabilitation support and continued the usual self-management without any supervision.

Textbox 1. Inclusion criteria for patient recruitment in the study.

Inclusion criteria:

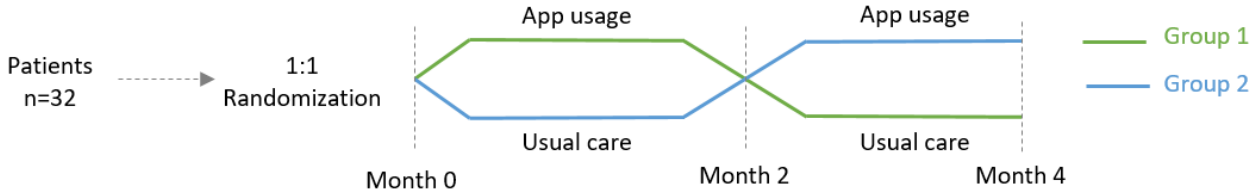
- History of coronary artery disease with or without intervention (percutaneous coronary intervention or coronary artery bypass grafting or conservative)
- History of a cardiac rehabilitation program
- Clinically stable without inducible ischemia or high-risk ventricular arrhythmia, confirmed by the last available maximal ergo-spirometry test
- Aged over 18 years
- Willing and physically able to follow an app-based telerehabilitation program and other study procedures in a 4 months follow-up period
- Possession of and/or able to use an Android smartphone
- Dutch speaking and understanding

Textbox 2. Exclusion criteria for patient recruitment in the study.

Exclusion criteria:

- Recent percutaneous coronary intervention or coronary artery bypass grafting procedure and still included in a cardiac rehabilitation program
- Orthopedic, neurologic, or any other pathologic condition that makes the patient physically unable to follow an app-based telerehabilitation program
- Planned interventional procedure or surgery in the next 4 months
- Pregnant females
- Present cardiovascular complaints
- Participation in other cardiac rehabilitation program trials, focusing on exercise outcome
- Any condition, which, in the opinion of the investigator, would make it unsafe or unsuitable for the patient to participate in this study or a life expectancy of less than 4 months based on investigators' judgment

Figure 1. The process of randomization showing change in phases of the study for both groups before and after crossover.



Qualitative Analysis

Qualitative data were acquired using pretest and intermediate questionnaires collected at baseline and end of 4 weeks in the app usage phase of the study (Figure 2). In addition, 2 HCI researchers conducted a final semistructured interview with patients at the end of their app usage phase. The questionnaires used are appended in Multimedia Appendix 2. All patients received an initial elicitation questionnaire to gather information on their familiarity and comfort with using smartphones and their current approach to log and monitor various aspects of their rehabilitation such as physical activities, medication, physiological parameters, and how they seek further information. During the course of their app use phase, at the end of 4 weeks, they were sent the intermediate questionnaire to gather information on how they perceive the influence of various modules of the app on their telerehabilitation progress, influence on their lifestyle, and motivation. These aspects were evaluated on a 5-point Likert scale. The values collected in the prestudy and intermediate questionnaires were used to probe and tailor questions about the app's influence on their overall health, rehabilitation, and motivation during the final semistructured interview. The interview gave a more nuanced insight on patient perceptions and reasoning behind their perceptions, which cannot be gathered merely through questionnaires. The researchers conducting these interviews took notes, and audio recordings were made of all interviews. We then translated these interviews from Dutch to English and transcribed the responses. The findings were categorized into contexts such as overall experience with HeartHab, insights per module of the app, understandability of visual representations, perceptions on caregivers' intervention, most helpful aspects of the app, features in the app that the patients felt were lacking or missing, and their motivation to continue using the app. We used a validated computer-assisted qualitative data analysis software (NVivo–developed by QSR International) [24,25] to do a context-wise sentiment analysis [26]. We then did a manual in-depth analysis of reasons associated with each positive or negative response.

Quantitative Analysis

Descriptive analyses were performed on data collected using the app and through the various questionnaires (Figure 2). Data collected from the app includes usage logs (logging all interactions with the app including features that were accessed, frequency, the time spent on a certain feature, etc), activities registered by patients using the app, their medication compliance, and evolution of various physiological parameters. Data collected from the questionnaires included overall physical activity levels and quality of life indicators using International Physical Activity Questionnaire (IPAQ) [27], HeartQoL [28], and the 5-level EuroQol 5-Dimensions (EQ-5D) questionnaires [29]. Clinical parameters such as weight, blood pressure, heart rhythm, exercise capacity (maximal oxygen consumption [VO_2 max]), glucose (glycated hemoglobin [HbA_{1c}]), and lipid profile (low-density lipoprotein and high-density lipoprotein [HDL] cholesterol) were collected using blood tests and ergo-spirometry tests at the hospital and rehabilitation center. All measurements were procured at baseline, the crossover point (month 2), and the end (month 4). The gathered data were used to observe mean differences between various parameters during different stages of the study, compare assessments, and make preliminary estimates of the influence of the app on the different evaluation metrics (parameters listed in Multimedia Appendix 3). To gain a precise insight on the actual influence of the app, we used the overall app usage percentage of each patient as a weighting variable against which all parameters were weighted. The overall app use percentage was computed using the number of days of actual app use against the total days that a patient was in the app usage phase. We then performed a detailed weighted statistical analysis on the data according to the intention-to-treat principle by the assigned treatment group. Nonparametric alternatives (such as Wilcoxon 2-sample test) were used for parametric statistics (*t* tests) in case assumptions for the latter were violated.

Test for Significance

For all the evaluations, the level of significance was 2-sided alpha=.05. First, we tested if the data were normally distributed using 4 tests for normality—Shapiro-Wilk test, Kolmogorov-Smirnov test, Cramer-von Mises test, and the Anderson-Darling test. When the data were normally distributed, we used Student *t* test, and when the data were not normally distributed, we used signed rank test to determine significant effects (Figure 3).

Test for Carryover Effects

Being a crossover study design, we also evaluated if there were any significant carryover effects. The primary objective to determine this was to see if any significance observed during the first phase with patients in group 1 (using the app in the first phase) was carried over when they switched to usual care in phase 2 as a means to get an indication on maintenance effects. The process followed to evaluate the carryover effect is detailed in the flowchart (Figure 4).

Figure 2. Qualitative and quantitative data collected across different time points of the study. IPAQ: International Physical Activity Questionnaire; EQ-5D: EuroQol 5-Dimensions questionnaire.

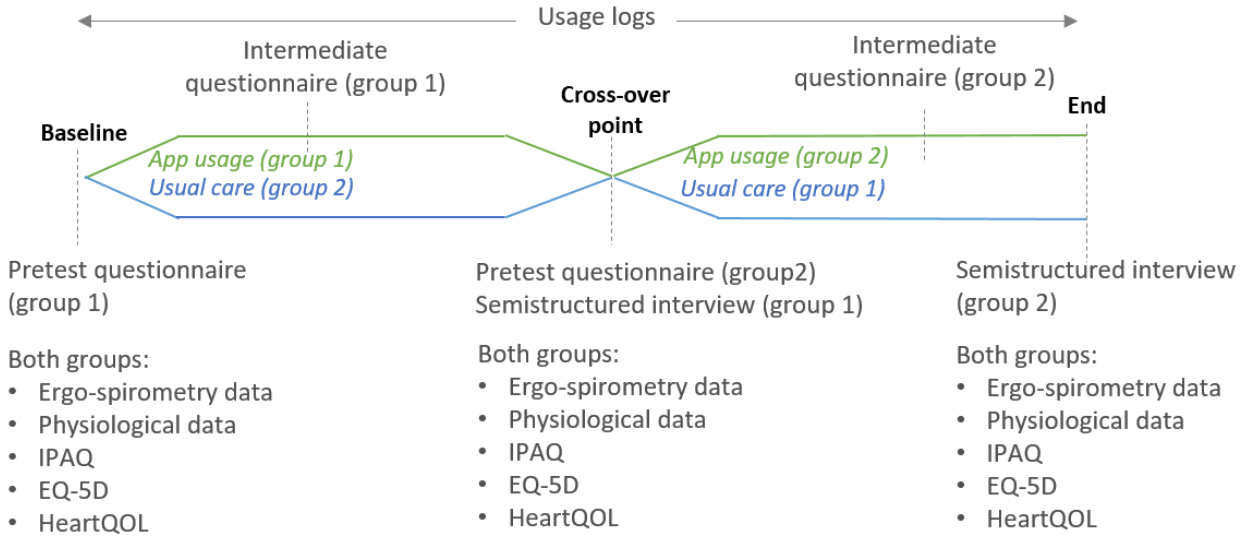


Figure 3. Flowchart showing the statistical analysis process to evaluate the effect of HeartHab on various health metrics.

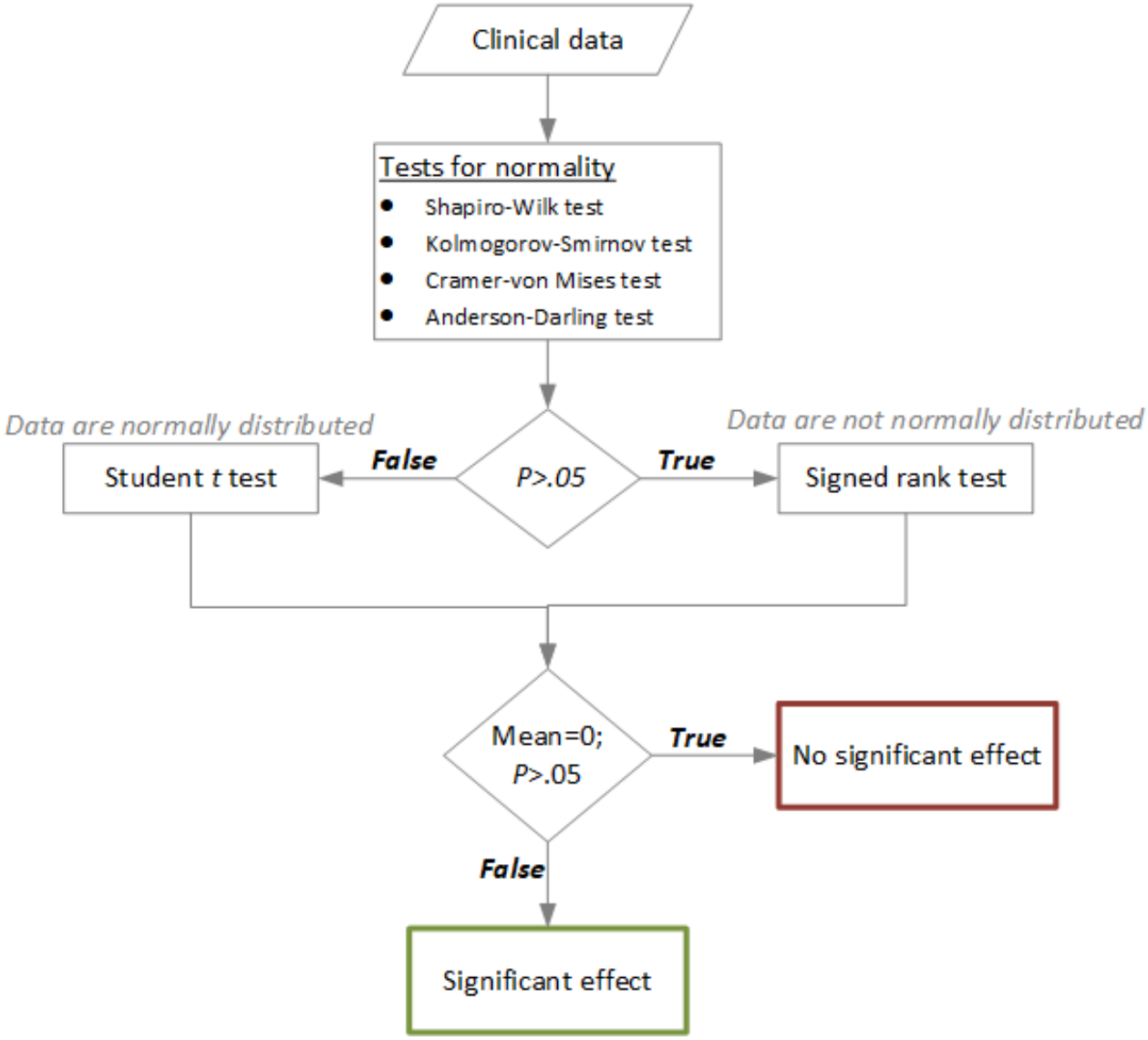
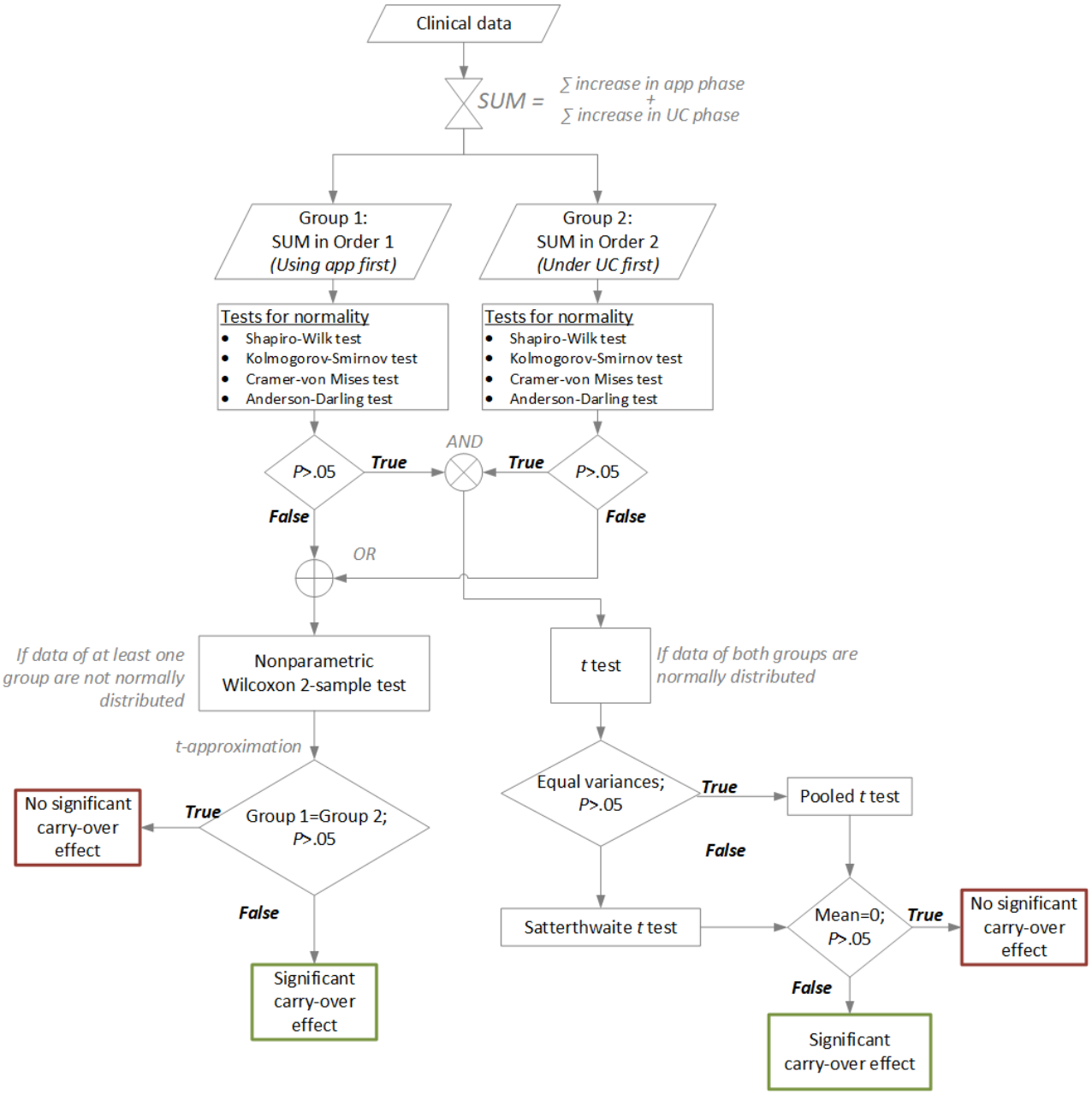


Figure 4. Flowchart showing the process of statistical analysis to identify significant carryover effects. UC: under control; SUM: summation.



Analysis of Subjective Measures

In addition, more information on physical activity and quality of life were procured through standardized questionnaires such as IPAQ, HeartQoL, and the 5-level EQ-5D. The methods followed to analyze these measures are detailed in the following subsections.

Physical Activity Data

The physical activity data collected using the IPAQ were translated into the volume of activity computed by weighting each type of activity by the energy requirements of that type of activity. The energy requirement is represented in terms of Metabolic Equivalents of Task (METs). This then yields a score in MET-minutes calculated by multiplying the METs of an activity by the duration for which it is performed and the resting metabolic rate as defined in IPAQ's scoring protocol [30]. In HeartHab, for prescribing exercise targets and to follow progress made by patients, the METs per activity were procured from the compendium of physical activities [31]. Then alongside resting metabolism and duration of an activity, factors such as intensity at which the activity was performed and the patient's weight were factored in, to get a more precise METs score per individual. The weekly METs scores achieved by each patient were compared against a personalized target range of METs that was prescribed by physiotherapists using the caregivers' dashboard application. The dashboard application integrates the European Association of Preventive Cardiology–supported EXPERT tool for exercise recommendations [32]. The EXPERT tool facilitates physiotherapists and other clinicians to generate tailored exercise recommendations based on the patient's individual pathology and risk factors.

Quality of Life Data

Quality-adjusted life years (QALYs) was used as a generic measure of effectiveness to assess change or evolution in the

quality of life. Estimates of QALYs were derived from the EQ-5D questionnaire. The EQ-5D scores were converted to QALY indices based on the time trade-off technique and conversion weights determined for the Belgian population as per the EQ-5D evaluation protocol [33]. The EQ-5D also uses a visual analog scale (VAS) where patients can pick a score from 0 to 100 (0 being the *worst imaginable* health state and 100 being the *best imaginable* health state) based on their perception of their current health status. The mean VAS scores across different phases were also computed to see if there was a perceived change in the quality of life of patients before or after using the app. In addition, the scores reported in the HeartQoL questionnaire were used to determine the influence of their heart disease on different aspects of physical and emotional well-being. The change in QALY index, VAS, and heart health score across different phases of the study is used to evaluate the impact of the app on the quality of life of patients.

Results

In this section, we present qualitative, quantitative, and comparative results for the 4 factors of CR identified in the goal of this study—F1) patients' perspectives, experience, and impact on motivation; F2) impact on physical activity; F3) impact on quality of life; and F4) impact on risk factors. Detailed statistical outcomes with respect to the impact of use of HeartHab on various physiological parameters, crossover impact, and correlations are presented in Multimedia Appendix 3.

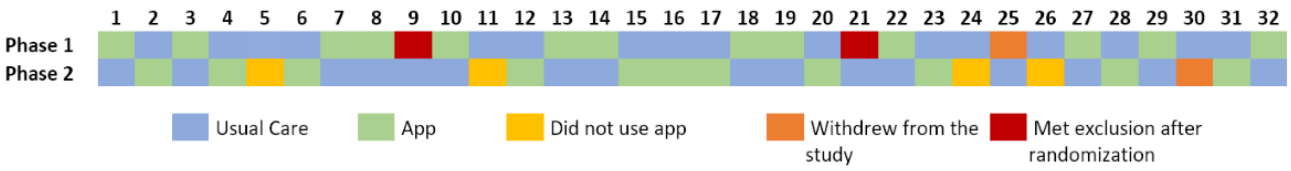
Data Exclusion

Overall, 2 patients were excluded after randomization when they developed other comorbidities such as back problems or encountered another cardiac incident. Therefore, they did not complete the study. In addition, 2 others were excluded as they chose to withdraw from the study during or before startup. The baseline patient information of the remaining 28 patients is presented in Table 1. Given the design of our study, the decision to exclude the data of these patients was made after discussions with a statistician. Patients that had missing data in the IPAQ questionnaire [27] or had overreported their physical activity effort according to IPAQ's scoring protocol [30] were excluded in the analysis of overall physical activity (F2). Patients with missing data in either the EQ-5D [29] or HeartQoL [28] questionnaires were excluded in the assessment of *"Quality of Life"* (F3). Figure 5 illustrates all the patients who were recruited with their assigned randomization, usage levels, and exclusion criteria.

Table 1. Presentation of baseline patient features, cardiac pathology, and medication information (n=28).

General patient features	Baseline values
Age (years), mean (SD)	60.9 (8.2)
Gender, n	24 males and 4 females
Body mass index (kg m s^{-1}), mean (SD)	28.7 (5.2)
Cardiac pathology, n	
Coronary artery disease, percutaneous coronary intervention, coronary artery bypass grafting, and endoscopic atraumatic coronary artery bypass	28
Cardiac resynchronization therapy, pacemaker, and implantable cardioverter defibrillator	1
Pulmonary arterial hypertension	1
Medication, n	
Calcium antagonists	2
Beta-blockers	17
Angiotensin-converting enzyme inhibitors	19
Statins	25
Antiplatelets	28

Figure 5. Representation of all patients recruited in the study, their randomization sequence, and app usage.



F1) Patients' Perspectives, Experience, and Impact on Motivation

Most patients were positive on the impact of using HeartHab on being more physically active and more medically compliant (Figure 6). Regarding how the app promoted being more physically active, P20 remarked:

> *I have been a bit more active than usual, and it also helped a lot that I could follow up all my activities. It pushed me to do a bit more than I usually tend to do.*

Another patient stated the following:

> *My efforts have increased. The flag [referring to the persuasive visualization of exercise target in the app's interface]was motivating. It [pointing to the goals and progress on the app] motivates you to exercise more.* [P18]

There were mixed responses on adopting a healthier lifestyle. Although some patients perceived the influence on physical activity and medication compliance as aspects of healthier lifestyle, some patients opined that they had already tried to adopt a healthier lifestyle as their diagnosis and the app did not change much:

> *Actually, the application should be used immediately in rehabilitation. Then the app is probably more useful than with me, but it's all right.* [P26]

On the other hand, P20 said:

> *I lost 7 kilograms. It [pointing to the phone/app] was one of the factors which pushed me towards a healthier life. You are aware of the fact that you are in a study and the fact that the app reminds you and nudges you at the correct time helps a ton.*

Most patients were positive about the support provided by the app in their rehabilitation, and 80% of the patients were willing to continue using the app to support long-term self-management (Figure 7). For example, when asked about their motivation to continue using HeartHab, P11 said:

> *Too bad it is being taken away[when the app was uninstalled from his phone at the end of the study].It was a luxury to have it. I had my entire heart history...good background information and it can also be useful for the cardiologist.*

Another patient said:

> *Yes of course. I don't have to hesitate about that [referring to willingness to continue using HeartHab] for a second. There exist many applications [referring to other generic health apps] that are less useful than this one.* [P7]

For patients who did not want to continue using the app, it was mainly because they believed that they already lead an active and healthy lifestyle and did not necessarily need an app. Other patients were not looking forward to continue using the app because of their discomfort in using smartphones. P2 also remarked:

> *My wife helps me now. The app is more interesting for singles [possibly referring to people that do not have an active informal caregiver] and poorly motivated people.*

Another patient said the following:

> *I am already quite active. I just have to watch my cholesterol but this [referring to the app] did not help me a lot further.* [P27]

Figure 6. Results of sentiment analysis using NVivo on increasing physical activity levels, promoting medication adherence and adopting healthier lifestyles after using the HeartHab app.

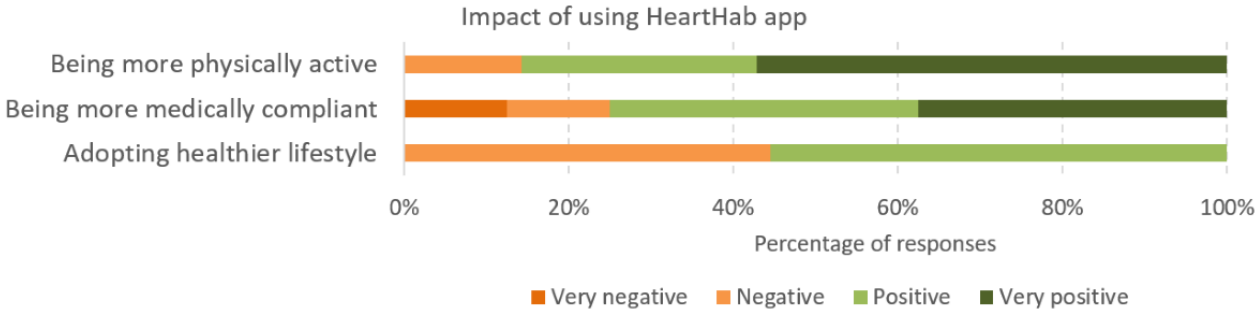
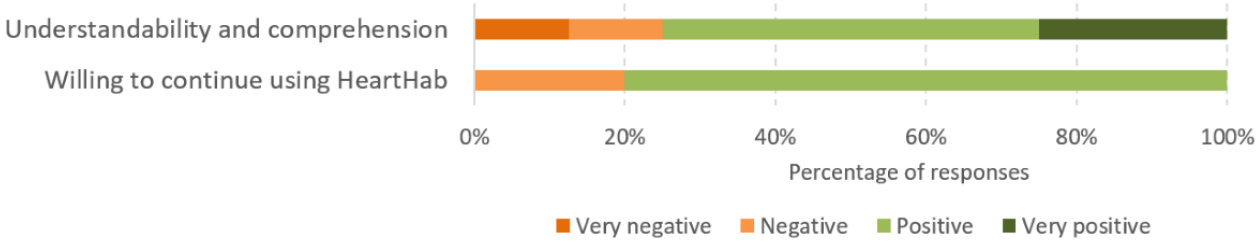


Figure 7. Results of sentiment analysis using NVIVO on patients' motivation to continue using HeartHab and the impact of the app on improving their overall understandability and comprehension.



F2) Impact on Physical Activity Levels

The progress and prediction visualizations facilitate in nudging patients to reach their targets. In both the intermediate questionnaire (collected after 4 weeks in the app use phase; Figure 8) and the final semistructured interview, most patients responded positively about both these elements (Figure 9). Therefore, the persuasive design approach applied in the app [21] certainly had an influence on the motivation of patients:

> *I try to exercise more. I would aim to get to the flag [referring to the visualization of targets in the app]. It was not always successful because of lack of time. But it certainly motivated me.* [P18]
>
> *I really liked the fact that I could explore the effect of different activities on the goals! So I chose my activity sometimes based on the effect it would have on my progress towards the goals.* [P10]

Here, the patient refers to the feature in the app that predicts the effect of an activity in terms of METs, thus facilitating the selection of activities that contribute enough to gradually reach the personalized target.

Even patients who were already active and did not necessarily increase their physical activity found the app motivating and felt reassured when they saw their progress:

> *I was already active. So, I did not do more. But it was nice to see progress... liked the motivational messages the app would give after entering a new activity such as "keep going", "good job"etc...* [P12]

These subjective outcomes were also reflected in the target achievement rates that were measured using the app. Overall, 84% (21/25) of patients either reached or exceeded their maximal target on average (Figure 10). The targets were tailored individually based on their exercise capacity as measured in the ergo-spirometry test. For example, P1 was prescribed to undertake exercise training for 7 times a week for a duration of 20 to 60 min per session based on his exercise capacity, which yielded a minimal weekly target of 921 METs (ie, 20 min/session and 7 sessions) and a maximal target of 4145 METs (ie, 60 min/session and 7 sessions). On the other hand, P7 who had a much lower exercise capacity was prescribed to undertake exercise training for 3 to 5 times a week for a duration of 20 to 45 min per session. This resulted in a minimal target of 269 METs (ie, 20 min/session and 3 sessions) and a maximal target of 1512 METs (ie, 45 min/session and 5 sessions). This personalized tailoring of targets motivated better target achievements and added to the credibility of the app as reflected by the following patient perceptions:

> *...it [the targets] seemed correct and trustworthy for me, realistic; otherwise I would not have looked at the activity goals.* [P10]
>
> *I assume that it is credible because it is prescribed by the physiotherapist. That motivates me.* [P18]
>
> *The goals were achievable and made for continuous use* [P2]

In the study, the duration of app usage varied between 7 to 10 weeks for all patients. The representation of weekly METs achievement in Figure 10 only shows values that were registered by patients in the app. There were weeks where some patients did not register all activities, and the incomplete data are not visualized in the chart.

As inferred from the graphs (Figure 11) for both groups of patients, there was a clear increase in total METs during the app phase. For some patients in group 1 (ie, patients who used the app in the first phase), we can also see the increasing trend continue onto the control phase. The increase we observe with patients in group 2 upon starting the app also suggests the positive impact of the app on increasing physical activity levels. For 1 patient in group 2, there is a substantial increase in the control phase and a steep drop in the app phase. The increase in control phase was due to a regional sporting event in which the patient participated. Therefore, the trend seen in this graph does not reflect on the patients' normal activity behavior.

Although the data of all patients are not visualized in the graphs because of incomplete data in the IPAQ questionnaires, the statistical evaluation of exercise capacity in terms of VO_2 max obtained from the ergo-spirometry tests across the phases showed a positive significant carryover effect (P=.008; Multimedia Appendix 3). This indicates the maintenance of acquired exercise capacity even upon completion of intervention (use of app), which was one of the key goals of this study.

Figure 8. Patients' perceptions on the impact of HeartHab on motivation to achieve physical activity targets as collected in the intermediate questionnaire.

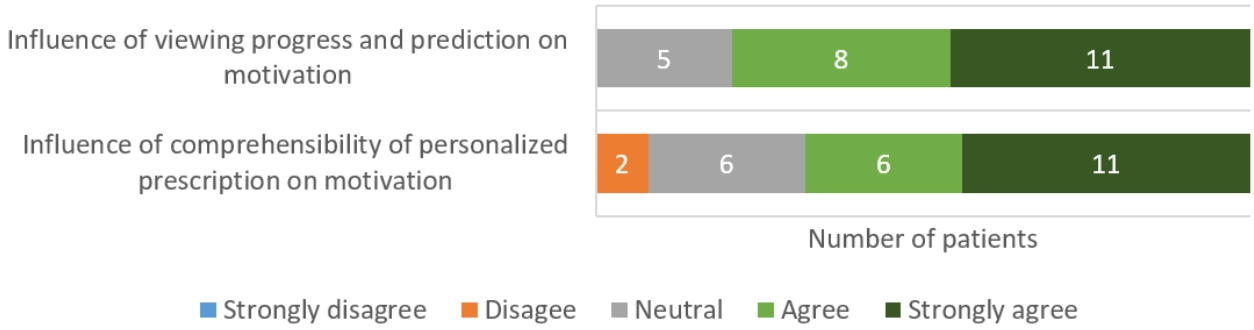


Figure 9. Sentiment analysis using NVivo on the impact of various aspects of HeartHab on motivation to be more physically active.

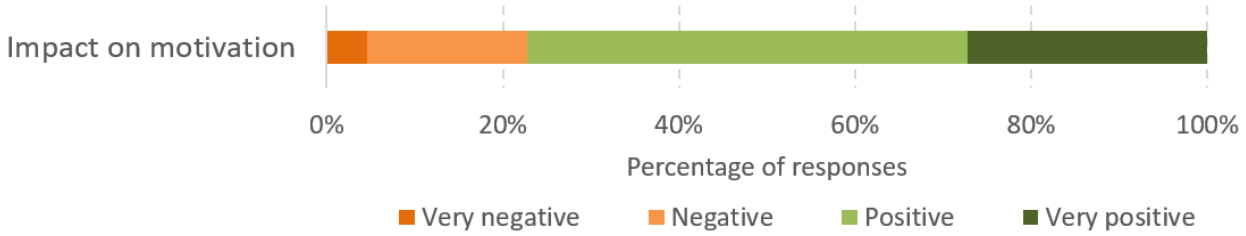


Figure 10. Chart showing weekly physical activity targets and totals METs achieved each week by patients that used the physical activity module of the HeartHab app. The chart depicts only complete data for weeks when a patient registered activities. MET: Metabolic Equivalents of Task.

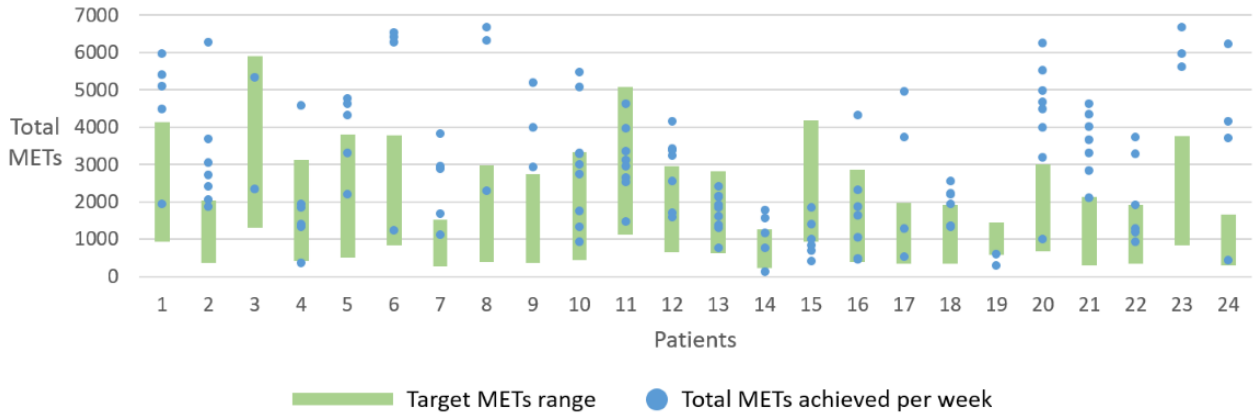


Figure 11. Graphs showing increase or decrease in total Metabolic Equivalents Of Task across different phases of the study for both groups of patients as collected from the International Physical Activity Questionnaire. MET: Metabolic Equivalents of Task.

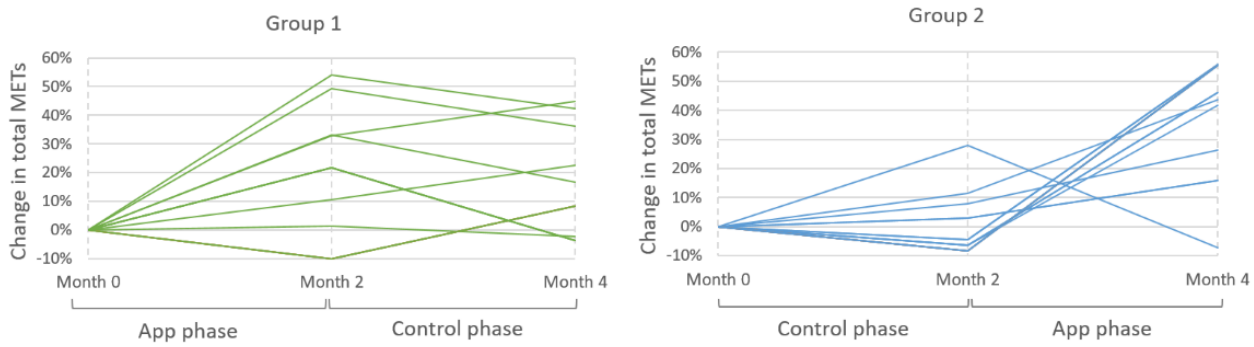
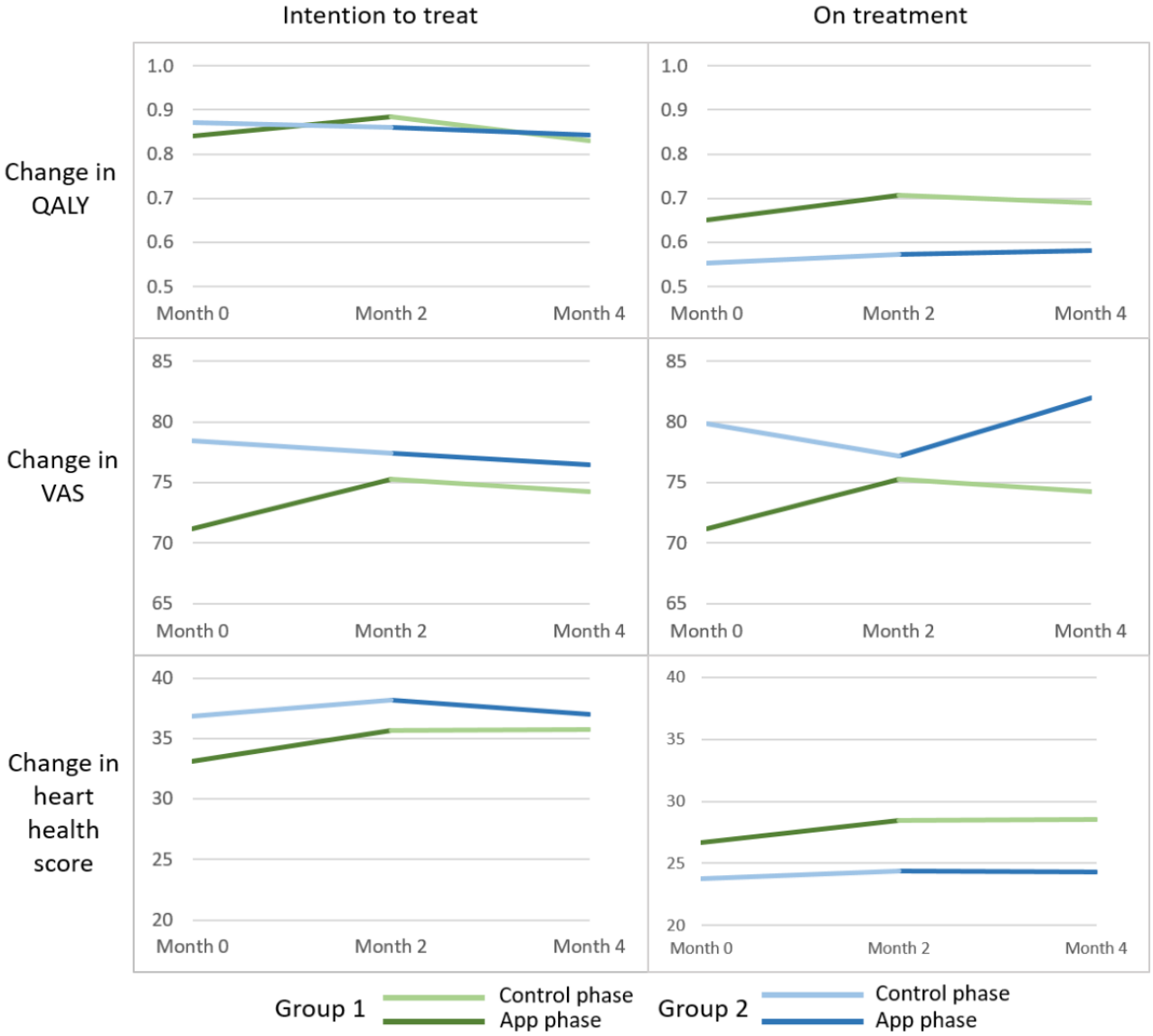


Figure 12. Changes in various assessments of quality of life across different phases of the study. QALY: quality-adjusted life years; VAS: visual analog scale.



F3) Impact on Quality of Life

Figure 12 depicts the quality of life data obtained from EQ-5D and HeartQoL questionnaires as assessed in 2 arms—intention-to-treat and on-treatment. In the intention-to-treat approach, the data of all patients that completed the study were included. For the on-treatment approach, the data of patients were excluded if there were any incomplete data in any of the phases of the study or if the patient developed a new symptom such as back pain, which was not directly caused by the app.

On the basis of the EQ-5D questionnaire, there was a mean increase of 0.06 QALY and a mean decrease of 7.14% in anxiety during the app usage phase for both groups of patients using the intention-to-treat analysis. From VAS in Figure 12, we can observe a clear trend in the perceived increase of health during the app phase by both groups of patients, which was lacking in the control phase. We could also see a marginal increase in heart health scores obtained from the HeartQoL questionnaire.

F4) Impact on Risk Factors

On the basis of the blood tests taken at baseline, crossover point (end of month 2), and final (end of month 4) appointments, we found a significant increase in HDL, which is also nicknamed *"the good cholesterol"* with *P*=.048 (*d*=0.78; 95% CI 0.02-1.55). There was a mean increase of 0.61 mg/dL in HDL cholesterol in the app usage phase as compared with a mean decrease of 2.28 mg/dL during the control period, demonstrating a clear impact of app use on this parameter. In addition, we found a significant decrease in HbA_{1c} with *P*=.01 (*d*=1.03; 95% CI 0.24-1.82). HbA_{1c} gives the glucose concentration in blood, which gives an indication of average blood sugar levels over a period. Although there was a mean decrease of 0.3 mg/dL in HbA_{1c} in the control period, there was a mean decrease of 1.5 mg/dL during the app usage period. Therefore, a significant decrease in this parameter in the app phase indicates a significant decrease in risk of diabetes-related complications. The increase in overall physical activity is reflected in the positive significant effects on these parameters as physical activity does have an

influence on both HbA_{1c} and HDL. With respect to other risk factors such as weight (d=0.4076) and pulse (d=0.4518), we observed a positive trend with a mean decrease in weight of 0.02 kg and a mean decrease in pulse of 3 bpm, although the results were not statistically significant.

Discussion

Reflections on Outcomes of the Study

The qualitative findings put forward a nuanced interpretation of how motivated patients feel across different phases of the study. The perceptions of patients may not always map to the quantitative results, but gathering those findings is equally important. For example, a patient who did not like to sport or engage in rigorous physical activity mentioned the following:

> *...I either walked or biked to the supermarket instead of taking my car so I can register it on the app.* [P10]

Such small changes in health behavior as gathered from qualitative evaluations cannot be observed purely from quantitative measures and are, nonetheless, potentially beneficial for this patient population. These observations are also suggestive of the impact of the persuasive elements of the HeartHab app on the motivation of patients.

The findings presented in this paper establish the positive effect of the app of study goals F1 (patients' perspectives, experience, and impact on motivation) and F2 (impact on physical activity). The impact on physical activity is also reflected in the significant effects on other parameters such as glucose and cholesterol. In addition, the carryover effects that we observed in this study are clinically important because one of the main challenges in telerehabilitation is the high rate of dropout and nonadherence after the intervention concludes [9]. Therefore, it is important for telerehabilitation to support and gradually train patients toward self-management. In this study, we found significant carryover effects on weight, HDL, and VO_2 max, indicating the sustenance of motivation and intervention effects at the end of the intervention. The willingness of patients to continue using the app at the end of the study is a promising indication toward minimizing attrition and dropouts. Moreover, patients' perceptions on the influence of viewing progress and prediction on their motivation and target achievement indicates how the app could support better self-awareness and promote self-management. Furthermore, we observed clinically significant effects on cholesterol and glucose. The effect on glucose concentration (d=1.03) is higher as compared with cholesterol (d=0.78) as patients have a better control over their insulin intake and sugar consumption. Effects on other parameters are not clinically significant and could be attributed to the relatively short duration of this study. Finally, the reduction in anxiety and perceived increase in quality of life emphasize on the positive effect of the app on goals F3 (impact on quality of life) and F4 (impact on risk factors) of the study.

Implications for Future Work

We acknowledge that the reported results are representative of only a small sample studied at 1 rehabilitation setting where all patients underwent the same supervised rehabilitation program and received similar advice on maintenance and follow-up with CR. It might, therefore, be interesting to evaluate the impact of the app on patients across different settings and in different rehabilitation phases. The subjective perceptions could also vary when studied under these different settings. However, by detailing our approach and findings, we hope to enable researchers to make an informed assessment of the adoption of similar methods within their practice or research contexts. Adding intelligibility to enable patients to better comprehend their rehabilitation progress and tailoring the intervention to suit individual needs seem highly useful in persuading patients. Future research could assess these factors in a longer duration to also study how patients' engagement varies across different phases of intervention and after the conclusion of the intervention. We believe that such integration of persuasive techniques in app development and evaluation using a mixed-methods approach can add to the existing knowledge base of technology-supported cardiac telerehabilitation and pave way for more such multidisciplinary research in this domain.

Conclusions

We conducted a multidisciplinary crossover study with 32 cardiac patients, and 28 patients successfully completed the study using the HeartHab app. We found positive significant effects of the app-based telerehabilitation approach on HbA_{1c} (P=.01; d=1.03; 95% CI 0.24-1.82) and HDL cholesterol (P=.04; d=0.78; 95% CI 0.02-1.55). We also found positive significant carryover effects on VO_2 max, HDL, and weight. The persuasive techniques applied in HeartHab yielded a positive outcome on patient motivation and facilitated them in achieving overall physical activity targets. The insights gathered from the study could help address some of the challenges faced by current systems targeting cardiac telerehabilitation and secondary prevention. A longer evaluation with a larger group of patients can potentially lead to other significant health outcomes and address remaining challenges that are currently hampering the effective widespread implementation of cardiac telerehabilitation.

Acknowledgments

The authors would like to acknowledge the contributions of Nele Mommen and Lisa Huygen in the drafting of the study protocol and analysis of the results, respectively. The authors are grateful for the assistance provided by the team at Jessa Hospital and ReGo (the rehabilitation center) at Hasselt, Belgium, where they conducted all studies. The authors would also like to extend their gratitude to their HCI colleagues at Expertise Center for Digital Media, Hasselt University, and the students for their assistance through various phases of research, development, and evaluation of the different studies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Video walkthrough of the HeartHab mobile application.

[MP4 File (MP4 Video), 6MB - mhealth_v7i4e10874_app1.mp4]

Multimedia Appendix 2

Human-computer interaction questionnaires that were used for collecting qualitative data across different phases of the study.

[PDF File (Adobe PDF File), 388KB - mhealth_v7i4e10874_app2.pdf]

Multimedia Appendix 3

Statistical data analysis for effects of HeartHab on various parameters and carryover effects.

[PDF File (Adobe PDF File), 128KB - mhealth_v7i4e10874_app3.pdf]

References

1. Kotseva K, Wood D, De Bacquer D, De Backer G, Rydén L, Jennings C, et al. EUROASPIRE IV: a European Society of Cardiology survey on the lifestyle, risk factor and therapeutic management of coronary patients from 24 European countries. Eur J Prev Cardiol Internet 2016 Apr:16. [doi: 10.1177/2047487315569401] [Medline: 25687109]
2. Anderson L, Oldridge N, Thompson DR, Zwisler A, Rees K, Martin N, et al. Exercise-based cardiac rehabilitation for coronary heart disease: Cochrane systematic review and meta-analysis. J Am Coll Cardiol 2016 Jan 5;67(1):1-12 [FREE Full text] [doi: 10.1016/j.jacc.2015.10.044] [Medline: 26764059]
3. Hansen D, Dendale P, Raskin A, Schoonis A, Berger J, Vlassak I, et al. Long-term effect of rehabilitation in coronary artery disease patients: randomized clinical trial of the impact of exercise volume. Clin Rehabil 2010 Apr;24(4):319-327. [doi: 10.1177/0269215509353262] [Medline: 20176771]
4. Turk-Adawi KI, Oldridge NB, Tarima SS, Stason WB, Shepard DS. Cardiac rehabilitation patient and organizational factors: what keeps patients in programs? J Am Heart Assoc 2013 Oct 21;2(5):e000418 [FREE Full text] [doi: 10.1161/JAHA.113.000418] [Medline: 24145743]
5. Piepoli M, Corrà U, Dendale P, Frederix I, Prescott E, Schmid J, et al. Challenges in secondary prevention after acute myocardial infarction: a call for action. Eur J Prev Cardiol 2016 Dec;23(18):1994-2006. [doi: 10.1177/2047487316663873] [Medline: 27600690]
6. Piotrowicz E, Piotrowicz R. Cardiac telerehabilitation: current situation and future challenges. Eur J Prev Cardiol 2013 Jun;20(2 Suppl):12-16. [doi: 10.1177/2047487313487483c] [Medline: 23702985]
7. Piepoli M, Corrà U, Benzer W, Bjarnason-Wehrens B, Dendale P, Gaita D, Cardiac Rehabilitation Section of the European Association of Cardiovascular PreventionRehabilitation. Secondary prevention through cardiac rehabilitation: from knowledge to implementation. A position paper from the Cardiac Rehabilitation Section of the European Association of Cardiovascular Prevention and Rehabilitation. Eur J Cardiovasc Prev Rehabil 2010 Feb;17(1):1-17. [doi: 10.1097/HJR.0b013e3283313592] [Medline: 19952757]
8. Jolly K, Taylor R, Lip G, Stevens A. Home-based cardiac rehabilitation compared with centre-based rehabilitation and usual care: a systematic review and meta-analysis. Int J Cardiol 2006 Aug 28;111(3):343-351. [doi: 10.1016/j.ijcard.2005.11.002] [Medline: 16316695]
9. Frederix I, Hansen D, Coninx K, Vandervoort P, Vandijck D, Hens N, et al. Effect of comprehensive cardiac telerehabilitation on one-year cardiovascular rehospitalization rate, medical costs and quality of life: a cost-effectiveness analysis. Eur J Prev Cardiol 2016 May;23(7):674-682. [doi: 10.1177/2047487315602257] [Medline: 26289723]
10. Sankaran S, Frederix I, Haesen M, Dendale P, Luyten K, Coninx K. A Grounded Approach for Applying Behavior Change Techniques in Mobile Cardiac Tele-Rehabilitation. In: Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments. New York, NY, USA: ACM; 2016 Presented at: PETRA'16; June 29-July 1, 2016; Corfu Island, Greece URL: https://dl.acm.org/citation.cfm?id=2910674.2910680 [doi: 10.1145/2910674.2910680]
11. Kraal JJ, Peek N, Van den Akker-Van Marle ME, Kemps HM. Effects of home-based training with telemonitoring guidance in low to moderate risk patients entering cardiac rehabilitation: short-term results of the FIT@Home study. Eur J Prev Cardiolog 2014 Oct 29;21(2_suppl):26-31. [doi: 10.1177/2047487314552606] [Medline: 25354951]
12. Hwang R, Bruning J, Morris N, Mandrusiak A, Russell T. A systematic review of the effects of telerehabilitation in patients with cardiopulmonary diseases. J Cardiopulm Rehabil Prev 2015;35(6):380-389. [doi: 10.1097/HCR.0000000000000121] [Medline: 26034937]

13. Frederix I, Vanhees L, Dendale P, Goetschalckx K. A review of telerehabilitation for cardiac patients. J Telemed Telecare 2015 Jan;21(1):45-53. [doi: 10.1177/1357633X14562732] [Medline: 25475219]
14. Frederix I, Solmi F, Piepoli M, Dendale P. Cardiac telerehabilitation: a novel cost-efficient care delivery strategy that can induce long-term health benefits. Eur J Prev Cardiol 2017 Dec;24(16):1708-1717. [doi: 10.1177/2047487317732274] [Medline: 28925749]
15. Thaler R, Sunstein C. Nudge: Improving Decisions About Health, Wealth, And Happiness. London, England: Penguin Books; 2008.
16. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Dec 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
17. Stawarz K, Cox A, Blandford A. Beyond Self-Tracking and Reminders: Designing Smartphone Apps That Support Habit Formation. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2015 Presented at: CHI'15; April 18-23, 2015; Seoul, Republic of Korea URL: https://dl.acm.org/citation.cfm?id=2702230 [doi: 10.1145/2702123.2702230]
18. Fogg BJ. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. New York, NY, USA: ACM; 2009 Presented at: Persuasive'09; April 26-29, 2009; Claremont, California, USA URL: https://dl.acm.org/citation.cfm?id=1541999 [doi: 10.1145/1541948.1541999]
19. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Comm Assoc Inf Syst 2009;24:500 [FREE Full text] [doi: 10.17705/1CAIS.02428]
20. Fogg BJ, Hreha J. Behavior wizard: a method for matching target behaviors with solutions. In: Proceedings of the 5th international conference on Persuasive Technology. 2010 Presented at: PERSUASIVE'10; June 7-10, 2010; Copenhagen, Denmark p. 117-131 URL: https://dl.acm.org/citation.cfm?id=2164030 [doi: 10.1007/978-3-642-13226-1_13]
21. Sankaran S, Frederix I, Haesen M, Dendale P, Luyten K, Coninx K. A Grounded Approach for Applying Behavior Change Techniques in Mobile Cardiac Tele-Rehabilitation. New York, NY, USA: ACM; 2016 Presented at: 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments; June 29 - July 01, 2016; Corfu island, Greece.
22. Sankaran S, Luyten K, Hansen D, Dendale P, Hospital J, Coninx K. Have You Met Your METs?- Enhancing Patient Motivation to Achieve Physical Activity Targets in Cardiac Tele-rehabilitation. In: Proceedings of the 32nd International BCS Human Computer Interaction Conference. London, UK: BCS Learning and Development Ltd; 2018 Presented at: HCI 2018; July 4-6, 2018; Belfast, UK. [doi: 10.14236/ewic/HCI2018.48]
23. Sankaran S, Bonneux C, Dendale P, Coninx K. Bridging Patients' Needs and Caregivers' Perspectives to Tailor Information Provisioning during Cardiac Rehabilitation. In: Proceedings of the 32nd International BCS Human Computer Interaction Conference. London, UK: BCS Learning and Development Ltd; 2018 Presented at: HCI'18; July 4-6, 2018; Belfast, UK. [doi: 10.14236/ewic/HCI2018.49]
24. Richards L. Using NVIVO in Qualitative Research. London, UK: Sage Publications; 1999.
25. Welsh E. Dealing with data: using NVivo in the qualitative data analysis process. Forum Qual Soc Res 2002;3(2) [FREE Full text] [doi: 10.17169/fqs-3.2.865]
26. Pawar AB, Jawale MA, Kyatanavar DN. Fundamentals of sentiment analysis: concepts and methodology. In: Pedrycz W, Chen SM, editors. Sentiment Analysis and Ontology Engineering. Switzerland: Springer; 2016:25-48.
27. International Physical Activity Questionnaire. URL: https://sites.google.com/site/theipaq/home [accessed 2018-02-19] [WebCite Cache ID 6xLUL6ydJ]
28. Oldridge N. Heartqol Questionnaire: a new patient-reported outcome in cardiology. Eur Heal Psychol 2014;16:379 [FREE Full text]
29. EuroQol Research Foundation. Rotterdam, Netherlands; 2017. EQ-5D-5L | About URL: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/ [accessed 2018-02-19] [WebCite Cache ID 6xLVPUEmH]
30. Google Docs. 2005. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) URL: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx0aGVpcGFxfGd4OjE0NDgxMDk3NDU1YWRlZTM [accessed 2019-03-05] [WebCite Cache ID 6xLVgoUHo]
31. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 2000 Sep;32(9 Suppl):S498-S504. [doi: 10.1097/00005768-200009001-00009] [Medline: 10993420]
32. Hansen D, Dendale P, Coninx K, Vanhees L, Piepoli M, Niebauer J, et al. The European Association of Preventive Cardiology Exercise Prescription in Everyday Practice and Rehabilitative Training (EXPERT) tool: A digital training and decision support system for optimized exercise prescription in cardiovascular disease. Concept, definitions and construction methodology. Eur J Prev Cardiol 2017 Dec;24(10):1017-1031. [doi: 10.1177/2047487317702042] [Medline: 28420250]
33. van Reenen M, Janssen B. EuroQol Research Foundation. 2015. EQ-5D-5L User guide URL: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf [accessed 2019-03-05] [WebCite Cache ID 76eBRitqu]

Abbreviations

CAD: coronary artery disease
CR: cardiac rehabilitation
EQ-5D: EuroQol 5-Dimensions
HbA_{1c}: glycated hemoglobin
HCI: human-computer interaction
HDL: high-density lipoprotein
IPAQ: International Physical Activity Questionnaire
MET: Metabolic Equivalents of Task
QALY: quality-adjusted life years
VAS: visual analog scale
VO_2 max: maximal oxygen consumption

Edited by G Eysenbach; submitted 25.04.18; peer-reviewed by M Varnfield, J Rawstorn, S Fraser; comments to author 30.08.18; revised version received 30.11.18; accepted 10.02.19; published 04.04.19.

Please cite as:
Sankaran S, Dendale P, Coninx K
Evaluating the Impact of the HeartHab App on Motivation, Physical Activity, Quality of Life, and Risk Factors of Coronary Artery Disease Patients: Multidisciplinary Crossover Study
JMIR Mhealth Uhealth 2019;7(4):e10874
URL: https://mhealth.jmir.org/2019/4/e10874/
doi:10.2196/10874
PMID:30946021

Original Paper

The Mindfulness App Trial for Weight, Weight-Related Behaviors, and Stress in University Students: Randomized Controlled Trial

Lynnette Nathalie Lyzwinski[1], MPhil Public Health, BA Health Science First Class with Distinction; Liam Caffery[1], PhD, Dip App Sci, BInfo Tech; Matthew Bambling[1], PhD, MSc, BSc; Sisira Edirippulige[1], PhD, MPH, MSc

University of Queensland, School of Medicine, Centre for Online Health, Woloongabba, Australia

Corresponding Author:
Lynnette Nathalie Lyzwinski, MPhil Public Health, BA Health Science First Class with Distinction
University of Queensland
School of Medicine
Centre for Online Health
Princess Alexandra Hospital
Building 33 Ground Floor
Woloongabba, 4102
Australia
Phone: 61 7 3176 7403
Email: lynnette.lyzwinski@uqconnect.edu.au

Abstract

Background: University students are at risk of weight gain during their studies. Key factors related to weight gain in this population include unhealthy weight-related behaviors because of stress. Mindfulness holds promise for weight management. However, there has not been any previous trial that has explored the effectiveness of a student-tailored mindfulness app for stress, weight-related behaviors, and weight. There is limited evidence that current mindfulness apps use evidence-based mindfulness techniques. A novel app was developed that combined evidence-based, mindfulness-based stress reduction and mindful eating (ME) techniques that were tailored to university students, with student-relevant themes for targeting weight behaviors, weight, and stress.

Objectives: The aim of this study was to test the effectiveness, acceptability, and feasibility of a student-tailored mindfulness app for weight, weight-related behaviors, and stress. Testing this app in a rigorous randomized controlled trial (RCT) for these outcomes is a novelty and contribution to this emerging field.

Methods: A 2-arm RCT of an 11-week duration was undertaken at the University of Queensland. Students were either randomized to the mindfulness app (n=45) or to a behavioral self-monitoring electronic diary (e-diary; n=45) for diet and exercise. Analysis of covariance was used to compare differences in weight, stress, mindfulness, ME, physical activity, and eating behaviors between both groups.

Results: Neither the mindfulness app group nor the e-diary group lost weight and there were no differences between the groups at follow-up. The mindfulness app group had significantly lower stress levels (P=.02) (adherers only), lower emotional eating (P=.02), and uncontrolled eating (P=.02) as well as higher mindfulness (P≤.001) and ME levels overall (P≤.001). The e-diary group had higher metabolic equivalents of moderate activity levels (P≤.01). However, the effect sizes were small. Regular adherence to mindfulness exercises in the app was low in the group. The majority of students (94%) liked the app and found it to be acceptable. Compared with other exercises, the most helpful reported meditation was the short breathing exercise observing the breath (39.4% [13/33] preferred it). This was the first RCT that tested a mindfulness app for weight and weight-related behaviors in students. The modest level of user adherence likely contributes to the lack of effect on weight loss. However, there was a small, albeit promising, effect on weight-related eating behavior and stress.

Conclusions: A mindfulness app demonstrated effectiveness for stress, eating behaviors, mindfulness, and ME, but the effect sizes were small. Future studies should be conducted over longer periods of time and with greater participant compliance.

Trial Registration: Australian New Zealand Trial Registry ACTRN12616001349437; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=371370 (Archived by WebCite at http://www.webcitation.org/761cc2K6ft)

(JMIR Mhealth Uhealth 2019;7(4):e12210) doi:10.2196/12210

KEYWORDS

mobile applications; mindfulness; body weight; feeding behavior; exercise; stress, psychological; students

Introduction

Background

Obesity and overweight are important international public health challenges that are critical to tackle across the globe given that they are leading risk factors for premature mortality and morbidity from a range of chronic diseases [1]. University students are a high-risk group for rapid and significant weight gain over a short period of time [2-5]. Recent reviews suggest that 60.9% of students gain weight in their freshman year [4]. The prevailing literature indicates that students who gain weight in their first year gain between 3.1 and 3.38 kg [3,4], and the reported rate of weight gain ranges from 110 to 156 grams per week [6,7]. This is significantly greater than the reported weight gain in young adults in the general population as the literature indicates that they gain anywhere between 605 grams and 1 kg per year [8-10]. Research also indicates that university students continue to gain weight during the remainder of their study years [11].

The leading causes of weight gain in this population have been reported to be an unhealthy diet, insufficient physical activity, and stress [3,5]. Stress has been reported to be highly prevalent in college students, with 80% experiencing some form of stress in their regular student lives [12,13]. The prevailing literature has linked stress with weight gain [14-17] and engagement in maladaptive weight-related behaviors in university students such as binge eating before exams, increased consumption of unhealthy food, or physical inactivity in students [18-30].

Emerging research suggests that mindfulness may hold potential for weight management and weight-related behaviors including reduced emotional and binge eating [31,32] as well as reduced stress [33]. The most recent systematic review and meta-analysis found that mindfulness assists with weight loss by 6.8 pounds, with moderate significant effect sizes (g=0.42; 95% CI excluding 0; P<.01), and that the effect size increased by 0.10 when both formal mindfulness practices were combined with informal practice [34]. They also found that the effect was greatest for weight-related eating behaviors than for weight (g=0.70; 95% CI excluding 0; P<.01) [34].

In addition, the most recent systematic review of mindful eating (ME) found that ME alone not only assists with weight management but also weight behaviors such as cravings and control of excess caloric intake [35]. Given that binge eating has been associated with weight and obesity [36,37], targeting these behaviors to promote healthy eating patterns and weight management is desired.

Mindfulness refers to a state of heightened awareness of oneself including one's feelings, perceptions, and senses in a nonself-critical manner while blocking out any thought processes that act as hindrances to present moment mediation [38]. ME involves a heightened awareness of one's bodily senses in response to the sight, taste, and sound of food as one eats by being attentive to one's salivary, olfactory, and visual responses and ceasing to eat when one feels full [39]. Details of the types of mindfulness-based therapies are beyond the scope of this study, but in short, the main therapies include mindfulness-based stress reduction (MBSR), mindfulness-based cognitive therapy (MBCT), dialectical behavior therapy, and acceptance and commitment therapy (ACT) [40-42]. Traditional MBSR comprises 8-week therapy sessions that include Hatha Yoga, walking and sitting meditation, body scans, and assignments at home through self-reflections coupled with group sessions [40-42]. MBCT is similar to MBSR with the addition of cognitive therapy [40]. Note that this study will focus specifically on MBSR and ME as the intervention has been found in both.

Our recent systematic review (manuscript under peer-review journal) [43], which focused on university students [44-69], found that that there were positive associations between mindfulness and physical activity, which included self-efficacy, overall levels, and time spent engaging in physical activity, though there were few studies, and most were limited by being cross-sectional. The review also found some support for a positive association among mindfulness and healthy eating behavior and diet including reduced binge eating, emotional eating, and some support for improved dietary intake of fruit and vegetables and reduced fat [43]. However, the review [43] found that were few studies on dietary intake, most studies were limited by being cross-sectional, and more longitudinal studies are needed [44-69].

Our review additionally found that there were very few mindfulness-based randomized controlled trials (RCTs) for weight loss in university students and that more high-quality trials are needed in tandem with trials that adopt diverse mindfulness approaches such as MBSR [43]. Research also suggests that combining approaches rather than using 1 single approach such as MBSR and ME is more successful [31] along with the fact that both formal and informal practices result in the greatest weight loss as described earlier [34].

Given that potential mindfulness holds for assisting with weight management, adoption of healthy weight-related behaviors, and stress reduction, promoting mindfulness in college students is important. However, research indicates that accessibility and availability to weight management services are limited across campuses and students are placed on waitlists for various counseling services [70,71]. As accessibility to general counseling services and weight management services is limited [70,71], it may be deduced that accessibility to special therapies like mindfulness on campus would be equally difficult to gain access to as well, although studies are needed to specifically explore access to mindfulness training on campus. Nonetheless, students do have busy schedules; therefore, developing an intervention that may be easily accessible to them at any time and place irrespective of their student roster and social calendar would be very helpful.

One novel platform for delivering weight loss and weight-related behavior change interventions has been mobile health (mHealth),

which includes the use of devices such as mobile phones personal digital assistants, tablets, and iPods [72]. mHealth holds numerous benefits, which include accessibility and portability to health care and promotion [73]. The author's previous published meta-analysis of mobile devices for weight loss found that they are effective in assisting with weight loss as a moderate significant effect size is exhibited (Cohen d=0.43; P<.05) [74]. None of these interventions had a mindfulness focus [74].

To date, there has not been a mobile mindfulness-based intervention that assessed the effectiveness of a mindfulness-based app or mindfulness-based short message service (SMS) text messages for weight management and healthy weight-related behaviors as well as stress in university students. Our second recent systematic review [75] reviewed the types of electronic mindfulness-based interventions for weight, eating, and stress [76-96], finding that there has not been a mobile mindfulness-based RCT that implemented an app or SMS text messages for weight management or weight-related health behaviors (diet and exercise) in university students. We identified 1 web-based intuitive eating intervention [93], which did not lead to body mass index (BMI) changes, as well as a multipurpose mHealth app with ME as 1 component, not focusing on mindfulness as a central theme [88]. There has also not been a mindfulness-based app that is tailored to university students including those at risk of the Freshman 15, which integrates ME and MBSR techniques.

Since our review and trial protocol registration, there have been a few studies on mHealth and mindfulness including an ACT app for eating behaviors and another mindfulness app for improving weight-related eating behaviors and physical activity in teenagers [97,98]. One found that using food as a reward was reduced in the app group [97]. Although the study did not find that stress moderated this relationship, the techniques used were applied on the basis of ACT and not traditional MBSR, which would have been valuable for assessing if stress moderated the effects [97]. The other study was a mindfulness app pilot with videos, which included some components of formal mindfulness such as walking meditation and the body scan, along with ME, as well as physical activity and stretching techniques [98]. The study was tested in 15 teens, finding that it helped them with their reported ME, the duration of their engagement in mindfulness practice, awareness of eating behaviors, as well as physical activity [98]. There was also a recent mobile phone study that involved mindfulness phone coaching, which found that it assisted with weight loss and was particularly beneficial for those with maladaptive eating behavior such as emotional eating [99].

This is a very new field and this trial study will contribute to expanding the literature. The novelty of this study is the combination of MBSR and ME in a student-tailored app RCT for university students for not only weight-related behaviors and weight but also stress.

Given that both mindfulness and mHealth have been found to assist with weight [31,32,34,35,74] and that university students experience challenges with issues that these interventions address [3,4], combining the 2 in a mindfulness-based mHealth intervention for weight, weight-related behaviors, and stress in the university student population could hold great potential and is needed. By targeting stress, which is a determinant of maladaptive weight-related behaviors in college students [18-30], students may potentially benefit more than simply using an ME app that does not target the root cause. Similarly, targeting stress alone through MBSR without teaching ME may also be limiting as research suggests that combination approaches are most effective [31]. Thus, this app will integrate ME and MBSR techniques. Developing and testing a tailored mindfulness-based app that meets the unique needs of university students are novel, and the app may hold potential as a medium for health promotion in this population. Previous reviews have found that only 5% of mindfulness apps in iTunes are truly based on mindfulness, and our review of ME apps also found that they were weak [100,101]. Therefore, this app will integrate proven evidence-based MBSR and ME techniques.

Aims and Objectives

The objective of this research was to develop and evaluate a mobile mindfulness-based app intervention that uses MBSR and ME techniques for weight loss, stress, and healthy weight-related behavior change in university students. The primary outcomes of interest are changes in weight and weight-related behaviors (eg, emotional eating and physical activity). Secondary outcomes of interest are changes in stress. Feasibility and acceptability will also be assessed.

Before the trial, the app was tested among the research group to ensure that its usability and functionality were intact. A subset of mindfulness-based messages had also been tested in a pilot exploratory qualitative pretrial study.

This study has the following main aims:

1. Determine if a novel mobile mindfulness-based app intervention is acceptable, feasible, and effective for weight loss and weight-related behavior change (eating behaviors and physical activity) when compared with a control receiving a self-monitoring electronic diary (e-diary) for diet and exercise with links to the World Health Organization (WHO) website
2. Determine if a novel mobile mindfulness-based app intervention reduces stress as a secondary outcome of interest when compared with the e-diary control

It is anticipated that the mindfulness app will be more effective for weight loss than the control e-diary as it targets stress-related weight behaviors in this specific population.

Methods

Participants and Setting

Students at the University of Queensland (UQ) at St Lucia and Herston campuses were recruited to participate in the Mindfulness App Weight Loss Trial of 11 weeks in duration. UQ is located in Brisbane, Australia.

Recruitment

From July 10, 2017 to August 14, 2017, undergraduate students were recruited for the study. Recruitment strategies for the trial

involved posters, flyers, social media, trial facts and UQ marketing and communications, and advertising via the website at UQ for the Centre for Online Health.

The recruitment material was categorized into general study information, which included posters and flyers on campus about the general study as well as detailed specific information that was in an information sheet. Trial facts advertised the information in a Web-based prescreener, and UQ marketing made a video about the trial and eligibility. The information sheet outlined the purpose of the study, the eligibility, the risks and benefits of participating, adverse event reporting, and rights to withdraw. Written informed consent was collected from participants before their enrollment in the study. This included consent to participate after participants read the information sheet with a clear description of the benefits, risks, including that the work will be published without releasing individual information. Participants were screened for eligibility in person and had signed the consent forms in person.

Incentives

Participants received an Aus $20 Merlo Coffee voucher for their participation in the study. They were also given a number for an iPad mini draw at the end of the study.

Eligibility

All healthy undergraduate students between 18 and 25 years of age who owned a mobile phone were eligible to participate if they wished to lose weight. Healthy was defined as anyone without a preexisting medical condition or history of serious psychiatric illness. Students <18 years or >25 years, pregnant women, and students with serious preexisting medical conditions were excluded. Students with a BMI<20 were also excluded. Students were asked if they had any serious medical conditions. Students with a preexisting medical condition were excluded.

Randomization

A computer-generated random number sequence randomized eligible participants to an intervention comprising a mindfulness app or to a control group receiving a self-monitoring e-diary for diet and exercise. A parallel simple randomization 1:1 sequence was undertaken.

Allocation was concealed to students before their entry into the mindfulness-based weight loss study, and a central random number generation method was used to conceal allocation from the principal investigator LNL and the research team. An independent statistician not involved with the study from the Centre for Online Health conducted the randomization process. Participants were notified of the group to which they were assigned, with instructions on downloading the app if assigned to the intervention.

Intervention Group: Mindfulness App

The intervention group received the mindfulness app (called My Student Mindfulness App), which contains ME and MBSR techniques. The app has been tailored to college students, with key themes on weight gain in college students (Freshman 15) and common college student stressors. Note that students tailored in this study are defined as being tailored specifically to the university student population, with student-relevant themes throughout rather than individual personal tailoring. The app seeks to educate, remind, prompt, and motivate students to practice ME [102] and stress reduction techniques [41]. MBSR techniques and meditation include body scan, diaphragmatic breathing, observing the breath, loving kindness meditation, concentration meditation, choiceless awareness mindfulness meditation, and Hatha yoga, which were adopted from key mindfulness books including elements from John Kabatt Zinn's MBSR program [41]. Several books were consulted to teach ME including those which cover mindfulness [103-106] and also specialist books on ME [103-108]. In brief, the app has audios and videos in addition to standard written educational content. It also has tailored practical tips and advice for achieving a mindful lifestyle and advices on addressing barriers and promoting facilitators to meditation when in university. Another feature that was added to the app was mindful exercise that focused on encouraging physical activity and movement from a mindfulness perspective [105].

The app home screen shot is attached in Multimedia Appendix 1. The key features of the app and the specific MBSR and ME techniques used are summarized in Multimedia Appendices 2 and 3. The app had organized written lectures and audios in order of difficulty and similarity of topics. For example, students were first introduced to the simple breathing exercises in the audios and articles before progressing gradually toward the more difficult and lengthy meditation techniques. Although students were informed that they could gradually learn how to meditate at their own pace, they were sent a supplementary document as a suggestion about how to use the app with recommended activities that they could work their way up to. This is attached as a supplementary multimedia file in Multimedia Appendix 4. The app also has push notification enabled functions. The push notifications aim to educate, prompt, and remind students to practice mindfulness. The messages were sent during eating times to maximize ME opportunities.

Control Group: E-Behavioral Self-Monitoring Diary

The control group received an electronic self-monitoring diary for its diet and exercise. The diary also had self-reflection space for students to think about key barriers they experience. They were also given links to WHO's dietary and physical activity guideline information in the e-diary.

Assessments and Measures

Participant characteristics and demographic information were collected at baseline. Anthropometric measures were assessed at 2 time points at baseline and at 11-week follow-up objectively by the candidate and volunteers. Height was self-reported. Weight was measured using a digital scale. Standard procedures were used to assess weight, which required participants to remove their shoes before being weighed. BMI was calculated as kg/m^2.

Physical activity was assessed using the International Physical Activity Questionnaire (IPAQ), comprising 27 questions on physical activity in the past week (IPAQ) [109]. A review of its validity and reliability across 12 countries found that it is both valid and reliable [109]. Previous college student weight

gain intervention studies have used the IPAQ to assess physical activity [88]. The short version [110] was used to minimize the time burden on students given all the other questionnaires they had to complete. Eating behavior was assessed using the 3-Factor Eating Behavior Questionnaire as it assesses cognitive restraint, emotional eating, and uncontrolled eating, and it has been found to be valid and reliable for these measures in adults [111]. Stress was assessed using the Perceived Stress Scale, a 10-item self-reported survey of perceptions of stress [112,113]. It has been previously validated and found to be reliable for assessing stress in college students in a study across 3 colleges [114]. Mindfulness was assessed using the Cognitive and Affective Mindfulness Revised Scale, which has been previously validated [115]. ME was assessed using the Mindful Eating Questionnaire (MEQ), which comprises 28 items that assess eating behaviors (68). The MEQ has been validated previously in college students [116].

Feasibility and Acceptability

Feasibility was assessed on the basis of participant retention and adherence. Participants were given a brief survey (Likert scale) upon completion of the study. It assessed adherence by asking participants how often they followed the app assignments through daily practice and whether they viewed all components of the app. The questionnaire asked whether they regularly used the app and practiced all the activities all the time, sometimes used it periodically, or did not use it at all.

The survey also assessed likability/acceptability of the intervention by assessing the degree of likability using a Likert scale. An extra questionnaire was added that asked about the most relevant meditation techniques used.

Ethics Approval

This study has been approved by the UQ Human Research Ethics Committee clearance number 2017000802).

Sample Size Calculation

As this is the first mobile mindfulness-based app intervention for weight and weight-related behaviors in university students, the anticipated effect size is not known. However, a sample size range may be estimated on the basis of certain assumptions. A previous mindfulness-based, Web-based stress reduction 2-arm RCT [89] estimated that a sample size of 50 in total would be needed for 80% power (alpha=.05) to detect a difference on the basis of a medium effect size in previous meta-analyses. The author's previous mHealth weight loss meta-analysis also found a medium effect size [74]. Assuming a medium effect size using a 1-side hypothesis (one-tailed) at alpha=0.05 and 80% power to detect an effect if one exists, a sample size of 58 would be required when allowing for a 10% dropout (for a 20% dropout, the size would need to be N=66) according to a consultation with a statistician. A sample size twice this size (N=114) would be required if one assumes a medium effect size at a lower end of the medium effect size distribution. Hence, this study aimed to recruit no less than 58 participants, with the goal of at least doubling this number as the target (N=114). Note that 80% power to detect a difference if one exists refers to a difference between both the app and control group in the main outcome of interest, which is weight change. A 1-tailed hypothesis was used as the hypothesis was directional, whereby we hypothesized that the app would be superior to the control e-diary for weight change. The sample size estimates are based on previous mindfulness traditional interventions for weight in tandem with previous mHealth interventions for weight.

Enrollment

At the end of the recruitment, a total of 90 students who met eligibility were enrolled in the study.

Data Analysis

Participant baseline characteristics were summarized using descriptive statistics (means, percentages, and standard deviation). To ensure that randomization was conducted correctly and that there were no significant differences between groups on any baseline characteristics, differences between groups were tested using a general linear model and independent sample two-tailed *t* tests (where data were nonnormally distributed or did not meet assumptions and nonparametric tests were undertaken) [117-119]. Residuals of the covariates were plotted to ensure that the data were normally distributed in tandem with checking for high correlations among covariates and undertaking Levene test for equality of variances. When assumptions were violated, data were transformed [120,121]. Differences in weight between the intervention and control were assessed using analysis of covariance (ANCOVA) where baseline weight values will be covariates [122]. The rational for this use is that Vickers and Altman argue that it is the preferred method when comparing differences between 2 groups in trials from baseline to follow-up over analysis of variance as it controls for baseline differences in a variable among groups (imbalance), which may influence the outcome [122]. Post hoc tests were conducted if the *F* statistic in the ANCOVA was significant for between-group differences in weight [119]. Changes in the other outcomes of interest were also assessed using ANCOVA. Categorical data were analyzed using the Chi-square test. Data were analyzed using SPSS 20 [123,124].

Reporting

This study was reported using the consolidated standards of reporting trials (CONSORT) guidelines. This trial was prospectively registered with the Australia and New Zealand Clinical Trials Registry ACTRN12616001349437.

Results

The CONSORT diagram outlines the recruitment process and retention is illustrated in Figure 1. Participant characteristics including baseline demographic and descriptive information are summarized in Table 1 for the 90 participants enrolled (45 in each group). The mean weight of all study participants was 76.29 kg (range 53-126 kg). The mean weight in the intervention group was 76.7 kg and 75.5 in the control. The mean BMI in the intervention was 26.09 kg/m^2 and 25.73 kg/m^3 in the control. The groups were comparable on baseline characteristics, and there were no significant differences in weight between the groups or BMI. The intervention group had slightly higher stress levels at baseline. Changes in key outcome variables at 11-week follow-up are summarized in Table 2.

Figure 1. Consolidated standards of reporting trials flow diagram. MAR: Missing at Random.

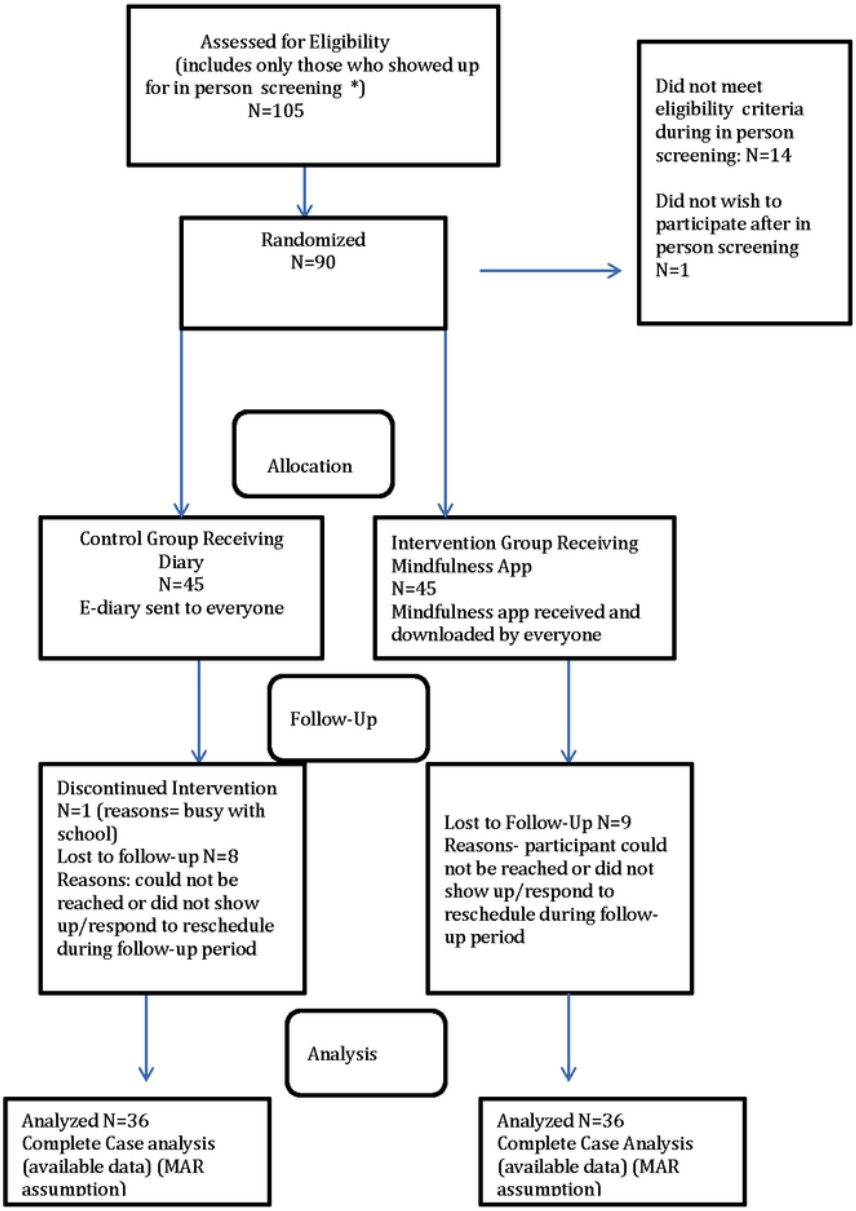


Table 1. Participant demographics and mean baseline anthropometrics.

Variable	Intervention app	Control e-diary	Total sample (N=90)
Age, mean; range	20.16	20.22	20.19; 18-24
Year	2.2	2.3	2.25=second year
Sex (female), %	74	61	67
Ethnicity (white), %	77	71	74
Weight, mean (SD)	76.4 (15.44)	76.18 (17.07)	76.29 (16.18)
Body mass index (kg/m^2), mean (SD); range	26.09 (4.795)	25.73 (4.75)	25.91 (4.74); 21-43
Stress[a], mean (SD)	18.18 (6.7)	18.82 (5.6)	18.5 (6.5)
Mindfulness[b], mean (SD)	24.29 (4.326)	25.56 (4.031)	24.92 (4.326)
Mindful eating[c], mean (SD)	2.588 (0.37)	2.639 (0.241)	2.61 (0.316)
Weight-related behaviors[d], mean (SD)			
Emotional Eating	7.78 (2.575)	6.64 (2.469)	7.21 (2.57)
Uncontrolled Eating	22.02 (4.490)	21.58 (4.33)	21.80 (4.394)
Cognitive Restraint	16.2 (3.9)	15.81 (3.8)	16 (3.87)
Physical activity[e], total minutes per week	173 (105)	156 (105)	164.5 (105)

[a]Perceived Stress Scale.

[b]Cognitive and Affective Mindfulness Questionnaire.

[c]Mindful Eating Questionnaire.

[d]Three Factor Eating Behavior Questionnaire.

[e]International Physical Activity Questionnaire (short).

Table 2. Analysis of covariance for differences between groups in key outcome variables.

Variable	Measure	*F* test	*P* value	Mean difference (I-C)	95% CI	Partial eta squared	R^2 (adjusted)
Weight	Digital scale	1.22 (df=1)	.27	0.76	−0.584 to 2.037	0.017	.971
		5.943 (df=1)[a,b]	.02[a,b]	−3.291[b]	−5.99 to −0.591[b]	0.090[b]	.326[b]
		2.775 (df=1)[c]	.10[c]	−2.151[c]	−4.727 to 0.425[c]	0.039[c]	.291[c]
Mindfulness	Cognitive and Affective Mindfulness Scale (revised)	14.580 (df=1)	<.001[a]	3.104[a]	1.482 to 4.726	0.174	.537
Mindful eating	Mindful Eating Questionnaire	21.035 (df=1)	<.0001[a]	0.295[a]	0.167 to 0.424	0.236	.402
Emotional eating	Three Factor Eating Behavior Questionnaire	5.893 (df=1)	.02[a]	−1.088	−1.98 to −0.194	0.079	.402
Uncontrolled eating	Three Factor Eating Behavior Questionnaire	5.974 (df=1)	.02[a]	−2.047	−3.717 to −0.376	0.080	.240
Cognitive restraint	Three Factor Eating Behavior Questionnaire	0.963 (df=1)	.330	−0.819	2.484 to 0.846	0.014	.290
MET[d] moderate (min/week)	IPAQ[e] (short)	7.885 (df=1)	.02[a]	−464[a]	−794 to −134	0.103	.113
Log transformed	IPAQ (short)	10.016 (df=1)	.01[a]	−4.07[a]	−0.671 to −0.143	0.278	.245
MET vigorous (min/week)	IPAQ (short)	0.737 (df=1)	.39	−306.933	−1020 to −406	0.394	.119
MET walk (min/week)	IPAQ (short)	0.095 (df=1)	.76	−76.150	−568 to 416	0.001	−.023
MET SUM	—	3.990 (df=1)	.05	−1006	−2011 to −1.296	0.055	.155
Log transformed	—	2.114 (df=1)	.15	0.127	−0.302 to −0.047	0.030	.103

[a]*P* value significant ≤.05.

[b]For adherers.

[c]Intention to treat.

[d]MET: metabolic equivalent.

[e]IPAQ: International Physical Activity Questionnaire.

Weight Changes

There were no statistically significant differences in weight between both the mindfulness app intervention and control e-diary groups at follow-up using the ANCOVA (*P*=.27). Although neither group lost weight, analyses of between-group changes and paired sample tests (within-group changes) results indicate that neither the mindfulness app group nor the electronic behavioral self-monitoring diary control group gained any weight when comparing weight change from baseline to follow-up within and between groups (*P*>.05).

Mindfulness

There were statistically significant differences in mindfulness between the mindfulness app intervention and control e-diary groups at 11 weeks using the ANCOVA (*F*=14.580(1); *P*<.01). After pairwise comparisons and adjustment for multiple comparisons using the Bonferroni correction, the intervention group had higher mindfulness levels by 3.104 on the mindfulness scale (*P*<.01; 95% CI 1.482-4.726; R^2=.537; partial eta=0.174).

Mindful Eating

There were statistically significant differences between both groups in ME levels at follow-up (*F*=21.035)(1); *P*<.01). The mindfulness app intervention group had higher ME levels by 0.295 points than the control e-diary group on the ME scale (*P*<.01; 95% CI 0.167-0.424; R^2=.402; partial eta=0.236).

Emotional Eating

It should be noted that at baseline, there were slightly higher levels of emotional eating in the mindfulness app intervention group compared with the control e-diary group (1.133; *P*<.04). However, ANCOVA controlled for these baseline differences, which may have influenced the outcome. The test on between-subjects effects using ANCOVA indicates that there is a significant difference between groups on emotional eating levels at follow-up (*F*=5.893 (1); *P*=.02). The pairwise comparison results indicate that the control group had significantly higher levels of emotional eating on the 3-Factor Eating Behavior Questionnaire by 1.088 points (*P*<.05; 95% CI −1.98 to −0.194), though the effect was small (adjusted R^2=.402; partial eta squared=0.079).

Uncontrolled Eating

The test on between-subjects effects using ANCOVA indicates that there is a significant difference between both groups on uncontrolled eating levels at follow-up (F=5.974(1); P=.02). The control e-diary group had higher uncontrolled eating levels than the intervention group by 2.047 points on the 3-Factor Eating Behavior scale (P<.05; 95% CI −3.717 to −0.376). Approximately 26% of the variability in uncontrolled eating levels is attributed to the intervention (R^2 and partial eta=0.26).

Cognitive Restraint

There were no statistically significant differences between both groups on cognitive restraint levels at follow-up (P=.33).

Stress

The mindfulness app intervention group had lower stress levels than the control e-diary group, but this was not significant in the intention-to-treat (ITT) analysis. The per-protocol analysis (after removing those who did not use the app) and 1 outlier (the student was under severe mental distress because of personal family circumstances), there were significant differences in stress levels between the groups (F=5.943; P=.02). Pairwise comparisons indicate that the control had stress levels that were 3.291 points higher on the Perceived Stress Scale (PSS) than the intervention group (P<.05; 95% CI 0.591-5.992; R^2=.326; partial eta squared=0.090).

Physical Activity

Metabolic Equivalent Mod

There were significant differences between both groups in terms of metabolic equivalent (MET) moderate physical activity levels (F=7.885; P=.02). The control e-diary group had higher MET moderate activity levels than the control by 464.233 min/week after controlling for baseline levels (P<.05; R^2=0.087; eta=0.103). However, Levene test for equality of variance did not meet assumptions, indicating that the results were less reliable. The residuals were also skewed. The data were log transformed and MET moderate vigorous physical activity levels remained significantly higher in the control e-diary group than in the app group (F=10.016; P=.01); R^2=0.245; partial eta=.278) mean difference −0.407 (−0.671 to −0.143).

Metabolic Equivalent Vigorous

There were no statistically significant differences in mean MET vigorous activity levels between both groups at follow-up (P=.39).

Metabolic Equivalent Walk

There was no significant difference in MET walking levels between both groups (P=.76).

Metabolic Equivalent Total

There was a marginally significant difference between groups in the total MET level (P=.05). The control e-diary group had higher overall MET levels min/week by 1,007 min/week. However, the normality assumptions of the residual and Levene test for equality of variance were not met. The data were log transformed and there were no significant differences between groups (F=2.114; P=.15).

Differences in Physical Activity Level Categories

The Chi-square test was undertaken to test whether there were differences between the groups in the proportion of individuals in low, moderate, and vigorous physical activity categories. There were no significant differences between the mindfulness app intervention and control e-diary groups (P=.41).

Mindfulness Meditation Practice Exercises

Participants in the app group listed several mindfulness meditation exercises that were most useful (up to three). The most useful was the observation of the breath mediation exercise (39% [13/33] reported it being useful). The second most useful exercise was walking meditation (27 %, 9/33), which was found to be helpful. The third-most useful exercise was diaphragmatic breathing (21%, 7/33) and concentration meditation (21%, 7/33). A total of 12% (4/33) found the loving kindness meditation exercises to be useful. The least useful was the choiceless awareness mindfulness meditation exercise as only 3% (1/33) found it to be useful. A total of 9% (3/33) of students found all the exercises to be useful and 3% (1/33) did not find any of the exercises to be useful.

Retention

A total of 72 out of 90 participants (80% retention) returned for the follow-up.

Acceptability/Likability

A total of 94% (32/34) of students reported that they liked the app overall and found it to be acceptable (either liked it, very much liked it, or somewhat), whereas 6% (2/34) did not like the app. Examples of mindfulness-based messages that were sent to students in the intervention are summarized in Textbox 1. The script is available in Multimedia Appendix 5.

Adherence

A total of 14% (5/34) of the students in the mindfulness app intervention group reported that they reviewed all the content and used it on a regular basis, whereas 61% (21/34) reported using it periodically. The remaining 23% (8/34) of students reported that they very seldom engaged with the app and did not review much of its content.

Feasibility

The app appears to be somewhat feasible as most students found it acceptable and the loss to follow-up was small. However, regular adherence was low.

Textbox 1. Sample mindfulness-based messages that were sent to the participants.

Theme: stress

- What issues are on your mind at the moment? Identify any stressors and visually image that these stressors are like bubbles that you acknowledge by gently touching. Imagine these stressors dissipating like bubbles.

Theme: mindful awareness of body

- Try and take a mindful bite or spoonful for lunch/dinner today. What does it taste like, smell like, feel like, and sound like? Chew the food slowly. Mindfully stop when your body indicates you are full.

  Fun eating tip: Fill your bowl with fruit that includes all the colors of the rainbow. Really taste the differences.

Theme: mindful awareness of body

- A person takes 16 breaths per minute and 23,040 a day. Take a mindful breath, really focusing on the physical feeling of inhaling and exhaling and not thinking about anything else.

Theme: formal practice

- Long practice: Did you practice your body scan yet? Try practicing the body scan by tuning in with your body and how it is feeling from the head to the toe. Breathe in awareness to those areas.

  Short practice: Take a moment for a 3 min breathing space. Focus your attention on the breath, breathing as you naturally do in and out. Allow any thoughts, feelings, or sensations to enter your awareness. Gently hold them without judgment. Then gently return your attention to your breath.

  Walking meditation is a great way to connect with your body. Find some space and walk slowly, really feeling your movement.

Theme: informal practice or present moment awareness

- Having cues throughout the day can help you with being mindful. Students identified the lake as being a mindful moment spot. Go checkout the lake-the choice of the ideal mindfulness spot selected by students.

Discussion

Principal Findings

This study is important because it is the first RCT that has assessed the feasibility and effectiveness of a student-tailored mindfulness app for stress, weight management, and weight-related behaviors. As mentioned earlier, previous reviews critiqued commercial mindfulness apps for their lack of use of rigorous evidence-based, mindfulness-based techniques along with issues relating to quality of content and esthetics [100,101]. The app was novel as it combined MBSR evidence-based techniques with ME techniques that were tailored to students with special student-relevant themes throughout the app. The app was also novel because of the use of mindfulness-based messages for reinforcing weight loss, stress reduction, and healthy eating patterns as well as incorporation of mindfulness of exercise in the components.

In summary, the mindfulness app did not assist with weight loss over the semester and there were not any significant differences in weight between the mindfulness app intervention group and the behavioral self-monitoring e-diary group. The app assisted with some weight-related behaviors including emotional eating and binge eating as well as with increasing mindfulness, ME, and reducing stress in adherers. The e-diary group had higher moderate-to-vigorous physical activity levels.

Although the mindfulness app did not help with weight loss, students did not gain weight over the semester in both the app and the electronic behavioral self-monitoring diary groups. Behavioral self-monitoring is an already established method of weight management in the literature [125]. The fact that students who are a high-risk group for weight gain over the course of their semester did not gain weight suggests that both an e-diary for behavioral self-monitoring and a mindfulness app could potentially be used to prevent weight gain. Thus, the mindfulness app is comparable with an e-diary to prevent weight gain or ensure weight maintenance in students. More studies are needed to confirm this relationship. It would be of interest to study students entering their first semester of studies to see if these methods could be effective for preventing the Freshman 15 phenomenon or not. It is also important to study whether students should have the option of either type of intervention, a combined one, or a selection on the basis of their individual needs by exploring the stress versus motivation hypothesis for instance.

In addition to this, there were small albeit significant improvements in emotional eating and uncontrolled eating in the mindfulness app of the intervention group. This is in agreement with a recent mHealth ACT app trial, which found that there were increases in eating for physical rather than emotional reasons in study participants [97].

There was also a small effect of the app on stress in the per-protocol analysis, suggesting that the app has potential for assisting with stress and stress-induced eating behavior. The effect sizes were small, but it is likely that they relate to the low levels of regular adherence in the mindfulness app group. A previous study on mindfulness found that stress was reduced in those who adhered to the app for stress [81]. Addressing the problem of low regular adherence in this student population is

needed. A further study is planned to better understand the barriers and facilitators of use to improve adherence to the app.

Since this study, there have been 2 studies in students using a mindfulness app for stress [126,127]. One 7-week intervention found that it significantly reduced stress [126], which is in agreement with this study, whereas the 4-week intervention found that it only improved emotional well-being [127]. It could be that the short length of the intervention, which is half the length of MBSR interventions [40], influenced the outcomes on stress, though the researchers mentioned that prespecified health issues may have influenced the result [127].

Early changes in eating behaviors and stress−related eating behaviors can take time before they have an effect on actual weight loss as maladaptive weight−related eating behaviors are indeed linked to weight gain and obesity [36,37]. Hence, it would be plausible that we may see weight loss in this population if studied over a longer period of time, especially if adherence will become more regular. It is also possible that it may take longer to observe weight change in mindfulness mHealth studies when compared with traditional behavior change mHealth studies as there have not been any previous comparisons in mHealth behavior change versus mHealth mindfulness apps for weight loss.

Mindful exercise was included as a module in the app with the aim of increasing physical activity. Mindfulness of exercise is an emerging field, which focuses on being present when one moves one's body [105,128]. A previous systematic review found some evidence for higher physical activity levels in more mindful students than in their less mindful counterparts [43]. Overall physical activity increased in the mindfulness groups. However, the behavioral self−monitoring control e-diary group exhibited higher levels of MET moderate physical activity in min/week upon follow-up. This is consistent with previous findings as electronic self−monitoring has already been shown to increase physical activity levels in the literature [129,130]. It may be of research interest to develop a mindfulness app that combines elements of behavioral self-monitoring as our app only had a mindfulness practice self-monitoring diary rather than a diet and exercise self-monitoring diary.

In addition to this, we found that mindfulness levels and ME increased in the mindfulness app intervention group. These findings are in agreement with the findings of 2 mindfulness mHealth trials [98,99]. The recent pilot in teens, which used a mindfulness app, also found a high level of engagement in ME, but it was limited to being a pilot; therefore, there were no comparisons overtime and between groups [98]. The mindfulness phone coaching study also found that it helped with ME and mindfulness, though with only 1 mindfulness subset [99]. The study also did not find improvements in weight [99].

The most useful mindfulness meditation exercise was observing the breath, whereas the least useful was Choiceless Awareness Mindfulness Meditation. This makes sense as observing the breath is a very short 5-min exercise whereby one focuses on one's breathing for up to 5 min in stillness [103,104]. It could be that simple, short meditation techniques are preferred in this population. Students also enjoyed walking meditation, which might seem to be practical on a green campus setting during breaks, as well as diaphragmatic breathing, which is longer than observing the breath and involves changing one's breathing pace to a deeper one [104]. Other relevant exercises were concentration meditation and loving kindness, which are generally practiced for 15-30 min [104]. By contrast, the Choiceless Awareness Meditation is the most difficult to undertake according to Dr. Jon Kabat-Zinn as one must extend one's awareness beyond the breath, to the body, and one's thoughts without becoming attached to them [103,104]. This is practiced for up to 45 min [104].

In terms of feasibility, we found the mindfulness app to be somewhat feasible. The retention in the study was 80%, which is relatively high. Overall, students liked the app and found it acceptable. As mentioned earlier, more information is needed to find out ways to improve adherence by ensuring regular usage.

Another interesting part of this study involved the use of mindfulness-based SMS text messages, which was a novelty. There was 1 recent study that used these messages to encourage mindfulness practice in depressed patients, though the results were mixed as they did not work for all patients in terms of feasibility [131]. This study found that participants enjoyed the intervention, and more information on message perceptions will be published in an upcoming posttrial paper. More studies on mindfulness-based SMS text messages for weight are needed in the field.

Strengths and Limitations

A strength is that this is the first RCT in the field to explore these relationships as our previous systematic review did not find such a study [75]. This study adds to the evidence base under the fields of mindfulness and mHealth (mobile mindfulness weight interventions). This is the first study that has also used mindfulness-based push notification messages. This study also measured mindfulness and ME, which is a strength over past studies [32] as it demonstrates that the app actually increased mindfulness.

A limitation is that this study was only for 11 weeks, and it may take longer to see the effects. However, there have been previous mHealth studies for weight loss that were of comparable duration [132,133] and still reported weight loss. Furthermore, traditional MBSR interventions are typically only 8 weeks in duration, indicating that this study's duration was significantly longer [40,103]. Moreover, many mHealth MBSR interventions for stress were of similar short duration [75]. Two studies were only 1-month long and still found improvements in stress [96], with the latter finding improvement after the study [90]. The rationale for undertaking a study that was just under a semester in duration was to ensure that students would return for follow-up during the examination period as the holiday break to the new year could have potentially led to a significant loss to follow-up.

In addition, we did not have a standard no intervention control as originally planned as ethics had advised to offer 2 interventions for both groups. Furthermore, although we ensured allocation concealment, we did not ensure blinding. Ideally,

study personnel as well as participants would be blinded to their intervention group [134]. It is possible that participants in both groups may have learned off each other. Another limitation involves the use of the *International Physical Activity Questionnaire*-Short Form, which has been found to be less reliable than the long form [135]. However, the short form was the most appropriate in our study given the sheer number of other measures.

Another minor limitation was receiving self-reported height rather than measuring height in person. However, students knew their height from their common IDs such as drivers' licenses. Research also indicates that the BMI may not always be the best indicator of weight given that muscle increases BMI and that it does not assess fat [136]. Weight was objectively measured in person, which is the strength of this study.

In terms of randomization, it was established that simple parallel randomization would be appropriate for this study rather than stratifying for clinical risk factors, which may affect outcomes [137] in this healthy population. This method also did not require the statistician who undertook randomization to have access to participant study data characteristics, ensuring complete independence from the study.

We also did not use imputation methods such as Last Observation Carried Forward to estimate the analyses without any dropouts as part of the ITT; however, given the relatively high retention and the inherent limitations of these methods [138], we did not feel it was necessary. Moreover, according to the 4-point ITT strategy [139], complete case analysis that involves analyzing the available data in our type of study design and form of analysis is acceptable if we assume that the dropouts were missing at random [139].

Future Directions

Future studies could further explore mHealth apps and mindfulness in relation to stress, weight, and weight-related behaviors over longer periods of time. In addition, we briefly explored push notification messages as well as mindfulness SMS text messages toward the end of the study that aimed to motivate students to use the app and practice their techniques. This has not been previously undertaken in any study. It would also be of interest to further conduct follow-up mindfulness-based SMS text message interventions for stress, weight, and weight-related behaviors such as simple observation of the breath audios sent as MPEG-1 Audio Layer 3 files with motivational mindfulness-based messages. It may also be of interest to combine traditional behavioral self-monitoring SMS text messages for weight and mindfulness-based messages.

Conclusions

In summary, this is the first mHealth mindfulness app RCT that has assessed the effectiveness of a student-tailored mindfulness app for stress, weight-related behaviors, and weight when compared with traditional, electronic behavioral self-monitoring of diet and exercise. We did not find that the app was effective for weight loss. However, neither the mindfulness app intervention group nor the control e-diary group gained weight over the course of the semester. We also found that the mindfulness app significantly assisted with stress, emotional eating, and uncontrolled eating relative to the control, though the effect sizes were small. We also found that the app increased ME and overall levels of mindfulness. The e-diary control group had higher levels of moderate MET min/week activity levels than the app group. We conclude that the mindfulness app holds promise for weight-related lifestyle behaviors related to stress and stress eating, but more studies are needed to confirm these relationships. Low regular adherence was an issue in this study and more studies are needed that can explore whether these relationships will have greater effects when used on a regular basis. Focus group studies are needed to better understand key barriers and facilitators of usage. Longer studies are also needed to study whether stress-induced eating behavior changes may lead to weight loss over longer periods of time during a student's studies. Furthermore, studies should examine whether a mindfulness app and electronic behavior self-monitoring via diary are effective measures for preventing weight gain, and possibly the "Freshman 15."

Acknowledgments

The authors would like to thank statistician Dr Michael David for assisting with randomization and Pascal Burger for his assistance with the app videos. This study received funding from Australia Research Training Scheme and UQ centennial Scholarship and the UQ Medicine Research Grant.

Conflicts of Interest

None declared.

Multimedia Appendix 1

App Home screenshot.

[PDF File (Adobe PDF File), 93KB - mhealth_v7i4e12210_app1.pdf]

Multimedia Appendix 2

Mindfulness techniques in the app.

[PDF File (Adobe PDF File), 111KB - mhealth_v7i4e12210_app2.pdf]

Multimedia Appendix 3

Mindfulness app media content.

[PDF File (Adobe PDF File), 116KB - mhealth_v7i4e12210_app3.pdf]

Multimedia Appendix 4

App suggestions.

[PDF File (Adobe PDF File), 119KB - mhealth_v7i4e12210_app4.pdf]

Multimedia Appendix 5

Mindfulness-based message options script.

[PDF File (Adobe PDF File), 107KB - mhealth_v7i4e12210_app5.pdf]

Multimedia Appendix 6

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 545KB - mhealth_v7i4e12210_app6.pdf]

References

1. World Health Organization. World Health Organization. Obesity and Overweight Fact Sheet URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2019-02-10] [WebCite Cache ID 764u5kfU4]
2. Mihalopoulos NL, Auinger P, Klein JD. The Freshman 15: is it real? J Am Coll Health 2008;56(5):531-533 [FREE Full text] [doi: doi: 10.3200/JACH.56.5.531-534] [Medline: 18400665]
3. Crombie A, Ilich J, Dutton G, Panton L, Abood D. The freshman weight gain phenomenon revisited. Nutr Rev 2009 Feb;67(2):83-94. [doi: 10.1111/j.1753-4887.2008.00143.x] [Medline: 19178649]
4. Vadeboncoeur C, Townsend N, Foster C. A meta-analysis of weight gain in first year university students: is freshman 15 a myth? BMC Obes 2015;2:22 [FREE Full text] [doi: 10.1186/s40608-015-0051-7] [Medline: 26217537]
5. Vella-Zarb R, Elgar F. The 'freshman 5': a meta-analysis of weight gain in the freshman year of college. J Am Coll Health 2009;58(2):161-166. [doi: 10.1080/07448480903221392] [Medline: 19892653]
6. Edmonds MJ, Ferreira KJ, Nikiforuk EA, Finnie AK, Leavey SH, Duncan AM, et al. Body weight and percent body fat increase during the transition from high school to university in females. J Am Diet Assoc 2008 Jun;108(6):1033-1037. [doi: 10.1016/j.jada.2008.03.002] [Medline: 18502240]
7. Levitsky DA, Halbmaier CA, Mrdjenovic G. The freshman weight gain: a model for the study of the epidemic of obesity. Int J Obes Relat Metab Disord 2004 Nov;28(11):1435-1442. [doi: 10.1038/sj.ijo.0802776] [Medline: 15365585]
8. Brown W, Hockey R, Dobson A. Effects of having a baby on weight gain. Am J Prev Med 2010;38(2):163-170. [doi: 10.1016/j.amepre.2009.09.044Get]
9. Drenowatz C, Cai B, Hand GA, Katzmarzyk PT, Shook RP, Blair SN. Prospective association between body composition, physical activity and energy intake in young adults. Eur J Clin Nutr 2016 Apr;70(4):482-487. [doi: 10.1038/ejcn.2015.133] [Medline: 26286639]
10. Gordon-Larsen P, Hou N, Sidney S, Sternfeld B, Lewis CE, Jacobs DR, et al. Fifteen-year longitudinal trends in walking patterns and their impact on weight change. Am J Clin Nutr 2009 Jan;89(1):19-26 [FREE Full text] [doi: 10.3945/ajcn.2008.26147] [Medline: 19056560]
11. Racette SB, Deusinger SS, Strube MJ, Highstein GR, Deusinger RH. Weight changes, exercise, and dietary patterns during freshman and sophomore years of college. J Am Coll Health 2005;53(6):245-251. [doi: 10.3200/JACH.53.6.245-251] [Medline: 15900988]
12. American Institute of Stress.: American Stress Society; 2018 Jul 02. Stress in College Students URL: https://www.stress.org/college-students [accessed 2019-02-10] [WebCite Cache ID 764omshRe]
13. Wu J. Stanford Daily News. Stanford , California ,USA: The Stanford Daily; 2000 Feb 11. Rising Levels of Stress Alarm Health Educators URL: https://stanforddailyarchive.com/cgi-bin/stanford?a=d&d=stanford20000229-01.2.51&e=-------en-20--21--txt-txIN-student+army+training+corps------ [accessed 2019-02-10]
14. Boyce J, Kuijer R. Perceived stress and freshman weight change: the moderating role of baseline body mass index. Physiol Behav 2015 Feb;139:491-496. [doi: 10.1016/j.physbeh.2014.12.011] [Medline: 25484356]
15. Economos C, Hildebrandt ML, Hyatt R. College freshman stress and weight change: differences by gender. Am J Health Behav 2008;32(1):16-25. [doi: 10.5555/ajhb.2008.32.1.16] [Medline: 18021030]
16. Epel E, Jimenez S, Brownell K, Stroud L, Stoney C, Niaura R. Are stress eaters at risk for the metabolic syndrome? Ann N Y Acad Sci 2004 Dec;1032:208-210. [doi: 10.1196/annals.1314.022] [Medline: 15677412]

17. Serlachius A, Hamer M, Wardle J. Stress and weight change in university students in the United Kingdom. Physiol Behav 2007 Nov 23;92(4):548-553. [doi: 10.1016/j.physbeh.2007.04.032] [Medline: 17537466]
18. Barker ET, Williams RL, Galambos NL. Daily spillover to and from binge eating in first-year university females. Eat Disord 2006;14(3):229-242. [doi: 10.1080/10640260600639079] [Medline: 16807216]
19. Costarelli V, Patsai A. Academic examination stress increases disordered eating symptomatology in female university students. Eat Weight Disord 2012 Sep;17(3):e164-e169. [Medline: 23086251]
20. Cruz S, Fabián C, Pagán I, Ríos JL, González AM, Betancourt J, et al. Physical activity and its associations with sociodemographic characteristics, dietary patterns, and perceived academic stress in students attending college in Puerto Rico. P R Health Sci J 2013 Mar;32(1):44-50. [Medline: 23556265]
21. El Ansari W, Berg-Beckhoff G. Nutritional correlates of perceived stress among university students in Egypt. Int J Environ Res Public Health 2015 Nov 06;12(11):14164-14176 [FREE Full text] [doi: 10.3390/ijerph121114164] [Medline: 26561825]
22. El Ansari W, Suominen S, Berg-Beckhoff G. Mood and food at the University of Turku in Finland: nutritional correlates of perceived stress are most pronounced among overweight students. Int J Public Health 2015 Sep;60(6):707-716. [doi: 10.1007/s00038-015-0717-4] [Medline: 26202822]
23. Kandiah J, Yake M, Jones J, Meyer M. Stress influences appetite and comfort food preferences in college women. Nutr Res 2006 Mar 01;26(3):118-123. [doi: 10.1016/j.nutres.2005.11.010]
24. Liu C, Xie B, Chou CP, Koprowski C, Zhou D, Palmer P, et al. Perceived stress, depression and food consumption frequency in the college students of China Seven Cities. Physiol Behav 2007 Nov 23;92(4):748-754. [doi: 10.1016/j.physbeh.2007.05.068] [Medline: 17585967]
25. Mikolajczyk R, El Ansari W, Maxwell A. Food consumption frequency and perceived stress and depressive symptoms among students in three European countries. Nutr J 2009 Jul 15;8:31 [FREE Full text] [doi: 10.1186/1475-2891-8-31] [Medline: 19604384]
26. Nguyen-Michel S, Unger JB, Hamilton J, Spruijt-Metz D. Associations between physical activity and perceived stress/hassles in college students. Stress Health 2006 Aug;22(3):179-188. [doi: 10.1002/smi.1094]
27. Papier K, Ahmed F, Lee P, Wiseman J. Stress and dietary behaviour among first-year university students in Australia: sex differences. Nutrition 2015 Feb;31(2):324-330. [doi: 10.1016/j.nut.2014.08.004] [Medline: 25442361]
28. Steptoe A, Wardle J, Pollard TM, Canaan L, Davies GJ. Stress, social support and health-related behavior: a study of smoking, alcohol consumption and physical exercise. J Psychosom Res 1996 Aug;41(2):171-180. [Medline: 8887830]
29. Sulkowski ML, Dempsey J, Dempsey AG. Effects of stress and coping on binge eating in female college students. Eat Behav 2011 Aug;12(3):188-191. [doi: 10.1016/j.eatbeh.2011.04.006] [Medline: 21741016]
30. Wolff GE, Crosby RD, Roberts JA, Wittrock DA. Differences in daily stress, mood, coping, and eating behavior in binge eating and nonbinge eating college women. Addict Behav 2000;25(2):205-216. [doi: 10.1016/S0306-4603(99)00049-0] [Medline: 10795945]
31. O'Reilly GA, Cook L, Spruijt-Metz D, Black DS. Mindfulness-based interventions for obesity-related eating behaviours: a literature review. Obes Rev 2014 Jun;15(6):453-461 [FREE Full text] [doi: 10.1111/obr.12156] [Medline: 24636206]
32. Olson KL, Emery CF. Mindfulness and weight loss: a systematic review. Psychosom Med 2015 Jan;77(1):59-67. [doi: 10.1097/PSY.0000000000000127] [Medline: 25490697]
33. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J Altern Complement Med 2009 May;15(5):593-600. [doi: 10.1089/acm.2008.0495] [Medline: 19432513]
34. Carrière K, Khoury B, Günak MM, Knäuper B. Mindfulness-based interventions for weight loss: a systematic review and meta-analysis. Obes Rev 2018 Dec;19(2):164-177. [doi: 10.1111/obr.12623] [Medline: 29076610]
35. Dunn C, Haubenreiser M, Johnson M, Nordby K, Aggarwal S, Myer S, et al. Mindfulness approaches and weight loss, weight maintenance, and weight regain. Curr Obes Rep 2018 Mar;7(1):37-49. [doi: 10.1007/s13679-018-0299-6] [Medline: 29446036]
36. Koenders PG, van Strien T. Emotional eating, rather than lifestyle behavior, drives weight gain in a prospective study in 1562 employees. J Occup Environ Med 2011 Nov;53(11):1287-1293. [doi: 10.1097/JOM.0b013e31823078a2] [Medline: 22027541]
37. Masheb RM, Lutes LD, Kim HM, Holleman RG, Goodrich DE, Janney CA, et al. High-frequency binge eating predicts weight gain among veterans receiving behavioral weight loss treatments. Obesity (Silver Spring) 2015 Jan;23(1):54-61 [FREE Full text] [doi: 10.1002/oby.20931] [Medline: 25385705]
38. Mayo Clinic Staff. Mayo Clinic.: Mayo Clinic Mindfulness Exercises URL: https://www.mayoclinic.org/healthy-lifestyle/consumer-health/in-depth/mindfulness-exercises/art-20046356 [accessed 2019-02-10] [WebCite Cache ID 764pYjeoO]
39. Tribole E. Intuitive Eating.: Intuitive Eating; 2010 Nov 17. The Difference Between Intuitive Eating and Mindful Eating URL: https://www.intuitiveeating.org/the-difference-between-intuitive-eating-and-mindful-eating/ [accessed 2019-02-10] [WebCite Cache ID 764pzCVCE]
40. Chiesa A, Malinowski P. Mindfulness-based approaches: are they all the same? J Clin Psychol 2011 Apr;67(4):404-424. [doi: 10.1002/jclp.20776] [Medline: 21254062]
41. Santorelli SF. Center for Mindfulness , Health Care, & Society University of Massachusetts Medical School. 2014. Mindfulness-Based Stress Reduction (MBSR): Standards of Practice URL: https://www.umassmed.edu/contentassets/

24cd221488584125835e2eddce7dbb89/mbsr_standards_of_practice_2014.pdf [accessed 2019-02-14] [WebCite Cache ID 76B0lokmf]
42. Sipe WE, Eisendrath SJ. Mindfulness-based cognitive therapy: theory and practice. Can J Psychiatry 2012 Feb;57(2):63-69. [doi: 10.1177/070674371205700202] [Medline: 22340145]
43. Lyzwinski LN, Caffery L, Bambling M, Edirippulige S. Relationship between mindfulness, weight, and weight-related behaviors in college students: a systematic review. Altern Complement Ther 2018 Oct;24(5):202-214. [doi: 10.1089/act.2018.29182.lnl]
44. Anderson A, Caine-Bish N, Gordon K, Falcone T. Effects of a mindful eating intervention on acute calorie and macronutrient intake in college students. J Nutr Educ Behav 2015 Jul;47(4):S88. [doi: 10.1016/j.jneb.2015.04.232]
45. Anderson DA, Schaumberg K, Anderson LM, Reilly EE. Is level of intuitive eating associated with plate size effects? Eat Behav 2015 Aug;18:125-130. [doi: 10.1016/j.eatbeh.2015.05.005] [Medline: 26042920]
46. Anderson LM, Reilly EE, Schaumberg K, Dmochowski S, Anderson DA. Contributions of mindful eating, intuitive eating, and restraint to BMI, disordered eating, and meal consumption in college students. Eat Weight Disord 2016 Mar;21(1):83-90. [doi: 10.1007/s40519-015-0210-3] [Medline: 26243300]
47. Anglin J. Assessing the effectiveness of intuitive eating for weight loss-pilot study. Nutr Health 2012 Apr;21(2):107-115. [doi: 10.1177/0260106012459994] [Medline: 23139388]
48. Bahl S, Milne GR, Ross SM, Chan K. Mindfulness: a long-term solution for mindless eating by college students. J Public Policy Mark 2013 Sep;32(2):173-184. [doi: 10.1509/jppm.11.008]
49. Bodenlos JS, Wells SY, Noonan M, Mayrsohn A. Facets of dispositional mindfulness and health among college students. J Altern Complement Med 2015 Oct;21(10):645-652. [doi: 10.1089/acm.2014.0302] [Medline: 26352341]
50. Bryan S. Mindfulness and nutrition in college age students. J Basic App Sci 2016;12:68-74.
51. Butryn M, Forman E, Hoffman K, Shaw J, Juarascio A. A pilot study of acceptance and commitment therapy for promotion of physical activity. J Phys Act Health 2011 May;8(4):516-522. [Medline: 21597124]
52. Cavanagh K, Vartanian L, Herman C, Polivy J. The effect of portion size on food intake is robust to brief education and mindfulness exercises. J Health Psychol 2014 Jun;19(6):730-739. [doi: 10.1177/1359105313478645] [Medline: 23471762]
53. Gao Y, Shi L. Mindfulness, physical activity and avoidance of secondhand smoke: a study of college students in shanghai. Int J Environ Res Public Health 2015 Aug 21;12(8):10106-10116 [FREE Full text] [doi: 10.3390/ijerph120810106] [Medline: 26308029]
54. Gast J, Campbell Nielson A, Hunt A, Leiker J. Intuitive eating: associations with physical activity motivation and BMI. Am J Health Promot 2015;29(3):e91-e99. [doi: 10.4278/ajhp.130305-QUAN-97] [Medline: 24459999]
55. Gast J, Madanat H, Nielson AC. Are men more intuitive when it comes to eating and physical activity? Am J Mens Health 2012 Mar;6(2):164-171. [doi: 10.1177/1557988311428090] [Medline: 22105066]
56. Gilbert D, Waltz J. Mindfulness and health behaviors. Mindfulness 2010 Oct 23;1(4):227-234. [doi: 10.1007/s12671-010-0032-3]
57. Grinnell S, Greene G, Melanson K, Blissmer B, Lofgren I. Anthropometric and behavioral measures related to mindfulness in college students. J Am Coll Health 2011;59(6):539-545. [doi: 10.1080/07448481.2011.555932] [Medline: 21660809]
58. Hawks S, Madanat H, Hawks J, Harris A. The relationship between intuitive eating and health indicators among college women. Am J Health Educ 2013 Feb 25;36(6):331-336. [doi: 10.1080/19325037.2005.10608206]
59. Herbert B, Blechert J, Hautzinger M, Matthias E, Herbert C. Intuitive eating is associated with interoceptive sensitivity. Effects on body mass index. Appetite 2013 Nov;70:22-30. [doi: 10.1016/j.appet.2013.06.082] [Medline: 23811348]
60. Humphrey L, Clifford D, Neyman Morris MN. Health at every size college course reduces dieting behaviors and improves intuitive eating, body esteem, and anti-fat attitudes. J Nutr Educ Behav 2015;47(4):354-60.e1. [doi: 10.1016/j.jneb.2015.01.008] [Medline: 25769516]
61. Katterman SN, Goldstein SP, Butryn ML, Forman EM, Lowe MR. Efficacy of an acceptance-based behavioral intervention for weight gain prevention in young adult women. J Contextual Behav Sci 2014 Jan;3(1):45-50. [doi: 10.1016/j.jcbs.2013.10.003]
62. Mantzios M, Wilson JC. Making concrete construals mindful: a novel approach for developing mindfulness and self-compassion to assist weight loss. Psychol Health 2014;29(4):422-441. [doi: 10.1080/08870446.2013.863883] [Medline: 24215123]
63. Moor KR, Scott AJ, McIntosh WD. Mindful eating and its relationship to body mass index and physical activity among university students. Mindfulness 2012 Aug 8;4(3):269-274. [doi: 10.1007/s12671-012-0124-3]
64. Murphy MJ, Mermelstein LC, Edwards KM, Gidycz CA. The benefits of dispositional mindfulness in physical health: a longitudinal study of female college students. J Am Coll Health 2012;60(5):341-348. [doi: 10.1080/07448481.2011.629260] [Medline: 22686356]
65. Roberts KC, Danoff-Burg S. Mindfulness and health behaviors: is paying attention good for you? J Am Coll Health 2010;59(3):165-173. [doi: 10.1080/07448481.2010.484452] [Medline: 21186446]
66. Slonim J, Kienhuis M, Di Benedetto M, Reece J. The relationships among self-care, dispositional mindfulness, and psychological distress in medical students. Med Educ Online 2015;20:27924 [FREE Full text] [doi: 10.3402/meo.v20.27924] [Medline: 26112354]

67. Smith TS, Hawks SR. Intuitive eating, diet composition, and the meaning of food in healthy weight promotion. American Journal of Health Education 2013 Jan 23;37(3):130-136. [doi: 10.1080/19325037.2006.10598892]
68. Framson C, Kristal AR, Schenk JM, Littman AJ, Zeliadt S, Benitez D. Development and validation of the mindful eating questionnaire. J Am Diet Assoc 2009;109(8):1439-1444. [doi: 10.1016/j.jada.2009.05.006] [Medline: 19631053]
69. Webb JB, Hardin AS. A preliminary evaluation of BMI status in moderating changes in body composition and eating behavior in ethnically-diverse first-year college women. Eat Behav 2012 Dec;13(4):402-405. [doi: 10.1016/j.eatbeh.2012.06.004] [Medline: 23121798]
70. Lynch S, Hayes S, Napolitano M, Hufnagel K. Availability and accessibility of student-specific weight loss programs and other risk prevention health services on college campuses. JMIR Public Health Surveill 2016 Jun 08;2(1):e29 [FREE Full text] [doi: 10.2196/publichealth.5166] [Medline: 27278261]
71. Novotney A. American Psychological Association. 2014. Students under pressure URL: https://www.apa.org/monitor/2014/09/cover-pressure
72. World Health Organization. Global Observatory for E Health Series. 2011. mHealth New Horizons for Health Through Mobile Technologies URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf
73. Lefebvre C. Integrating cell phones and mobile technologies into public health practice: a social marketing perspective. Health Promot Pract 2009 Oct;10(4):490-494. [doi: 10.1177/1524839909342849] [Medline: 19809002]
74. Lyzwinski LN. A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis. J Pers Med 2014 Jun 30;4(3):311-385 [FREE Full text] [doi: 10.3390/jpm4030311] [Medline: 25563356]
75. Lyzwinski LN, Caffery L, Bambling M, Edirippulige S. A systematic review of electronic mindfulness-based therapeutic interventions for weight, weight-related behaviors, and psychological stress. Telemed J E Health 2018 Dec;24(3):173-184. [doi: 10.1089/tmj.2017.0117] [Medline: 28885896]
76. Warnecke E, Quinn S, Ogden K, Towle N, Nelson MR. A randomised controlled trial of the effects of mindfulness practice on medical student stress levels. Med Educ 2011 Apr;45(4):381-388. [doi: 10.1111/j.1365-2923.2010.03877.x] [Medline: 21401686]
77. Spadaro K, Hunker D. Exploring The effects of an online asynchronous mindfulness meditation intervention with nursing students On Stress, mood, And Cognition: A descriptive study. Nurse Educ Today 2016 Apr;39:163-169. [doi: 10.1016/j.nedt.2016.02.006] [Medline: 27006051]
78. Räsänen P, Lappalainen P, Muotka J, Tolvanen A, Lappalainen R. An online guided ACT intervention for enhancing the psychological wellbeing of university students: A randomized controlled clinical trial. Behav Res Ther 2016 Mar;78:30-42. [doi: 10.1016/j.brat.2016.01.001] [Medline: 26848517]
79. Pflugeisen BM, Drummond D, Ebersole D, Mundell K, Chen D. Brief video-module administered mindfulness program for physicians: a pilot study. Explore (NY) 2016;12(1):50-54. [doi: 10.1016/j.explore.2015.10.005] [Medline: 26725471]
80. O'Leary K, Dockray S. The effects of two novel gratitude and mindfulness interventions on well-being. J Altern Complement Med 2015 Apr;21(4):243-245. [doi: 10.1089/acm.2014.0119] [Medline: 25826108]
81. Muuraiskangas S, Harjumaa M, Kaipainen K, Ermes M. Process and effects evaluation of a digital mental health intervention targeted at improving occupational well-being: lessons from an intervention study with failed adoption. JMIR Ment Health 2016 May 11;3(2):e13 [FREE Full text] [doi: 10.2196/mental.4465] [Medline: 27170553]
82. Morledge TJ, Allexandre D, Fox E, Fu AZ, Higashi MK, Kruzikas DT, et al. Feasibility of an online mindfulness program for stress management--a randomized, controlled trial. Ann Behav Med 2013 Oct;46(2):137-148 [FREE Full text] [doi: 10.1007/s12160-013-9490-x] [Medline: 23632913]
83. Mak WW, Chan AT, Cheung EY, Lin CL, Ngai KC. Enhancing Web-based mindfulness training for mental health promotion with the health action process approach: randomized controlled trial. J Med Internet Res 2015 Jan 19;17(1):e8 [FREE Full text] [doi: 10.2196/jmir.3746] [Medline: 25599904]
84. Ly K, Asplund K, Andersson G. Stress management for middle managers via an acceptance and commitment-based smartphone application: A randomized controlled trial. Internet Interv 2014 Jul;1(3):95-101. [doi: 10.1016/j.invent.2014.06.003]
85. Levin ME, Pistorello J, Seeley JR, Hayes SC. Feasibility of a prototype web-based acceptance and commitment therapy prevention program for college students. J Am Coll Health 2014;62(1):20-30 [FREE Full text] [doi: 10.1080/07448481.2013.843533] [Medline: 24313693]
86. Krusche A, Cyhlarova E, King S, Williams JM. Mindfulness online: a preliminary evaluation of the feasibility of a web-based mindfulness course and the impact on stress. BMJ Open 2012;2(3) [FREE Full text] [doi: 10.1136/bmjopen-2011-000803] [Medline: 22614170]
87. Kemper KJ, Lynn J, Mahan JD. What is the impact of online training in mind-body skills? J Evid Based Complementary Altern Med 2015 Oct;20(4):275-282. [doi: 10.1177/2156587215580882] [Medline: 25962637]
88. Gow RW, Trace SE, Mazzeo SE. Preventing weight gain in first year college students: an online intervention to prevent the "freshman fifteen". Eat Behav 2010 Jan;11(1):33-39 [FREE Full text] [doi: 10.1016/j.eatbeh.2009.08.005] [Medline: 19962118]
89. Glück TM, Maercker A. A randomized controlled pilot study of a brief web-based mindfulness training. BMC Psychiatry 2011 Nov 08;11:175 [FREE Full text] [doi: 10.1186/1471-244X-11-175] [Medline: 22067058]

90. Drozd F, Raeder S, Kraft P, Bjørkli CA. Multilevel growth curve analyses of treatment effects of a web-based intervention for stress reduction: randomized controlled trial. J Med Internet Res 2013 Apr 22;15(4):e84 [FREE Full text] [doi: 10.2196/jmir.2570] [Medline: 23607962]
91. Cavanagh K, Strauss C, Cicconi F, Griffiths N, Wyper A, Jones F. A randomised controlled trial of a brief online mindfulness-based intervention. Behav Res Ther 2013 Sep;51(9):573-578. [doi: 10.1016/j.brat.2013.06.003] [Medline: 23872699]
92. 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]
93. Boucher S, Edwards O, Gray A, Nada-Raja S, Lillis J, Tylka TL, et al. Teaching intuitive eating and acceptance and commitment therapy skills via a web-based intervention: a pilot single-arm intervention study. JMIR Res Protoc 2016 Oct 14;5(4):e180 [FREE Full text] [doi: 10.2196/resprot.5861] [Medline: 27742602]
94. Allexandre D, Bernstein AM, Walker E, Hunter J, Roizen MF, Morledge TJ. A web-based mindfulness stress management program in a corporate call center: a randomized clinical trial to evaluate the added benefit of onsite group support. J Occup Environ Med 2016 Mar;58(3):254-264 [FREE Full text] [doi: 10.1097/JOM.0000000000000680] [Medline: 26949875]
95. Aikens K, Astin J, Pelletier K, Levanovich K, Baase C, Park Y, et al. Mindfulness goes to work: impact of an online workplace intervention. J Occup Environ Med 2014 Jul;56(7):721-731. [doi: 10.1097/JOM.0000000000000209] [Medline: 24988100]
96. 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]
97. Järvelä-Reijonen E, Karhunen L, Sairanen E, Muotka J, Lindroos S, Laitinen J, et al. The effects of acceptance and commitment therapy on eating behavior and diet delivered through face-to-face contact and a mobile app: a randomized controlled trial. Int J Behav Nutr Phys Act 2018 Dec 27;15(1):22 [FREE Full text] [doi: 10.1186/s12966-018-0654-8] [Medline: 29482636]
98. Turner T, Hingle M. Evaluation of a mindfulness-based mobile app aimed at promoting awareness of weight-related behaviors in adolescents: a pilot study. JMIR Res Protoc 2017 Apr 26;6(4):e67 [FREE Full text] [doi: 10.2196/resprot.6695] [Medline: 28446423]
99. Carpenter K, Vickerman K, Salmon E, Javitz H, Epel E, Lovejoy J. A randomized pilot study of a phone-based mindfulness and weight loss program. Behav Med 2017 Oct 06:1-11 (forthcoming). [doi: 10.1080/08964289.2017.1384359] [Medline: 28985151]
100. Lyzwinski LN, Caffery LC, Bambling M, Edirippulige S. A critical appraisal of mindful eating apps. - 2018 Under Review.
101. Mani M, Kavanagh D, Hides L, Stoyanov S. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
102. Somov P. Eating the Moment. Mindful Practices to Overcome Overeating One Meal at a Time. Oakland, California: New Harbinger Publications; 2008:141.
103. Kabat-Zinn J. Mindfulness for Beginners: Reclaiming the Present Moment and Your Life. Colorado: Sounds True; 2016.
104. Labee E. Psychology Moment by Moment: A Guide to Enhancing Your Practice with Mindfulness Meditation. Oakland, California: New Harbinger Publications; 2011.
105. Stahl B, Goldstein E. A Mindfulness-Based Stress Reduction Workbook. Oakland, California: New Harbinger Publications; 2010.
106. Stahl B, Goldstein E. MBSR Everyday: Daily Practices from the Heart of MBSR. Oakland, California: New Harbinger Publications; 2015.
107. Albers S. Mindful Eating 101: A Guide to Healthy Eating in College and Beyond. London: Routledge; 2006.
108. Tribole E. Intuitive Eating: A revolutionary Program that Works. New York: Griffin publishing (St Martin's Press); 2012:320.
109. 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]
110. International Physical Activity Questionnaire. 2001. IPAQ: Short Last 7 Days Self Adminsitered Format URL: http://www.sdp.univ.fvg.it/sites/default/files/IPAQ_English_self-admin_short.pdf [accessed 2019-02-10] [WebCite Cache ID 7652ENZcv]
111. Cappelleri J, Bushmakin A, Gerber R, Leidy N, Sexton C, Lowe M, et al. Psychometric analysis of the Three-Factor Eating Questionnaire-R21: results from a large diverse sample of obese and non-obese participants. Int J Obes (Lond) 2009 Jun;33(6):611-620. [doi: 10.1038/ijo.2009.74] [Medline: 19399021]
112. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
113. Lee EH. Review of the psychometric evidence of the perceived stress scale. Asian Nurs Res (Korean Soc Nurs Sci) 2012 Dec;6(4):121-127 [FREE Full text] [doi: 10.1016/j.anr.2012.08.004] [Medline: 25031113]
114. Roberti J, Harrington LN, Storch EA. Further psychometric support for the 10-item version of the perceived stress scale. J Coll Couns 2006;9(2):135-147. [doi: 10.1002/j.2161-1882.2006.tb00100.x]

115. Feldman G, Hayes A, Kumar S, Greeson J, Laurenceau J. Mindfulness and emotion regulation: the development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R). J Psychopathol Behav Assess 2006 Nov 7;29(3):177-190. [doi: 10.1007/s10862-006-9035-8]
116. Hulbert-Williams L, Nicholls W, Joy J, Hulbert-Williams N. Initial validation of the Mindful Eating Scale. Mindfulness 2013 Jun 10;5(6):719-729. [doi: 10.1007/s12671-013-0227-5]
117. Hatcher L. Step by Step Guide to Basic Statistics Using SAS. In: SAS Institute. North Carolina: SAS Institute Inc; 2003.
118. Ntoumanis N. A Step by Step Guide to SPSS for Sport and Exercise Studies. London: Routledge; 2011.
119. O'rourke R, Hatcher L, Stepanski EJ. A Step by Step Guide to Using SAS for Univariate and Multivariate Statistics , Second Edition. North Carolina: SAS Institute Inc; 2005.
120. Yockey R. SPSS Demystified : A Step by Step Guide to Successful Data Analysis. Second Edition. New Jersey: Prentice Hall; 2011.
121. Rothwell Joanne. University of Sheffield.: Stats Tutor; 2017. ANCOVA URL: https://www.sheffield.ac.uk/polopoly_fs/1.531229!/file/MASH_ANCOVA_SPSS.pdf [accessed 2019-02-10] [WebCite Cache ID 7652qENyA]
122. Vickers AJ, Altman DG. Analysing controlled trials with baseline and follow up measurements. Br Med J 2001;323(7321):1123-1124. [doi: 10.1136/bmj.323.7321.1123 (]
123. PAWS Statistics (formerly SPSS). 2009. PASW Statistics for Windows, ; Version 18.0. C URL: http://www.spss.com.hk/statistics/ [accessed 2019-02-10] [WebCite Cache ID 76536Pc7Y]
124. SAS. Statiscal Analysis Software (SAS).: SAS Institute Incorporated; 2004. URL: https://www.sas.com/en_au/software/stat.html [accessed 2019-02-10] [WebCite Cache ID 7653VCWBy]
125. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
126. Hall B, Xiong P, Guo X, Sou E, Chou U, Shen Z. An evaluation of a low intensity mHealth enhanced mindfulness intervention for Chinese university students: A randomized controlled trial. Psychiatry Res 2018 Dec;270:394-403. [doi: 10.1016/j.psychres.2018.09.060] [Medline: 30300870]
127. Lee RA, Jung ME. Evaluation of an mHealth app (DeStressify) on university students' mental health: pilot trial. JMIR Ment Health 2018 Jan 23;5(1):e2 [FREE Full text] [doi: 10.2196/mental.8324] [Medline: 29362209]
128. Blair Kennedy A, Resnick PB. Mindfulness and physical activity. Am J Lifestyle Med 2015 Jan 06;9(3):221-223. [doi: 10.1177/1559827614564546]
129. Conroy MB, Yang K, Elci OU, Gabriel KP, Styn MA, Wang J, et al. Physical activity self-monitoring and weight loss: 6-month results of the SMART trial. Med Sci Sports Exerc 2011 Aug;43(8):1568-1574 [FREE Full text] [doi: 10.1249/MSS.0b013e31820b9395] [Medline: 21200337]
130. Gleeson-Kreig JM. Self-monitoring of physical activity: effects on self-efficacy and behavior in people with type 2 diabetes. Diabetes Educ 2006;32(1):69-77. [doi: 10.1177/0145721705284285] [Medline: 16439495]
131. Kraft S, Wolf M, Klein T, Becker T, Bauer S, Puschner B. Text message feedback to support mindfulness practice in people with depressive symptoms: a pilot randomized controlled trial. JMIR Mhealth Uhealth 2017 May 02;5(5):e59 [FREE Full text] [doi: 10.2196/mhealth.7095] [Medline: 28465278]
132. Napolitano M, Hayes S, Bennett G, Ives A, Foster G. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31 [FREE Full text] [doi: 10.1002/oby.20232] [Medline: 23505165]
133. Turner-McGrievy G, Campbell M, Tate D, Truesdale K, Bowling J, Crosby L. Pounds Off Digitally study: a randomized podcasting weight-loss intervention. Am J Prev Med 2009 Oct;37(4):263-269 [FREE Full text] [doi: 10.1016/j.amepre.2009.06.010] [Medline: 19765496]
134. Higgins J, Altman D, Gøtzsche PC, Jüni P, Moher D, Oxman A, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Br Med J 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
135. Lee P, Macfarlane D, Lam T, Stewart S. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8:115 [FREE Full text] [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
136. Nuttall FQ. Body mass index: obesity, BMI, and health: a critical review. Nutr Today 2015 May;50(3):117-128 [FREE Full text] [doi: 10.1097/NT.0000000000000092] [Medline: 27340299]
137. Kernan W, Viscoli C, Makuch R, Brass L, Horwitz R. Stratified randomization for clinical trials. J Clin Epidemiol 1999 Jan;52(1):19-26. [Medline: 9973070]
138. Lachin JM. Fallacies of last observation carried forward analyses. Clin Trials 2016 Apr;13(2):161-168 [FREE Full text] [doi: 10.1177/1740774515602688] [Medline: 26400875]
139. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clin Trials 2012 Aug;9(4):396-407 [FREE Full text] [doi: 10.1177/1740774512450098] [Medline: 22752633]

Abbreviations

ACT: acceptance and commitment therapy
ANCOVA: analysis of covariance
BMI: body mass index
CONSORT: consolidated standards of reporting trials
e-diary: electronic diary
IPAQ: International Physical Activity Questionnaire
ITT: intention-to-treat
MBCT: mindfulness-based cognitive therapy
MBSR: mindfulness-based stress reduction
ME: mindful eating
MET: metabolic equivalent
MEQ: Mindful Eating Questionnaire
mHealth: mobile health
RCT: randomized controlled trial
SMS: short message service
UQ: University of Queensland
WHO: World Health Organization

Edited by G Eysenbach; submitted 15.09.18; peer-reviewed by L Becker, S Boucher, D George; comments to author 29.11.18; revised version received 07.12.18; accepted 29.12.18; published 10.04.19.

Please cite as:
Lyzwinski LN, Caffery L, Bambling M, Edirippulige S
The Mindfulness App Trial for Weight, Weight-Related Behaviors, and Stress in University Students: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(4):e12210
URL: https://mhealth.jmir.org/2019/4/e12210/
doi:10.2196/12210
PMID:30969174

Original Paper

Neighborhood Deprivation and the Effectiveness of Mobile Health Coaching to Improve Periconceptional Nutrition and Lifestyle in Women: Survey in a Large Urban Municipality in the Netherlands

Dionne V Gootjes[1], MD; Matthijs R van Dijk[1], MD, PhD; Maria PH Koster[1], MD, PhD; Sten P Willemsen[2], MSc; Eric AP Steegers[1], MD, PhD; Régine PM Steegers-Theunissen[1], MD, PhD

[1]Department of Obstetrics and Gynecology, Erasmus Medical Center, University Medical Center Rotterdam, Rotterdam, Netherlands

[2]Department of Biostatistics, Erasmus Medical Center, University Medical Center Rotterdam, Rotterdam, Netherlands

Corresponding Author:
Régine PM Steegers-Theunissen, MD, PhD
Department of Obstetrics and Gynecology
Erasmus Medical Center
University Medical Center Rotterdam
Doctor Molewaterplein 40
Rotterdam, 3015 GD
Netherlands
Phone: 31 107043598
Fax: 31 107036815
Email: r.steegers@erasmusmc.nl

Abstract

Background: In 2011, we launched the Smarter Pregnancy mobile health (mHealth) coaching program, which has shown to effectively improve inadequate nutrition and lifestyle behaviors in women before and during pregnancy. It is known that in deprived neighborhoods, risk factors for adverse pregnancy outcomes like inadequate nutrition and lifestyle behaviors accumulate. However, it has not yet been investigated whether the Smarter Pregnancy program is equally effective in women living in deprived neighborhoods.

Objective: This paper aimed to study the associations between neighborhood deprivation and improvement of inadequate nutrition and lifestyle behaviors of women who were either contemplating pregnancy or already pregnant and subscribed to the Smarter Pregnancy program.

Methods: We performed an additional analysis on data from women who used the Smarter Pregnancy program from 2011 to 2016. The program comprised 24 weeks of coaching on 5 nutrition and lifestyle behaviors, of which adequate intakes or lifestyle behaviors were defined as an intake of 200 grams or above of vegetables, 2 pieces of fruit, daily folic acid supplement use of 400 µg per day, and no smoking or alcohol consumption. Neighborhood deprivation was determined according to the status scores of the Netherlands Institute for Social Research. Logistic regression analyses and generalized estimating equation models were used to assess the associations between the neighborhood status score (NSS) and the improvement of inadequate nutrition and lifestyle behaviors, taking into account the behaviors at baseline. We adjusted the analyses for maternal age, body mass index, geographic origin, pregnancy status, and participation as a couple.

Results: Of the 2554 women included, 521 participated with their male partner. Overall, daily vegetable intake was most frequently inadequate at the start of the program (77.72, 1985/2554). Women with a higher NSS (ie, nondeprived neighborhood) smoked less often (adjusted odds ratio [OR] 0.85; 95% CI 0.77-0.93), consumed alcohol more often (adjusted OR 1.14, 95% CI 1.04-1.24), and were less likely to complete the 24 weeks of coaching (OR 0.91, 95% CI 0.88-0.95) compared with women who lived in a neighborhood with a low NSS (ie, deprived). In the total group, the relative improvement of inadequate nutrition and lifestyle behaviors after 24 weeks of coaching was between 26% and 64%. NSS was negatively associated with this improvement, indicating that women with a higher NSS were less likely to improve inadequate nutrition and lifestyle behaviors, especially vegetable intake (adjusted OR 0.89, 95% CI 0.82-0.97).

Conclusions: The Smarter Pregnancy mHealth coaching program empowers women to improve inadequate nutrition and lifestyle behaviors. Unexpectedly, the program seemed more effective in women living in deprived neighborhoods. It is important to

unravel differences in needs and behaviors of specific target groups to further tailor the mHealth program on the basis of demographic characteristics like neighborhood deprivation.

(JMIR Mhealth Uhealth 2019;7(4):e11664) doi:10.2196/11664

KEYWORDS

pregnancy; telemedicine; lifestyle; nutritional status; preconception care

Introduction

Background

Worldwide, there are substantial differences in perinatal morbidity and mortality rates between and within countries, which may indicate inequalities in perinatal as well as population health [1,2]. Several underlying factors can explain these differences such as maternal-specific (eg, age, body mass index, BMI, and parity), environmental (eg, air pollution and extreme temperature), and community-derived (eg, housing conditions and poverty) factors [3-6]. As in other countries, perinatal morbidity and mortality rates in the Netherlands also differ among districts, with particularly high mortality rates in the country's 4 largest cities. This is mainly because of the large number of deprived neighborhoods in these cities [7-9].

Risk factors for adverse pregnancy outcomes, such as poor nutrition, lifestyle and housing conditions as well as lower health literacy, often accumulate in residents of deprived neighborhoods [6,9,10]. However, living in a deprived neighborhood itself has also been described as an independent risk factor for poor health outcomes [11]. Exposure to the abovementioned risk factors during the periconception period (ie, the 14 weeks before conception until 10 weeks after conception) [12,13] can have a detrimental effect on maternal and neonatal outcome. Moreover, on the longer term, the effect of these adverse outcomes is not limited to perinatal health; it also extends to the child's health later in life [14,15]. Therefore, it is important to change inadequate nutrition and lifestyle behaviors during the periconception period.

According to the transtheoretical model of behavioral change, intentional behavioral change can be achieved after passing 6 different stages, from precontemplation to maintenance and termination [16].

However, behavioral change is more challenging for individuals who have limited health literacy or impaired financial resources, who are less educated, and live in more deprived neighborhoods [3,17,18]. From this background, we hypothesize that women who live in more deprived neighborhoods are less likely to improve inadequate nutrition and lifestyle behaviors before and during pregnancy compared with women who live in less deprived neighborhoods.

Currently, mobile health (mHealth) apps are widely available and used for health improvement. mHealth apps can be designed to a specific population and target of interest and may be offered anytime and anywhere at low costs. Therefore, mHealth is a promising medium to support people to improve nutrition and lifestyle behaviors [19,20]. In 2011, after more than 30 years of research on the impact of nutrition and lifestyle behaviors on reproduction, we developed and launched the Smarter Pregnancy mHealth coaching program [21] for women, together with their male partners, who are contemplating pregnancy or are already pregnant [22,23]. Smarter Pregnancy is a Web-based program that can be used on a mobile device, comprising screening questions, thereafter comprising personal coaching through short message service (SMS) and email (Multimedia Appendix 1).

Previously, van Dijk et al analyzed survey data of all subscribers to the Smarter Pregnancy program to assess compliance, feasibility, usability, and first effectiveness of the program. It was shown that the program contributes to significant improvements of inadequate nutrition and lifestyle behaviors—that is, vegetable (+26%) and fruit intake (+38%), folic acid supplement use (+56%), smoking (−35%), and alcohol consumption (−42%)—in couples before and during pregnancy, which also resulted in an enhanced pregnancy chance in both fertile and subfertile couples up to 40%. Besides, a high compliance (65%) and usability were reported [22,24].

Objectives

However, it has not yet been investigated whether the Smarter Pregnancy program is equally effective in women who live in deprived neighborhoods. Therefore, our current aim was to investigate in an additional analysis, associations between neighborhood deprivation and the improvement of inadequate nutrition and lifestyle behaviors of women before and during pregnancy who previously subscribed to the Smarter Pregnancy program.

Methods

Study Design

We used the data of an epidemiological survey conducted among all women who subscribed to the Smarter Pregnancy program for an additional analysis [21]. Women and their partners living in the Netherlands were invited to subscribe to the Smarter Pregnancy program. Inclusion criteria were the following: aged between 18 and 45 years, an active wish to contemplate pregnancy, pregnant less than 13 weeks, the possession of a mobile phone with internet access, and a sufficient knowledge or understanding of the Dutch language. For male partners, the same inclusion criteria had to be met but without an upper age limit. Registration to this mHealth program was recommended to patients who visited the Division of Reproductive Medicine of the Department of Obstetrics and Gynecology of the Erasmus Medical Centre (MC) and to women who attended a community midwife in the Rotterdam area. However, as the website had an open access policy, other visitors were able to register. Although women could either participate alone or together with their male partner, because of the small sample size of couples, in this

study, we only analyzed data from women, and participation as a couple was taken into account as a covariate.

The coaching model developed for the Smarter Pregnancy program is based on the most recent knowledge on the effect of vegetable, fruit, and folic acid supplement intake, smoking and alcohol consumption on pregnancy chance, course, and outcome [18,25,26]. For the content of the platform, the stage of the model of Prochaska and Diclemente's was taken into account, which describes the readiness for behavioral change. This was implemented by informing participants about the positive effects of adequate nutrition and lifestyle behaviors on pregnancy course and outcome, which could affect the readiness to improve these behaviors [16]. Characteristics of the attitude, social influence, and self-efficacy model were implemented by enabling individuals as well as their partners to improve behavior [27]. Fogg's behavior model was applied by including triggers throughout the program to support motivation and thereby increase the ability to change nutrition and lifestyle behaviors [28]. Furthermore, the Smarter Pregnancy program meets the highest rules of legislation for medical devices in Europe, and it received the Conformité Européenne, classe 1 classification (2013). Effectiveness of the program has previously been demonstrated and described by van Dijk et al [24].

Intervention

The coaching program starts with a baseline screening on nutrition (ie, vegetable and fruit intake and folic acid supplement use) and lifestyle (ie, smoking and alcohol consumption) behaviors that significantly affect fertility and pregnancy course and outcome [24]. The mHealth coaching lasts for a period of 24 weeks and only targets the nutrition and lifestyle behaviors that are inadequate at the start of the program. Coaching comprises a maximum of 3 interventions per week, comprising SMS text messages and email messages containing recommendations, vouchers, and seasonal recipes. Follow-up screening takes place at 6, 12, 18, and 24 weeks after registration (Multimedia Appendix 2). Besides nutrition and lifestyle behaviors, there are additional questions addressing pregnancy status and BMI. The technical programming is executed by Peercode BV. A detailed description of the content of the Smarter Pregnancy program has previously been published by van Dijk et al [22].

Data Collection

Data were collected through the Smarter Pregnancy program itself. Demographic characteristics and anthropometric measurements of the participants were retrieved from the Smarter Pregnancy database—zip code, sex (male or female), age (continuous), pregnancy status (pregnant or not pregnant), and BMI (calculated from self-reported height and weight). Geographic origin was not reported by participants themselves. Therefore, we used the surnames of the participants to ascribe them a geographic origin, a method that is considered valid when self-identification is not available [29]. Classification was performed by 3 investigators (DG, MRD, and MPHK) who separately categorized all participants' surnames into 2 groups, that is, Western (Europe, excluding Turkey, North America, Oceania, Indonesia, and Japan) and non-Western (Africa, Latin America, Asia excluding Indonesia and Japan, and Turkey) origin. Any disagreement was resolved by discussion among the 3 investigators, which was the case within 7.6% (195/2554) of the surnames.

Outcomes

Compliance to the Smarter Pregnancy program was defined as the percentage of participants who filled in the last questionnaire of the program after 24 weeks of coaching. At baseline and after 24 weeks of coaching, the reported nutrition and lifestyle behaviors were classified as adequate or inadequate. Adequate behavior was defined as a daily intake of at least 200 grams of vegetables and at least 2 pieces of fruit, daily folic acid supplement use of at least 400 µg starting before conception and lasting until the 12th week of pregnancy, and no smoking or alcohol consumption [30].

To adjust for nutrition and lifestyle behaviors, a total risk score (TRS) was calculated. For vegetable and fruit intake, folic acid supplement, and alcohol use, 0 points were assigned in case a participant had an adequate intake or use [24]. For inadequate intake or use, 3 points were assigned. For smoking, 6 points were assigned in case of inadequate use, because of its known strong negative impact on pregnancy course and outcome [25,31]. Consequently, TRS in this study ranges from 0 (most adequate) to 18 (most inadequate).

To assign participants a neighborhood deprivation state, the status scores of the Netherlands Institute for Social Research were used. These scores follow a standard normal distribution by design and are calculated for all 4-digit zip codes in the Netherlands on the basis of 4 neighborhood characteristics—the average income, the number of nonemployed residents, the number of lower educated residents, and the number of households with a low income [32]. When the neighborhood status score (NSS) is low, this indicates a deprived neighborhood. A high NSS indicates a nondeprived neighborhood [33]. Since 1998, NSS is calculated every 4 years. For this study, the NSS of the year 2014 was used to determine the classification of the neighborhood participants lived in while using the Smarter Pregnancy program. In 2014, the interquartile range (IQR) of the NSS in the Netherlands was –0.57 to 0.71.

Data Analysis

All participants who started the program were included in the analysis at baseline. However, improvement of nutrition and lifestyle behaviors was only examined in those individuals who scored inadequate at any of these behaviors at the start of the program. To minimize selection bias, multiple imputation using chained equations was performed to handle missing data of women who prematurely resigned from the program. For those women, it was assumed that the adequacy of their nutrition and lifestyle behaviors at the last reported screening moment would not have changed until the end of the program (24 weeks).

Univariate linear and logistic regression analysis was used to study associations between demographic characteristics of the study population (maternal age, BMI, geographic origin, pregnancy status, whether a woman participated as a couple or alone, and TRS) and the NSS at the start of the program. Logistic regression analysis was used to examine the association

between the NSS and (in) adequate nutrition and lifestyle behaviors at the start of the program. To study the improvement of inadequate nutrition and lifestyle behaviors after 24 weeks of coaching, generalized estimating equations with an independent working correlation matrix were used to model the fraction of the study population that scored inadequate at baseline, taking into account that less improvement may be expected when less women show inadequate behavior at baseline. Interaction tests were performed to study interactions of geographic origin, participation as a couple, or being pregnant at the start of the program on the association between NSS and nutrition and lifestyle behaviors.

Statistical analyses were performed using SPSS version 21 software package (IBM Corp) and R version 3.4 (Foundation for Statistical Computing). $P<.05$ values were considered statistically significant. No alpha adjustment for multiple comparisons was made.

Ethical Approval

Details of ethical approval included the following: This survey was conducted according to the guidelines laid down in the Declaration of Helsinki. All procedures involving patients were approved by the Medical Ethical and Institutional Review Board of the Erasmus MC, University Medical Centre, Rotterdam, the Netherlands (MEC-2011–524, approved on 22 December 2011). Digital informed consent was obtained from all participants.

Results

General Characteristics

A total of 3776 women registered to the Smarter Pregnancy program, out of which 32.36% (1222/3776) of the women were excluded because of absence of activating the registration, incomplete registration, or incomplete data entry at the start of the program (Figure 1). Consequently, a total of 2554 women were included in the analysis, out of which 521 participated with their male partner. The median age of women at the start of the program was 31 years and most women were of Western geographic origin (72.91% (1862/2554)). Of all nutrition and lifestyle behaviors, daily vegetable intake was most frequently inadequate at the start of the program (77.72%(1985/2554) Table 1).

Women with a higher NSS (ie, who lived in a less deprived neighborhood) were older (beta=.04; 95% CI 0.03-0.05) and more often participated as a couple (beta=.18; 95% CI 0.11-0.25). Moreover, these women were more often pregnant at the start of the program (beta=−.30; 95% CI −0.41 to −0.19), had a lower BMI (beta=−.03; 95% CI −0.04 to −0.02), and were less often of non-Western geographic origin (beta=−0.78; 95% CI −0.85 to −0.70; Table 2).

Compliance to the Smarter Pregnancy program was 68.17% (1741/2554; Figure 1). Women with a higher NSS were less likely to finish the 24 weeks of coaching (odds ratio [OR] 0.91, 95% CI 0.88-0.95).

Figure 1. Flowchart of study participants that completed or resigned from the Smarter Pregnancy mobile health coaching program.

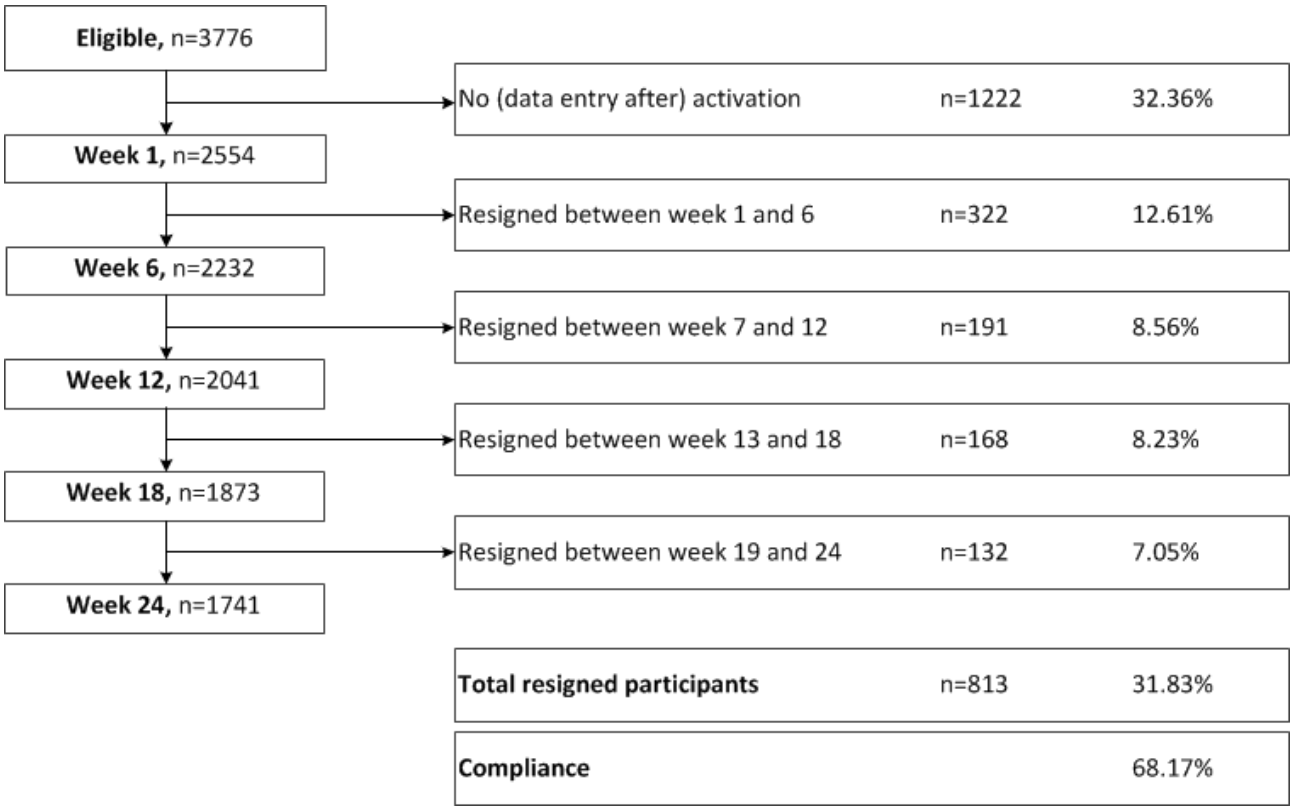


Table 1. Demographics of the study population and nutrition and lifestyle behaviors at the start and after 24-weeks of coaching with the Smarter Pregnancy mobile health coaching program (N=2554).

Characteristics	Statistics
Demographics	
Age[a] (years), median (IQR[b])	31 (28 to 34)
Neighborhood status score, median (IQR)	−0.18 (−1.14 to 0.69)
Pregnant at baseline (yes), n (%)	1300 (50.90)
Body mass index[a] (kg/m^2), median (IQR)	23.9 (21.4 to 27.5)
Participating as couple (yes), n (%)	521 (20.4)
Geographic origin[a] (Western), n (%)	1862 (72.91)
Nutrition and lifestyle behaviors	
Vegetable intake (inadequate), n (%)	
At start of the program	1985 (77.72)
At 24 weeks	1462 (57.24)
Fruit intake (inadequate), n (%)	
At start of the program	1024 (40.09)
At 24 weeks	576 (22.6)
Folic acid supplement use (inadequate), n (%)	
At start of the program	316 (12.4)
At 24 weeks	114 (4.5)
Smoking (yes), n (%)	
At start of the program	252 (9.9)
At 24 weeks	182 (7.1)
Alcohol consumption (yes), n (%)	
At start of the program	605 (23.7)
At 24 weeks	339 (13.3)
Total risk score, median (IQR)	3 (1-6)

[a]Age, body mass index, and geographic origin were missing in 1.2%, 0.4%, and 9.6% of the study population, respectively.
[b]IQR: interquartile range.

Table 2. Univariate associations between the neighborhood status score and demographic factors (N=2554).

Characteristic	B[a] (95% CI)	*P* value
Age[b] (years)	0.04 (0.04 to 0.05)	<.001
Pregnant at baseline (yes)	−0.30 (−0.41 to −0.19)	<.001
Body mass index[b] (kg/m^2)	−0.03 (−0.04 to −0.02)	<.001
Participating as couple (yes)	0.18 (0.11 to 0.25)	<.001
Geographic origin[b] (non-Western)	−0.78 (−0.85 to −0.70)	<.001
Total risk score	−0.01 (−0.02 to 0.001)	.42

[a]β: effect size.
[b]Age, body mass index, and geographic origin were missing in 1.2%, 0.4%, and 9.6% of the study population, respectively.

Table 3. The association between the neighborhood status score and inadequate nutrition and lifestyle behaviors in all participating women at the start of the program (N=2554).

Nutrition and lifestyle behaviors	Crude, OR[a] (95% CI)	*P* value	Adjusted[b], OR (95% CI)	*P* value
Vegetable intake (inadequate)	1.04 (0.98-1.11)	.21	1.04 (0.98-1.12)	.20
Fruit intake (inadequate)	1.03 (0.97-1.09)	.29	1.01 (0.95-1.07)	.74
Folic acid supplement use (inadequate)	1.00 (0.92-1.08)	.94	1.00 (0.90-1.09)	.85
Smoking (yes)	0.85 (0.78-0.92)	<.001	0.85 (0.77-0.93)	<.001
Alcohol consumption (yes)	1.23 (1.15-1.32)	<.001	1.14 (1.04-1.24)	.004

[a]OR: odds ratio.

[b]Adjusted for body mass index, age, geographic origin, pregnancy status, and participation as a couple.

Table 4. The association between the neighborhood status score and improvement of inadequate nutrition and lifestyle behaviors after 24 weeks of coaching in all women who scored inadequately at the start of the mobile health program.

Nutrition and lifestyle behaviors	Crude, OR[a] (95% CI)	*P* value	Adjusted[b], OR (95% CI)	*P* value
Vegetable intake (inadequate) (n=1462)	0.86 (0.79-0.94)	.001	0.89 (0.82-0.97)	.02
Fruit intake (inadequate) (n=576)	0.90 (0.81-1.00)	.051	0.93 (0.84-1.04)	.21
Folic acid supplement use (inadequate) (n=114)	1.00 (0.80-1.24)	.97	1.02 (0.80-1.30)	.87
Smoking (yes) (n=182)	0.87 (0.69-1.10)	.23	0.90 (0.69-1.16)	.40
Alcohol consumption (yes) (n=339)	1.04 (0.9-1.19)	.57	1.05 (0.91-1.21)	.49

[a]OR: odds ratio.

[b]Adjusted for body mass index, age, geographic origin, pregnancy status and participation as a couple.

Nutrition and Lifestyle Behaviors

As coaching was only aimed at nutrition and lifestyle behaviors that were reported as inadequate at the start of the program, improvement of these behaviors was only studied in subsets of women. Overall, women who used the Smarter Pregnancy program improved all nutrition and lifestyle behaviors (Table 1). At the start of the program, vegetable intake was most frequently inadequate (77.72% (1985/2554)). After 24 weeks of coaching, this was reduced to 57.24% (1462/2554), which is a relative improvement of 26%. The largest improvement (relative improvement of 64%) was achieved for folic acid supplement use; this was inadequate in 12.4% (316/2554) women at the start of the program and reduced to 4.5% (114/2554) after 24 weeks.

At the start of the program, no statistically significant association between NSS and inadequate vegetable and or fruit intake was found. However, women with a higher NSS were significantly less likely to smoke (adjusted OR 0.85; 95% CI 0.77-0.93) but more likely to consume alcohol (adjusted OR 1.14, 95% CI 1.04-1.24; Table 3). NSS was not associated with the amount of improvement in smoking and alcohol consumption after 24 weeks of coaching (Table 4). However, NSS was significantly negatively associated with improvement of vegetable intake after 24 weeks of coaching—women with a higher NSS improved their vegetable intake less than women with a lower NSS (adjusted OR 0.89, 95% CI 0.82-0.97). Improvement of the other nutrition and lifestyle behaviors did not significantly depend on NSS (Table 4).

Interaction tests showed that the association between NSS and nutrition and lifestyle behaviors was not significantly different in women who did or did not participate as a couple and who were pregnant or not pregnant at the start of the program. However, at the start of the program, the association between NSS and alcohol consumption was stronger in non-Western (adjusted OR 1.74, 95% CI 1.33-2.28) compared with Western women (adjusted OR 1.12; 95% CI 1.00-1.25). This difference between non-Western and Western women was not observed for the association between the NSS and improvement of alcohol consumption.

Discussion

Principal Findings

Following the results of van Dijk et al, this study demonstrated that women improve their inadequate nutrition and lifestyle behaviors after 24 weeks of mHealth coaching using the Smarter Pregnancy program [22]. However, especially with regard to vegetable intake, this improvement is less in women living in a lesser deprived neighborhood (higher NSS). Although women with a higher NSS were less likely to smoke and more likely to consume alcohol at the start of the program compared with women with a lower NSS, we observed no significant differences in the amount of improvement of these lifestyle behaviors. Furthermore, NSS was significant and negatively associated with compliance to the Smarter Pregnancy program; women with a higher NSS were less likely to complete the 24 weeks of coaching than women with a lower NSS.

Comparison With Previous Work

Currently, a growing number of mHealth apps are developed for personal lifestyle and medical health care support. These apps provide interaction and targeted information on particular

domains for specific target groups, and improvement in self-reported health behaviors because these apps are observed. Specifically, decreased tobacco use, increased vitamin intake, and more frequent healthy food intake have been reported after coaching by apps designed to encourage healthy behavior. Therefore, in our opinion, it is important to conduct profound research both in low- and middle- as well as in high income countries before these apps can be implemented in medical health care [34].

In this study's population, women who lived in less deprived neighborhoods were less likely to smoke but more likely to consume alcohol, which is in line with previous studies [9,35,36]. Despite recent studies stating that residents who live in deprived neighborhoods are difficult to motivate to change unhealthy behaviors [3,17,18], in this study, those women were more likely to complete the 24 weeks of mHealth coaching and improve their nutrition and lifestyle behaviors more than women who live in less deprived neighborhoods. This is rather surprising as we expected the opposite, namely that higher educated women, more often living in a neighborhood with a higher NSS, have generally higher health literacy skills compared with women from a neighborhood with a lower NSS and therefore improve behaviors more quickly [37]. An explanation may be that higher educated women believe that they already have healthy behaviors and do not need to change [37,38]. In addition, our previously conducted focus group study among women participating in the Smarter Pregnancy program reported that higher educated women showed a lower compliance and appreciated the program less than middle- and low-educated women, who often live in neighborhoods with lower NSS [39]. This is in line with the fact that the content of the coaching is compiled so that it matches the skills and knowledge of the largest population of middle- and low-educated women who generally have a higher prevalence of unhealthy lifestyle behaviors.

Strengths and Limitations

Strengths of this study are the large number of included participants (N=2554), the high overall compliance of 68.17% (1741/2554) of women who completed the 24 weeks of coaching, the fact that several potential confounders were taken into account in the adjusted models, and the imputation of missing data. In this study, NSS—based on a well-defined index—was used as a proxy for socioeconomic health inequality among neighborhoods. This continuous measure of neighborhood deprivation was used instead of a dichotomous measure (ie, deprived vs nondeprived), which provides a more precise evaluation of the effect of neighborhood deprivation. The use of area-based indices as a proxy for socioeconomic health is well supported in the literature; thus, the used neighborhood deprivation index can be considered a valid indicator [40,41]. NSS is a measure based on factors that are specific for (the residents in) that particular neighborhood. Indeed, we found that NSS is a representative measure for deprivation characteristics on the individual level; in less deprived neighborhoods, participating women had a lower BMI and were more likely to be of non-Western geographic origin. Furthermore, the distribution of NSS in this study cohort (IQR −1.14 to 0.69, data not shown) was comparable with the national NSS in the year 2014 (IQR −0.57 to 0.71).

Despite the fact that the inclusion period of the study population and the coaching with the Smarter Pregnancy program covers several years, the NSS of 2014 was used as the measure of neighborhood deprivation for the whole study population. As the NSS and the ratio of score among the neighborhoods do not change much over time, we consider this a valid determinant of the neighborhood deprivation within the study population.

Although geographic origin is known to be a potential confounding variable for associations with deprivation, information regarding geographic origin was not directly available from our database. To take geographic origin into account, we retrospectively performed geographic classification. This approach is considered a valid method for ascribing individuals to geographic groups when self-identification is not available [29], but unfortunately, it does not permit any further subdivision into more specific geographic groups besides Western and non-Western.

Limitations of this study are the absence of validation of nutritional status by biomarkers and the absence of a control group, although this is inherent to this study's design. Furthermore, the Smarter Pregnancy program was only available in Dutch and on multiple devices with internet access and preferably a mobile phone. Consequently, only those familiar with the Dutch language and in possession of a mobile phone with internet access participated. Over 95% of all women and men of reproductive age living in the Netherlands have internet access on their mobile phone, making the program properly accessible [42]. However, a selection may have occurred of only those familiar with the Dutch language, who are mainly of Western origin. This is reflected by the fact that over 80% of the women in this study were of Western geographic origin, although, on the basis of the population distribution of the city of Rotterdam, a percentage of 62% was to be expected [43]. Misclassification of the geographic origin because of incorrect assignment cannot be excluded. However, the surname-based method for ascribing individuals to geographic groups when self-identification is not available is previously described as a valid method. Another form of selection bias may have been induced as the Smarter Pregnancy program was not routinely used or recommended as part of (pre) pregnancy care, and participants mostly subscribed upon their own initiative. Therefore, women could have been mainly women who are already intrinsically motivated to change nutrition and lifestyle behaviors before starting the mHealth program. Together, these limitations may contribute to the generalizability of this study's results.

Conclusions and Future Perspectives

Overall, we can conclude that the Smarter Pregnancy mHealth coaching program is able to motivate and support women from more and less deprived neighborhoods to improve their nutrition and lifestyle behaviors. However, women who live in more deprived neighborhoods seem to improve their nutrition and lifestyle behaviors more compared with women from less deprived neighborhoods.

Together, these findings underline the need for a more tailored version of the program, adapted to the needs of its participants on the basis of demographic characteristics, so that the program can adequately and optimally empower all women to improve their nutrition and lifestyle behaviors.

Acknowledgments

This research was funded by the Department of Obstetrics and Gynecology, Erasmus University Medical Centre, Rotterdam, the Netherlands, a grant of the health care insurance company ZilverenKruis and Achmea, the Netherlands (StichtingAchmeaGezondheidszorg, 2011), and by the Erasmus MC Medical Research Advisory Committee (Mrace) program Health Care Efficiency Research.

Authors' Contributions

DVG, MRvD, and MPHK analyzed and interpreted all data and wrote the first version of the manuscript. SPW contributed to the data analysis. RPMST designed the study and contributed to all aspects of the study. MRvD acquired the data. DVG, MRvD, MPHK, SPW, EAPS, and RPMST revised all versions of the manuscript. All authors approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Smarter Pregnancy program.

[PDF File (Adobe PDF File), 85KB - mhealth_v7i4e11664_app1.pdf]

Multimedia Appendix 2

Description of the coaching program.

[PDF File (Adobe PDF File), 76KB - mhealth_v7i4e11664_app2.pdf]

References

1. Mackenbach JP. The persistence of health inequalities in modern welfare states: the explanation of a paradox. Soc Sci Med 2012 Aug;75(4):761-769. [doi: 10.1016/j.socscimed.2012.02.031] [Medline: 22475407]
2. World Health Organization. 2018. Key concepts: Social determinants of health URL: http://www.who.int/social_determinants/thecommission/finalreport/key_concepts/en/ [accessed 2018-07-10] [WebCite Cache ID 70o62UC7h]
3. de Graaf JP, Steegers EA, Bonsel GJ. Inequalities in perinatal and maternal health. Curr Opin Obstet Gynecol 2013 Apr;25(2):98-108. [doi: 10.1097/GCO.0b013e32835ec9b0] [Medline: 23425665]
4. Leddy MA, Power ML, Schulkin J. The impact of maternal obesity on maternal and fetal health. Rev Obstet Gynecol 2008;1(4):170-178. [Medline: 19173021]
5. Weck RL, Paulose T, Flaws JA. Impact of environmental factors and poverty on pregnancy outcomes. Clin Obstet Gynecol 2008 Jun;51(2):349-359. [doi: 10.1097/GRF.0b013e31816f276e] [Medline: 18463465]
6. Metcalfe A, Lail P, Ghali WA, Sauve RS. The association between neighbourhoods and adverse birth outcomes: a systematic review and meta-analysis of multi-level studies. Paediatr Perinat Epidemiol 2011 May;25(3):236-245. [doi: 10.1111/j.1365-3016.2011.01192.x] [Medline: 21470263]
7. de Graaf JP, Ravelli AC, Wildschut HI, Denktaş S, Voorham AJ, Bonsel GJ, et al. [Perinatal outcomes in the four largest cities and in deprived neighbourhoods in The Netherlands]. Ned Tijdschr Geneeskd 2008 Dec 13;152(50):2734-2740. [Medline: 19192587]
8. de Graaf JP, Ravelli AC, de Haan MA, Steegers EA, Bonsel GJ. Living in deprived urban districts increases perinatal health inequalities. J Matern Fetal Neonatal Med 2013 Mar;26(5):473-481. [doi: 10.3109/14767058.2012.735722] [Medline: 23039164]
9. Timmermans S, Bonsel GJ, Steegers-Theunissen RP, Mackenbach JP, Steyerberg EW, Raat H, et al. Individual accumulation of heterogeneous risks explains perinatal inequalities within deprived neighbourhoods. Eur J Epidemiol 2011 Feb;26(2):165-180 [FREE Full text] [doi: 10.1007/s10654-010-9542-5] [Medline: 21203801]
10. Algren MH, Bak CK, Berg-Beckhoff G, Andersen PT. Health-risk behaviour in deprived neighbourhoods compared with non-deprived neighbourhoods: a systematic literature review of quantitative observational studies. PLoS One 2015;10(10):e0139297 [FREE Full text] [doi: 10.1371/journal.pone.0139297] [Medline: 26506251]
11. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. J Epidemiol Community Health 2001 Feb;55(2):111-122 [FREE Full text] [doi: 10.1136/jech.55.2.111] [Medline: 11154250]

12. Griffin J, Emery BR, Huang I, Peterson CM, Carrell DT. Comparative analysis of follicle morphology and oocyte diameter in four mammalian species (mouse, hamster, pig, and human). J Exp Clin Assist Reprod 2006 Mar 1;3:2 [FREE Full text] [doi: 10.1186/1743-1050-3-2] [Medline: 16509981]
13. Steegers-Theunissen RP, Twigt J, Pestinger V, Sinclair KD. The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. Hum Reprod Update 2013;19(6):640-655 [FREE Full text] [doi: 10.1093/humupd/dmt041] [Medline: 23959022]
14. van Uitert EM, van der Elst-Otte N, Wilbers JJ, Exalto N, Willemsen SP, Eilers PH, et al. Periconception maternal characteristics and embryonic growth trajectories: the Rotterdam Predict study. Hum Reprod 2013 Dec;28(12):3188-3196. [doi: 10.1093/humrep/det375] [Medline: 24105824]
15. Barker DJ. The origins of the developmental origins theory. J Intern Med 2007 May;261(5):412-417 [FREE Full text] [doi: 10.1111/j.1365-2796.2007.01809.x] [Medline: 17444880]
16. 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]
17. Weightman AL, Morgan HE, Shepherd MA, Kitcher H, Roberts C, Dunstan FD. Social inequality and infant health in the UK: systematic review and meta-analyses. BMJ Open 2012;2(3):e000964 [FREE Full text] [doi: 10.1136/bmjopen-2012-000964] [Medline: 22700833]
18. Hammiche F, Laven JS, van Mil N, de Cock M, de Vries JH, Lindemans J, et al. Tailored preconceptional dietary and lifestyle counselling in a tertiary outpatient clinic in The Netherlands. Hum Reprod 2011 Sep;26(9):2432-2441 [FREE Full text] [doi: 10.1093/humrep/der225] [Medline: 21752799]
19. Silva BM, Rodrigues JJ, de la Torre Diez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272. [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
20. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of M-health technologies for improving health and health services: a systematic review protocol. BMC Res Notes 2010;3:250 [FREE Full text] [doi: 10.1186/1756-0500-3-250] [Medline: 20925916]
21. [Slimmer Pregnant]. URL: https://www.slimmerzwanger.nl/ [accessed 2018-06-28] [WebCite Cache ID 70VYJtRZ6]
22. Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth platform for pregnancy on nutrition and lifestyle of the reproductive population: a survey. JMIR Mhealth Uhealth 2016;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
23. van Dijk MR, Oostingh EC, Koster MP, Willemsen SP, Laven JS, Steegers-Theunissen RP. The use of the mHealth program Smarter Pregnancy in preconception care: rationale, study design and data collection of a randomized controlled trial. BMC Pregnancy Childbirth 2017 Dec 26;17(1):46 [FREE Full text] [doi: 10.1186/s12884-017-1228-5] [Medline: 28125970]
24. van Dijk MR, Koster MP, Willemsen SP, Huijgen NA, Laven JS, Steegers-Theunissen RP. Healthy preconception nutrition and lifestyle using personalized mobile health coaching is associated with enhanced pregnancy chance. Reprod Biomed Online 2017 Oct;35(4):453-460. [doi: 10.1016/j.rbmo.2017.06.014] [Medline: 28688924]
25. Cnattingius S. The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. Nicotine Tob Res 2004 Apr;6(Suppl 2):S125-S140. [doi: 10.1080/14622200410001669187] [Medline: 15203816]
26. Englund-Ögge L, Brantsæter AL, Sengpiel V, Haugen M, Birgisdottir BE, Myhre R, et al. Maternal dietary patterns and preterm delivery: results from large prospective cohort study. Br Med J 2014 Mar 4;348:g1446 [FREE Full text] [doi: 10.1136/bmj.g1446] [Medline: 24609054]
27. Steenbeek R, Schellart AJ, Mulders H, Anema JR, Kroneman H, Besseling J. The development of instruments to measure the work disability assessment behaviour of insurance physicians. BMC Public Health 2011 Jan 3;11:1 [FREE Full text] [doi: 10.1186/1471-2458-11-1] [Medline: 21199570]
28. Fogg BJ. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Apr Presented at: Persuasive'09; April 26-29, 2009; Claremont, California, USA. [doi: 10.1145/1541948.1541999]
29. Mateos P. A review of name‐based ethnicity classification methods and their potential in population studies. Popul Space Place 2007 May 22;13(4):243-263. [doi: 10.1002/psp.457]
30. [The Netherlands Nutrition Centre]. 2018. [Pregnancy] URL: https://www.voedingscentrum.nl/professionals/kindervoeding-0-4-jaar/zwangerschap.aspx [accessed 2018-07-10] [WebCite Cache ID 70o6ebrcz]
31. Kleinman JC, Pierre MB, Madans JH, Land GH, Schramm WF. The effects of maternal smoking on fetal and infant mortality. Am J Epidemiol 1988 Feb;127(2):274-282. [doi: 10.1093/oxfordjournals.aje.a114803] [Medline: 3337083]
32. [Social and Cultural Planning Office]. 2012 Nov 26. [Status development of districts in the Netherlands 1998-2010] URL: https://www.scp.nl/nl/Publicaties/Alle_publicaties/Publicaties_2012/Statusontwikkeling_van_wijken_in_Nederland_1998_2010 [accessed 2019-03-02] [WebCite Cache ID 76ZgbZnlf]
33. [Social and Cultural Planning Office]. 2018. [Status scores] URL: https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores [accessed 2018-07-10] [WebCite Cache ID 70o6MmoGD]
34. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology–based lifestyle and medical intervention apps supporting health care during pregnancy:

systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
35. Taylor AW, Bewick BM, Makanjuola AB, Qian L, Kirzhanova VV, Alterwain P. Context and culture associated with alcohol use amongst youth in major urban cities: a cross-country population based survey. PLoS One 2017;12(11):e0187812 [FREE Full text] [doi: 10.1371/journal.pone.0187812] [Medline: 29155847]
36. Hansen AB, Hvidtfeldt UA, Grønbæk M, Becker U, Nielsen AS, Tolstrup JS. The number of persons with alcohol problems in the Danish population. Scand J Public Health 2011 Mar;39(2):128-136. [doi: 10.1177/1403494810393556] [Medline: 21382854]
37. Pampel FC, Krueger PM, Denney JT. Socioeconomic disparities in health behaviors. Annu Rev Sociol 2010 Aug;36:349-370 [FREE Full text] [doi: 10.1146/annurev.soc.012809.102529] [Medline: 21909182]
38. Smith AM, Baghurst K, Owen N. Socioeconomic status and personal characteristics as predictors of dietary change. J Nutr Educ Behav 1995 Jul;27(4):173-181. [doi: 10.1016/S0022-3182(12)80424-9]
39. van Dijk MR, Koster MP, Rosman AN, Steegers-Theunissen RP. Opportunities of mHealth in preconception care: preferences and experiences of patients and health care providers and other involved professionals. JMIR Mhealth Uhealth 2017 Aug 17;5(8):e123 [FREE Full text] [doi: 10.2196/mhealth.7834] [Medline: 28818820]
40. Rajaratnam JK, Burke JG, O'Campo P. Maternal and child health and neighborhood context: the selection and construction of area-level variables. Health Place 2006 Dec;12(4):547-556. [doi: 10.1016/j.healthplace.2005.08.008] [Medline: 16188483]
41. Janevic T, Stein CR, Savitz DA, Kaufman JS, Mason SM, Herring AH. Neighborhood deprivation and adverse birth outcomes among diverse ethnic groups. Ann Epidemiol 2010 Jun;20(6):445-451 [FREE Full text] [doi: 10.1016/j.annepidem.2010.02.010] [Medline: 20470971]
42. Stephenson J, Patel D, Barrett G, Howden B, Copas A, Ojukwu O, et al. How do women prepare for pregnancy? Preconception experiences of women attending antenatal services and views of health professionals. PLoS One 2014;9(7):e103085 [FREE Full text] [doi: 10.1371/journal.pone.0103085] [Medline: 25058333]
43. [City of Rotterdam]. [Rotterdam in Figures] URL: https://rotterdam.buurtmonitor.nl/jive?cat_open=Beleidsthema%27s/Demografie [accessed 2019-03-02] [WebCite Cache ID 76ZgvhyVK]

Abbreviations

BMI: body mass index
IQR: interquartile range
MC: Medical Centre
mHealth: mobile health
NSS: neighborhood status score
OR: odds ratio
SMS: short message service
TRS: total risk score

Edited by C Dias; submitted 24.07.18; peer-reviewed by O Ilozumba, R de Leeuw, I de la Torre; comments to author 03.10.18; revised version received 27.11.18; accepted 19.12.18; published 11.04.19.

Please cite as:
Gootjes DV, van Dijk MR, Koster MPH, Willemsen SP, Steegers EAP, Steegers-Theunissen RPM
Neighborhood Deprivation and the Effectiveness of Mobile Health Coaching to Improve Periconceptional Nutrition and Lifestyle in Women: Survey in a Large Urban Municipality in the Netherlands
JMIR Mhealth Uhealth 2019;7(4):e11664
URL: https://mhealth.jmir.org/2019/4/e11664/
doi:10.2196/11664
PMID:30973345

Original Paper

Effectiveness and Feasibility of a Remote Lifestyle Intervention by Dietitians for Overweight and Obese Adults: Pilot Study

Karin Haas[1], Dr rer nat; Stefanie Hayoz[1], Dr phil nat; Susanne Maurer-Wiesner[2], Dr med

[1]Applied Research & Development Nutrition and Dietetics, Department of Health Professions, Bern University of Applied Sciences, Bern, Switzerland
[2]Center of Obesity and Metabolism Winterthur GmbH, Winterthur, Switzerland

Corresponding Author:
Karin Haas, Dr rer nat
Applied Research & Development Nutrition and Dietetics
Department of Health Professions
Bern University of Applied Sciences
Murtenstrasse 10
Bern, 3014
Switzerland
Phone: 41 31 848 3553
Email: karin.haas@bfh.ch

Abstract

Background: To tackle the problem of obesity and related diseases in Switzerland, cost-efficient, effective, and innovative primary health care interventions for weight management are required. In this context, Oviva has developed a scalable technology for registered dietitians to counsel overweight and obese patients via a mobile phone app.

Objective: The aim of this study was to evaluate the effectiveness and feasibility of weight loss counseling by dietitians using a mobile phone app for patients with overweight and obesity.

Methods: In this pre- and posttest pilot study, overweight and obese adults participated in a 1-year behavioral intervention to lose weight through remote counseling by dietitians in the German-speaking part of Switzerland. The study started in April 2016 and finished in May 2018. Participants received individual counseling through the app and the exchange with the dietitian focused on regular feedback on photo-based food log, motivation, and education. The contents were tailored to the individual lifestyle goal set. The predefined intensity of remote counseling decreased during the year. Group chat could be used. The outcomes examined were changes in weight (primary outcome), hemoglobin A_{1c}, fasting glucose, fasting insulin, triglyceride, high-density lipoprotein cholesterol, blood pressure (BP), body mass index (BMI), waist circumference, body fat, and responses to a self-administered questionnaire with questions regarding participants' physical activity, dietary assessment, and health-related quality of life. Changes were tested at baseline, after 3 months, and after 12 months, as well as between the third and the 12th month.

Results: In total, 36 women and 7 men, with a mean age of 40.6 years, participated and 36 participants completed the study. Median weight change after the first 12 weeks was −3.8 kg (range: −15 to 2.4 and P<.001), between week 12 and week 52 it was −1.1 kg (range: −9.7 to 7 and P=.08), and the median change during the entire period of intervention was −4.9 kg (range: −21.9 to 7.5 and P<.001). Furthermore, changes in BMI, waist circumference, body fat, and BP between baseline and 12 weeks and between baseline and 52 weeks were also significant. Significant changes in certain eating habits were also demonstrated (higher frequency of vegetable, fruit, and breakfast consumption and lower frequency of alcohol, sweet, and fat consumption).

Conclusions: In addition to the professional skills of a dietitian, a profession-specific app such as Oviva can provide effective support that meets the needs of dietitians and clients on the long path of behavioral change and sustainable weight reduction.

Trial Registration: ClinicalTrials.gov NCT02694614; https://clinicaltrials.gov/ct2/show/NCT02694614 (Archived by WebCite at http://www.webcitation.org/76gYkGOIc)

(JMIR Mhealth Uhealth 2019;7(4):e12289) doi:10.2196/12289

KEYWORDS

remote consultation; obesity; weight loss; mobile app; behavior therapy; healthy lifestyle; healthy diet

Introduction

Background

Obesity has become one of the biggest public health challenges in recent decades [1]. In Europe, more than half of the population is classified as overweight or obese [2]. The prevalence is also high in Switzerland. Currently, 41.6% and 13.9% of men and 19.7% and 11.3% of women are overweight or obese, respectively [3]. Both overweight and obesity are associated with the incidence of noncommunicable diseases including type 2 diabetes, cardiovascular diseases, certain types of cancers, as well as premature death [4,5]. Overweight and obesity are also associated with psychosocial problems such as disordered eating, depression, anxiety, body image dissatisfaction, and low self-esteem in relation to weight stigma [6]. In addition to the negative effects on individual health, obesity is responsible for a large proportion of cost for the health care system and society [7,8]. In 2012, the direct costs of overweight and obesity associated diseases represented 7.2% of the total Swiss health care expenses [9].

International guidelines recommend a weight loss of at least 5% for the treatment of individuals with overweight or obesity to achieve positive clinical effects. The magnitude of weight loss is associated with dose-effect improvements in cardiovascular risk factors such as high blood pressure (BP), hyperlipidemia, and hyperglycemia [10]. In addition, waist circumference and body composition (maintenance or increase in muscle mass and decrease in fat mass) must be considered as important outcomes [11]. To achieve sustainable positive changes in these parameters, the first step must be a lifestyle modification concerning eating behavior and physical activities [10,11]. Behavior-based interventions have been shown to be effective for weight loss outcomes including change in weight and waist circumference [12,13]. In addition, the effectiveness of a weight loss program may be influenced by the intensity of an intervention and the continuous provision of a varied program [10,14]. However, behavior change requires time and must be maintained. At this maintenance period, numerous barriers, including the absence of social support, a lack of time management, health status changes, life transitions, and lack or decline of motivation can impact on the original success of weight loss [15,16].

Therefore long-term, cost-efficient, effective, and innovative primary health care interventions for weight management are needed.

Web-Based Nutrition Counseling

New digital technologies could be a possible solution as they provide innovative counseling possibilities with greater flexibility to coach overweight and obese clients. Such technologies can help reach clients who are unable and/or unwilling to engage in face-to-face counseling and make it possible to assist clients more conveniently and as required (eg, in the weight maintenance phase). Studies show that both face-to-face and remote counseling are effective methods for losing weight, as well as for maintaining weight loss [17-20]. Mobile phone app and short message service (SMS) text messaging could offer effective intervention strategies in this context. For example, SMS text messaging is a cheap, portable, convenient, and innovative medium facilitating goal setting, self-monitoring, and information exchange [21] and allows individualization, which are important components of behavior change. For obese or overweight adults, personalization through counseling, individualized feedback, as well as social support and self-monitoring system seem to be important when using mobile phones to assist behavior change and therefore promote weight loss [22,23]. Small successes can be visualized with these tools, which may have an impact on patient self-efficacy. This may in turn support further changes in behavior and lead to long-term successful weight management [24].

Although these innovative tools offer dietitians new potential for counseling and supporting eating and physical behavior change, their use of these tools remains limited. One reason may be that few studies on the long-term impact of Web-based nutrition counseling exist. The number of commercial health and nutrition apps available is huge; however, specific apps for remote counseling, which are aligned to the needs of a dietitian, are limited. In Switzerland, Oviva, a new type of health care provider, has come to the market using new technology to scale the dietitian workforce. Oviva has developed a scalable technology for dietitians to counsel overweight and obese patients via a mobile phone app. Through this app, dietitians can organize their sessions through messaging and video calls, have access to photo food diary, activity, and weight tracking features allowing them to monitor client progress through their photos, activity, and weight logs, as well as set goals and share documents and surveys. Assessing the implementation of this novel kind of intervention is important [25,26] as similar innovative solutions for use by dietitians have not yet been investigated in Switzerland. The aim of this pilot study was to evaluate the effectiveness and feasibility of weight loss counseling using the mobile phone app Oviva for overweight or obese adults.

Methods

Study Design

This study was carried out as a 1-year, single arm pre- and postpilot intervention with overweight or obese adults in Switzerland from March 2016 (first participant in) to May 2018 (last participant out). All study procedures were approved by the local institutional review board and the study is registered under ClinicalTrials.gov (NCT02694614). Follow-up will be conducted 1 year after the intervention.

Study Population and Recruitment

Sample size was calculated with a significance level of 0.05 and power of 80%. A sample size of 36 was needed to detect a weight loss of 0.5 SD, which corresponds to a medium effect size according to Cohen, using a 2-sided Wilcoxon signed rank test. To account for dropouts, 50 participants were planned. After screening, 43 participants were included in the study.

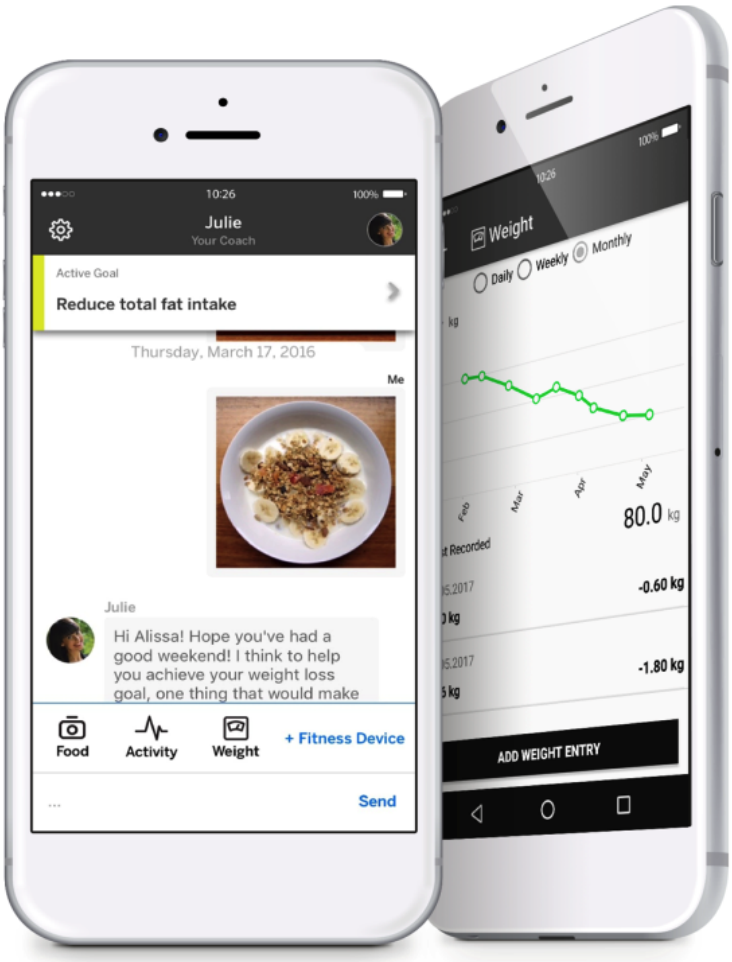
Participants were included in the study if they met the following inclusion criteria: adults (aged 18 years and over), body mass index (BMI) between 26 and 33 kg/m^2, fluent in German, mobile phone user (iOS or Android), and capable of sending and

receiving SMS text messages and pictures. Participants were excluded if they were pregnant or breastfeeding, were diagnosed with conditions other than dyslipidemia, hypertension, and insulin resistance requiring nutrition therapy, had serious disease requiring continuous drug therapy, were on a weight reduction diet during the last 6 months, took medication for weight loss in the past, or enrolled in another weight loss program. Subjects were invited to participate with flyers distributed through the Center for Obesity and Metabolism Medicine Winterthur (in Canton Zurich), via general practitioners, advertisements on the websites of the participating research institutions, local newspapers, and through word of mouth advertising. Interested persons received written and verbal information about the study. Screening took place following informed consent to determine the eligibility of participants on the basis of the inclusion and exclusion criteria at the Center for Obesity and Metabolism Medicine Winterthur.

Oviva App

The program is designed to facilitate personal coaching by registered dietitians. This digital communication system connects patients remotely with their dietitian who has all the tools required for remote counseling. Via customer-facing apps for iOS and Android, all clients are able to collect relevant information (eg, food intake, physical activity, and weight) and communicate easily with their dietitian. All features of the app are modeled on typical activities in a dietitian's everyday practice and include chat-like communication with dietitians, group chats for support from peers, dietitian's profile to create a more personal connection, a photo-based food log, activity and weight logs, a goal scorecard, showing past goals and future options, a content database, feedback to the dietitian, and links to standard learning materials. Client exchanges with the dietitian focused on regular feedback regarding the photo-based food log, motivation, and education (eg, recipes, nutritional facts, and challenges). Clients were also able to join group sessions within the same app. All data were transferred and saved via secure channels. For the dietitians, all information was displayed in a secure cloud-based platform, similar to Webmail. They could review all relevant information for each client individually, take notes, create customized surveys, and use a content database, including relevant links. Figure 1 shows the client view of a chat communication including a photo-based food log and weight log.

Figure 1. Client view of a chat communication including a photo-based food log and weight log.



Intervention

This behavior-based intervention was carried out entirely remotely by 3 registered dietitians using Oviva technology. The intervention comprised 3 different coaching phases as described in Textbox 1.

The content of the remote counseling was tailored to the specific needs of each participant, but it was tailored with the overall aim to lose weight. The behavioral strategies used were self-monitoring, goal setting, stimulus control, elaboration of alternative strategies, social support, positive reinforcement, and relapse prevention (particularly toward the end of the intervention). Motivational interview techniques were used throughout the entire intervention process.

Key counseling topics were based on the needs of the participants and included negative energy balance, plate model (Swiss tool for the correct proportions of each food group at every meal), portion size, meal rhythm, energy density of food, quality of carbohydrate, fat, carbohydrate-modified diet, adequate protein sources, dealing with special situations (such as emotional eating, invitations to dinner and parties, holidays, eating triggers, lapses), grocery shopping, and physical activity.

Outcome Measures: Anthropometric and Metabolic Risk Factors

The primary outcome was defined as weight change between baseline (M0) and 1 year (M12). All measurements were examined at baseline (M0), at 3 months (M3), and 1 year (M12), and it included anthropometric, clinical, and questionnaire assessments. Anthropometric and clinical assessments included body height, body weight, BMI, waist circumference, body fat, hemoglobin A_{1c} (HbA_{1c}), fasting glucose, fasting insulin, triglyceride, high-density lipoprotein (HDL) cholesterol, and BP. Measurements were carried out by a clinical nurse in the morning. Participants were instructed to fast overnight (8 hours), abstain from smoking, consuming alcohol, and drinking caffeine-containing beverages, as well as from vigorous exercise 48 hours before blood samples were taken and body composition was measured. Weight was measured using a clinically validated and calibrated scale (Seca mBCA 515, medical Body Composition Analyzer). Participants were lightly dressed, without shoes, and weight was recorded to within 0.01 kg. Height was measured using a calibrated stadiometer (Forma Seca). The measurement was carried out in an upright position with feet positioned on the floor board of the stadiometer. Heels, buttocks, and the back of the head touched the back board of the stadiometer with arms on the sides. BMI was calculated by dividing body weight in kilograms by height in meters squared. For the measurement of the waist circumference, an inelastic tape was placed directly on the skin between the lower rib and iliac crest. Each measurement was performed twice after exhalation and recorded within 0.5 cm. If the difference between the 2 measurements was greater than 1 cm, a third measurement was performed and the mean of the 2 closest measurements was calculated. Body composition was measured by bioelectrical impedance analysis (Seca mBCA 515, medical Body Composition Analyzer) with participants in upright position, lightly dressed, and barefoot. BP measurements were performed using the Riva-Rocci method with a special instrument (Firma boso-medicus SN 768 00 440 729 Bosch+Sohn GmbH und Co) and a stethoscope. The measurement was performed 3 times using the right arm after participants had been sitting for at least 10 min. The average of these measurements was used for analysis. Blood samples were obtained after overnight fasting from the antecubital vein to determine HbA_{1c}, fasting glucose, fasting insulin, triglyceride, and HDL cholesterol. All laboratory assays were performed by Labor Risch Schaffhausen using standardized methods as described in the study protocol.

Textbox 1. Phases of intervention.

Prephase or initiation (week −2 to 0)

- Complete initial assessment; implementation: photo-assisted dietary and activity recording; setting overall aim.

Phase 1: Transition (month 1 to 3)

- Mobile phone assisted patient coaching 5 times per week; feedback on nutrition and activity; setting 1 or 2 specific goals for 2 weeks at the time; assessing goals and adapting them if necessary; providing information and education materials that are appropriate to the goals; 1 Skype call at the end; clinical assessment at the end.

Phase 2: Stabilization (month 4 to 6)

- Mobile phone assisted patient coaching 3 times per week; feedback on nutrition and activity; strengthening new behavior; setting new goals if necessary; assessing goals and adapting them if necessary; providing information and education materials that are appropriate to the goals; exchange with peer groups (optional and anonymous, coach guided).

Phase 3: Maintenance (month 7 to 12)

- Mobile phone assisted patient coaching once every 2 weeks; feedback on nutrition and activity if required; exchange with peer groups (optional and anonymous, noncoach guided or coach guided); access to Web-based education materials; final chat; clinical assessment at the end.

Additional Outcome Measures

A questionnaire was used to assess socioeconomic data, dietary assessment, physical activity, and health related quality of life. Physical activity was assessed using the self-administered Global Physical Activity Questionnaire (GPAQ), developed and validated by the World Health Organization. Dietary intake was investigated using a brief 11-item simplified food frequency questionnaire. The English version of this questionnaire was developed and tested in a diverse population as a relatively

simple, valid, and efficient tool for dietary assessment [27]. The English version was translated into German, tested for comprehensibility in a pretest and adapted. The questionnaire includes questions on the following categories: fast food and convenience foods, fruits, vegetables, sweetened beverages, unsweetened beverages, alcoholic beverages, salty snacks, sweets and desserts, fat for food preparation, fatty spreads, and breakfast consumption. Response options for the items are organized into 3 response categories: first category represents the most healthful dietary consumption (0 points), second category represents less healthful consumption (1 point), and the last category represents the least healthful consumption (2 points). The total score (out of 22 points) provides an indication of how healthy the diet is, with a low score indicating a healthy diet. Quality of life was assessed using the 12-item Short-Form Health Survey (SF-12; short version of SF-36, instrument for assessing health-related quality of life in obesity research), which includes 12 items covering physical functioning, role limitations because of physical health problems, bodily pain, general health, energy or fatigue, social functioning, and role limitations because of emotional problems and mental health [28]. The SF-12 generates a physical and a mental health component summary score (PCS and MCS) with a total maximum of each score of 100 (higher scores indicate higher quality of life). The SF-12 scores compare favorably with those obtained using the SF-36 [29].

Statistical Analysis

All outcomes were analyzed on the basis of the intention-to-treat (ITT) population, defined as all participants who started the intervention. Supportive analyses based on the per-protocol (PP) population were performed for selected outcomes. For all outcomes, the changes were calculated from baseline to the specified time points. The changes over the whole observation period were investigated using nonparametric analysis of variance methods for longitudinal data [30]. Posthoc tests for changes at all measured time points were performed using Wilcoxon signed rank tests. Two-tailed tests with significance level .05 were performed for all analyses. No adjustment for multiple testing was performed. All outcome variables were summarized using descriptive statistics. All analyses were performed using R version 3.4.3 (R Foundation for Statistical Computing).

Results

Study Population

Following screening, a total of 43 people aged between 20 and 67 years were included in the study. 36 participants completed the 1-year intervention. Overall, the proportion of women was higher (84%, 36/43) than men, with 6 women and 1 man not completing the intervention. Socioeconomic data of the study population are summarized in Table 1.

In total, 33 people were initially diagnosed with dyslipidemia, 17 with insulin resistance, and 4 with hypertension. No participant was diagnosed with diabetes and 1 participant was diagnosed with prediabetes because of a slightly elevated HbA_{1c} level. At the beginning of the study, 3 people took medication to treat hypertension and 2 people took lipid-lowering medication. The medication did not change during intervention.

Outcome Measures: Anthropometric

The results are reported as median values unless otherwise indicated. The median weight at different timepoints (M0: 83.5 kg, M3: 80.3 kg, and M12: 78.7 kg) are shown in Figure 2. The weight loss between baseline (M0) and month 12 (M12) was −4.9 kg or 6% (P<.001). Weight loss was also significant between M0 and M3 (−3.8 kg, P<.001) but not between M3 and M12 (−1.1kg, P=.08; Table 2). Overall, 58% (21/36) of participants achieved a weight loss of 5% or more from their initial weight, and 5 participants gained weight (between 2.1 kg and 7.5 kg). Of the 5 participants who gained weight during the total intervention, 4 also lost no weight in the first 3 months.

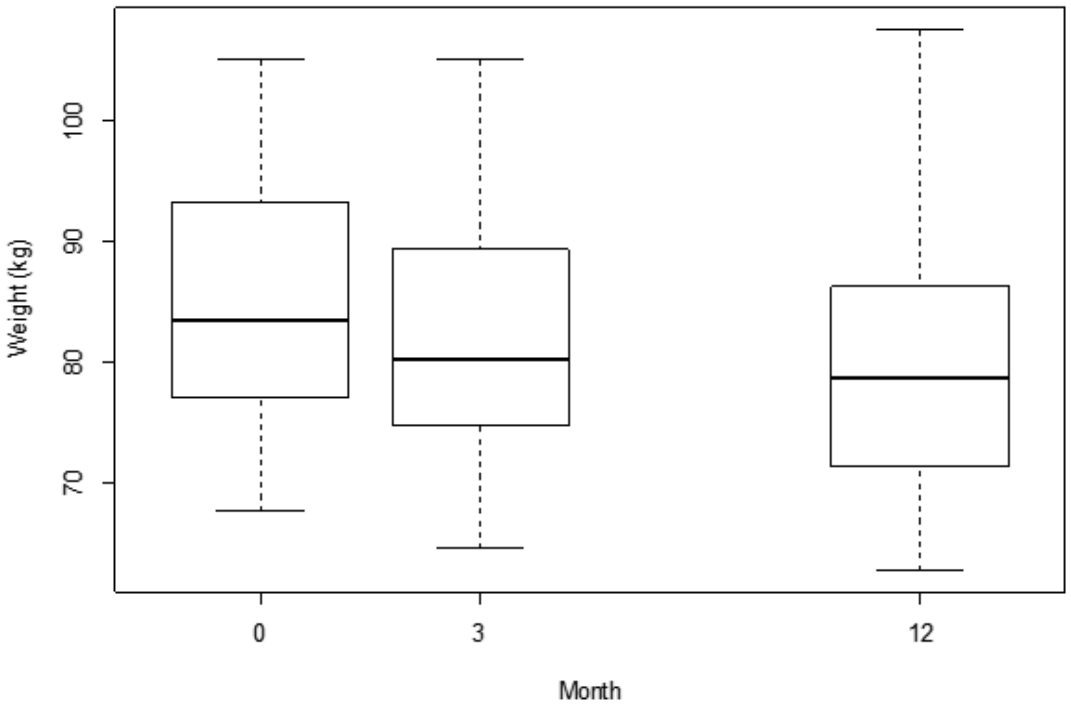
The BMI decreased significantly between M0 and M12 (by 1.8 kg/m^2 [P<.001]) and between M0 and M3 (by 1.4 kg/m^2 [P<.001]). Between M3 and M12, the decrease was 0.4 kg/m^2 (P=.06). Waist circumference, body fat (kg), or body fat (%) showed similar trends; a significant reduction between M0 and M12 (P<.001), as well as between M0 and M3 (P<.001), but no significant effects between M3 and M12 (waist circumference P=.72, body fat [kg] P=.14, body fat [%] P=.20; Table 2). Percentage changes in weight, BMI, and fat (%) are shown in Table 3.

Table 1. Socioeconomic data (N=43).

Characteristics	Statistics
Women, n (%)	36 (84)
Age (years), mean (SD)	40.6 (12.4)
Marital status, n (%)	
Married	18 (42)
Single	15 (35)
Divorced	6 (14)
Separated	3 (7)
Widowed	1 (2)
Children living in the household, n (%)	
0	19 (44)
1	9 (21)
2	11 (26)
3	4 (9)
Net household income per month, n (%)	
Less than 3000 CHF[a]	1 (2)
3000 to 4500 CHF	6 (14)
4500 to 6000 CHF	6 (14)
6000 to 9000 CHF	13 (30)
More than 9000 CHF	17 (40)
Highest level education, n (%)	
Compulsory education	0 (0)
Apprenticeship or Matura	19 (44)
Universities or Universities of Applied Sciences or Higher Vocational Education	24 (56)
Nationality, n (%)	
Swiss	36 (84)
Other (German, Great Britain, United States, Austria, and Italy)	7 (16)

[a]CHF: Swiss Francs.

Figure 2. Median weight (in kg) at different time points: month 0 (N=43), month 3 (N=40), and month 12 (N=36).



Outcome Measures: Other Metabolic Risk Factors

Weight loss observed among different timepoints appeared to have no effect on glucose metabolism (HbA_{1c}, blood glucose, and insulin) and blood lipids (triglyceride and HDL cholesterol). In contrast, a significant reduction in systolic BP was observed among all timepoints (−11.2 mm Hg between M0 and M12, *P*<.001), as well as in diastolic BP between M0 and M12 (−5.5 mm Hg, *P*<.001; see Table 2). The supportive analyses based on the PP population showed similar results to the ITT analysis. The results of anthropometric measurements and other metabolic risk factors at baseline (M0), 3 months (M3), and 12 months (M12) are presented in Multimedia Appendix 1.

Additional Outcome Measures

The SF-12 score (MCS and PCS) decreased during intervention, especially during the first 3 months (PCS: 53 to 55.2 and MCS: 53.2 to 54.9), but the changes were not significant (PCS: *P*=.08 and MCS: *P*=.09). The values for health-related quality of life at month 12 (PCS: 55.4 and MCS: 55.1) showed little change in comparison with the values at month 0 and the changes were also nonsignificant (PCS: *P*=.58 and MCS: *P*=.58). The overall result of GPAQ showed a small but nonsignificant increase in total moderate-vigorous physical activity metabolic equivalent minutes per week (MET-min/week) among different assessment times (median M0: 1920, M3: 2360, M12: 2740; M0-M3: *P*=.06, M3-M12: *P*=.78, M0-M12: *P*=.15). An analysis according to the 3 subcategories showed the following results. In subcategories *activity at work / travel to and from places*, no changes were apparent over the entire course of the year, whereas there were significant changes in the category *recreational activities* during this period (M0: 960 MET-min/week, M12: 1700 MET-min/week; M0-M12: *P*=.007). The results for certain eating habits (higher frequency of fruit, vegetable, and breakfast consumption and lower frequency of alcohol, sweet, and fat consumption) were significant between M0 and M3 and M0 and M12 and for salty snacks between M0 and M3. The total score for dietary consumption (median M0: 6 points, M3: 4 points, M12: 4 points) showed a significant reduction respective improvement toward a healthier diet between M0 and M3 (*P*<.001) and M0 and M12 (*P*<.001).

Table 2. Changes in anthropometric measurements and other metabolic risk factors between baseline (M0) 3 months (M3), and 12 months (M12).

Outcomes	Median	Range (min-max[a])	*P* value
Weight (kg)			
Difference between M0 and M3 (N=40)	−3.8	−15 to 2.4	<.001
Difference between M3 and M12 (N=36)	−1.1	−9.7 to 7	.08
Difference between M0 and M12 (N=36)	−4.9	−21.9 to 7.5	<.001
Body mass index (kg/m^2)			
Difference between M0 and M3 (N=40)	−1.4	−4.5 to 1.1	<.001
Difference between M3 and M12 (N=36)	−0.4	−3.1 to 2.3	.06
Difference between M0 and M12 (N=36)	−1.8	−6.9 to 2.5	<.001
Waist circumference (cm)			
Difference between M0 and M3 (N=40)	−3.5	−23 to 5	<.001
Difference between M3 and M12 (N=36)	0	−6.6 to 13.7	.72
Difference between M0 and M12 (N=36)	−3.8	−17.8 to 9	<.001
Body fat (kg)			
Difference between M0 and M3 (N=40)	−3.3	−10.6 to 2.5	<.001
Difference between M3 and M12 (N=36)	−0.6	−7.4 to 7.4	.14
Difference between M0 and M12 (N=36)	−4.0	−16.9 to 6.4	<.001
Body fat (%)			
Difference between M0 and M3 (N=40)	−2.3	−7.6 to 2.5	<.001
Difference between M3 and M12 (N=36)	−0.3	−9.9 to 5	.20
Difference between M0 and M12 (N=36)	−2.5	−11.9 to 3.7	<.001
Hemoglobin A_{1c} (%)			
Difference between M0 and M3 (N=40)	0	−0.5 to 0.3	.36
Difference between M3 and M12 (N=36)	0	−0.6 to 0.4	.84
Difference between M0 and M12 (N=36)	0	−0.5 to 0.5	.08
Blood glucose (mmol/L)			
Difference between M0 and M3 (N=40)	0	−1 to 0.6	.44
Difference between M3 and M12 (N=36)	0	−1.2 to 0.8	.66
Difference between M0 and M12 (N=36)	0	−1 to 1	.05
Insulin (mlU/L)			
Difference between M0 and M3 (N=40)	−1.35	−12 to 10.5	.39
Difference between M3 and M12 (N=36)	−0.15	−12.7 to 33.2	.69
Difference between M0 and M12 (N=36)	−1.35	−9.1 to 37.7	.33
Triglyceride (mmol/L)			
Difference between M0 and M3 (N=40)	−0.14	−2.09 to 1.81	.13
Difference between M3 and M12 (N=36)	0.01	−2.38 to 0.99	.61
Difference between M0 and M12 (N=36)	−0.14	−2.46 to 1.33	.07
High-density lipoprotein cholesterol (mmol/L)			
Difference between M0 and M3 (N=40)	−0.06	−0.64 to 0.3	.09
Difference between M3 and M12 (N=36)	0.11	−0.31 to 0.55	.01
Difference between M0 and M12 (N=36)	0.02	−0.32 to 0.78	.63

Outcomes	Median	Range (min-max[a])	*P* value
Blood pressure systolic (mm Hg)			
Difference between M0 and M3 (N=40)	−6.2	−28.4 to 20.7	.003
Difference between M3 and M12 (N=36)	−6.4	−31.3 to 22.3	.05
Difference between M0 and M12 (N=36)	−11.2	−29 to 10.3	<.001
Blood pressure diastolic (mm Hg)			
Difference between M0 and M3 (N=40)	−1.8	−22 to 13.7	.16
Difference between M3 and M12 (N=36)	−1.4	−26.3 to 16	.16
Difference between M0 and M12 (N=36)	−5.5	−20.3 to 12	<.001

[a]min-max: minimum-maximum.

Table 3. Percentage changes between baseline (M0), 3 months (M3), and 12 months (M12) in weight, body mass index, and body fat percentage.

Measurements	Median	Range (min-max[a])
Weight		
Percentage difference between M0 and M3 (N=40)	−4.6	−15.6 to 3.3
Percentage difference between M3 and M12 (N=36)	−1.5	−10.7 to 7.9
Percentage difference between M0 and M12 (N=36)	−6.0	−21.3 to 8.6
Body mass index		
Percentage difference between M0 and M3 (N=40)	−4.8	−15.6 to 3.9
Percentage difference between M3 and M12 (N=36)	−1.4	−10.7 to 7.9
Percentage difference between M0 and M12 (N=36)	−6.2	21.3 to 8.6
Body fat percentage		
Percentage difference between M0 and M3 (N=40)	−5.4	−25.5 to 6.8
Percentage difference between M3 and M12 (N=36)	−0.9	−23.9 to 13.7
Percentage difference between M0 and M12 (N=36)	−6.5	−35.8 to 9.8

[a]min-max: minimum-maximum.

Discussion

Principal Findings

This study investigated whether a 1-year remote counseling intervention by dietitian targeting weight loss was effective and feasible for people with overweight and obesity. The results show that this form of behavior counseling leads to significant weight loss, both in the first 3 months during an intensive remote counseling period, as well as over the entire course of 1 year. Weight loss also occurred between the third and 12th month, although this was not significant. The positive results in this study are in line with other studies investigating Web-based counseling and weight loss [17,31-33]. In 1 study, participants receiving remote support by weight loss coaches attained a mean (SD) change in weight from baseline of −4.6 (0.7) kg after 24 months, with weekly calls during the first 3 months and after that, with 1 call each month [17]. In a pilot randomized controlled trial with the aim to evaluate the feasibility and acceptability of a Web-based weight loss intervention in men with type 2 diabetes with feedback from a dietitian and exercise expert, the median weight loss was −4.3 kg (−7.8 to −1.0) from baseline to 12 months [31]. Other Web-based interventions that have focused on weight loss with shorter duration have demonstrated positive results [32-35]. For example, 1 study tested the effectiveness of a behaviorally-based mobile phone app combined with SMS text messaging from a health coach (participants could choose the frequency) and demonstrated significant weight loss after 3 months in the group with both mobile phone and health coach (−1.8 kg, *P*=.01) [33]. However, these interventions were only performed for 3 months and lack a long-term follow-up measurement. Changes in other anthropometric parameters associated with cardiovascular risk, including BMI, body fat percentage, and waist circumference were also significant in this study, similar to other remote interventions [32,33]. However, a direct comparison is difficult because of methodological differences such as duration, intensity, and frequency of counseling sessions among interventions. In addition, communication took place via different channels, such as SMS text messages, emails, phoning, and on the Web, and with counsellors who had different qualifications.

Over half of the participants lost at least 5% of their initial weight. This amount of weight loss is typically associated with positive clinical effects, namely an improvement in cardiovascular risk factors. A systematic review demonstrated

the following effects of dietary advice: reduction in systolic BP by 2.61 mm Hg (95% CI 1.31-3.91), diastolic BP by 1.45 mm Hg (95% CI 0.68-2.22), and 24 - hour urinary sodium excretion by 40.9 mmol (95% CI 25.3-56.5) after 3 to 36 months, between intervention and control group. However, no changes in HDL cholesterol, triglyceride levels, and fasting glucose were seen [36]. Although group comparisons were made in contrast to this pre- and postsingle arm study, the results indicate a similar trend. This study's results showed significant reduction in systolic BP and diastolic BP, but no significant reduction was seen in triglycerides, HDL cholesterol, HbA_{1c}, blood glucose, and insulin despite weight loss. One reason for this could be the small sample. The sample size calculation was based on the primary outcome. In addition, for some participants, the initial values of these parameters, particularly HbA_{1c}, were already within a normal range; therefore, a change in these values was not expected. Furthermore, the initially diagnosed cardiovascular risk factors (dyslipidemia or/and insulin resistance) in many participants could also have an impact on these results, although participants were free of serious diseases and only 5 participants took medication to treat hypertension and/or lipid-lowering medication during intervention.

Weight loss and improvement in other parameters were achieved by lifestyle intervention. The most important components of such weight loss interventions include a moderate calorie-reduced diet, increased physical activity, and the application of behavioral strategies to facilitate compliance with nutrition and activity recommendations [10]. In this study, the diet was analyzed with a very short frequency questionnaire to minimize the burden on participants. A qualitative analysis of individual food groups did not take place and energy intake was not measured. Nevertheless, positive changes were observed for most food groups and included increased intake of fruits and vegetables and lower intake of fat, sweets, and alcohol. These changes in food consumption would have resulted in lower energy intakes and consequent weight loss. An evaluation of the photo diaries and physical activity data available on the mobile phone were not conducted in the context of this study. However, this might lead to additional interesting findings. The measurement on movement behavior also indicated positive effects, particularly moderate and vigorous-intensity sports, fitness or recreational (leisure) activities in the *recreational activities* subcategory, whereas no changes were observed in the subcategories *activity at work / travel to and from places*. Increase in recreational activity is important as this form of activity may contribute to the prevention of cardiovascular diseases and cancer [37]. With regard to quality of life, no significant changes were observed for the PCS and MCS. This could be attributed to the very high quality of life of in Switzerland. In a similar study, the scores for MCS or PCS were lower and/or the initial weight higher and improvement mainly took place in the first half of the year [38]. In this study, the largest improvement in MCS and PCS occurred in the first 3 months.

The authors attribute the reasons for the successful weight reduction in this study to the following aspects. Dietitians used a wide range of behavioral change and counseling techniques (eg, self-monitoring, goal setting, relapse prevention, and motivational interviewing) and applied evidence-based nutritional knowledge according to individual needs (eg, energy density, meal rhythm, quality of carbohydrate, and adequate protein sources). They also provided simple recommendations for physical activity. The results of other studies demonstrate that interventions aimed at changing eating habits or sedentary lifestyle are more effective if they include behavioral change techniques [39,40]. Another review has confirmed the effectiveness of dietetic counseling for adults (not Web-based counseling) aimed at improving diet quality, diabetes, and weight loss outcomes [41].

The intensity of the intervention typically affects study outcomes [14]. In this study, the intensity of the intervention in the first half of the year might also have had an important impact (5 times per week in the first 3 months and 3 times per week during the following 3 months; average duration of 5 min per contact). Staying in contact with participants in the second half of the year to prevent relapses to maintain weight or further reduce it may also explain the successful weight reduction in this study [42]. Furthermore, a good working relationship with health care professionals is described as an important factor in long-term weight management [43]. Therefore, apps such as Oviva have the potential to play an important role in maintaining the relationship between client and dietitian and act as a simple means of supporting weight loss. In a Web-based counseling intervention for overweight or obese people, many effective components of a traditional face-to-face counseling session can be used. In addition, the app Oviva represents a simple and time-saving method of communicating eating behavior to the dietitians. Remote counseling can reduce barriers, such as lack of time, geographical distance, inconvenience, and embarrassment, that prevent possible clients from attending a face-to-face counseling session. This technique should also allow people with lower socioeconomic status to be reached. This is especially important as the prevalence of overweight and obesity is higher in this population group [3]. Although we tried to reach people with lower education levels through the obesity center, most people recruited to this study had undertaken higher education. The needs of people with lower socioeconomic status need to be better understood to facilitate access to such innovative and low-threshold solutions for weight loss.

Limitations

This study was a 1-arm pilot study; therefore, conclusions concerning effectiveness are limited. Generalization is limited because of the small sample size, individual focus in counseling, and the fact that it was a single-center trial. Randomized control trials are needed to examine efficacy and cost-effectiveness of Oviva and remote counseling in general. Pilot studies are described in the literature as an important first step in assessing the implementation of a novel kind of intervention [25,26]. As similar innovative solutions for use by dietitians have not yet been investigated in Switzerland or are rarely available, this pilot study was an initial milestone. By comparison, in the United States, for example, such tools have increasingly found their way into the everyday working life of dietitians [44].

For this study, it was important to ensure that the counseling process took place in a real-life setting. For example, the study physician held a brief motivating discussion during the clinical assessment at M3 with the participants about the course of the intervention. This might also have had a positive effect on weight progression. To meet the client's needs, small deviations from the contact frequencies were accepted. However, the tolerable limit was specified in the study protocol. Furthermore, 2 clients requested a change of dietitian as they were dissatisfied. This request was granted to prevent these participants from dropping out of the study. This has been considered in the PP analysis. Nevertheless, this indicates that the working relationship with the dietitian has an important influence on weight loss. In addition to the scientific and practical contents, the working relationship with the dietitian must be developed in a different way in comparison with face-to-face counseling. A qualitative study in which health care professionals were interviewed suggested that it is a challenge to establish and maintain an empathic relationship when conducting electronic health (eHealth) lifestyle coaching compared with face-to-face coaching [43]. As eHealth lifestyle coaching does not allow for facial expressions and gestures, specific competencies and experience in remote counseling are required to ensure that what would otherwise be transmitted in conversation is perceived by the client in the same way as it would in face-to-face coaching.

Conclusions

A mobile phone app for overweight or obese patients represents a new approach to weight loss counseling and has the potential to make a valuable contribution to tackling this important public health nutrition problem. This study has shown that in addition to the professional skills of dietitians, a profession-specific app such as Oviva can provide effective support that meets the needs of both dietitians and clients on the long path of behavioral change and sustainable weight reduction.

Acknowledgments

This study was cofunded by Innosuisse-Suisse Innovation Agency (project #17236.2 PFLS-LS) and Oviva.

Authors' Contributions

All authors were involved in the study design. SM-W led data collection. SH was responsible for statistical analysis. KH was responsible for study monitoring, performed interpretation, and drafted the manuscript. All other authors critically reviewed and revised versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

SM-W was paid for her functions as study physician by Oviva. For the study, registered dietitians were hired who performed remote counseling. They have also worked for Oviva. However, they received no additional payment by Oviva for the study.

Multimedia Appendix 1

Results of anthropometric measurements and other metabolic risk factors at baseline (M0), 3 months (M3) and 12 months (M12).

[PDF File (Adobe PDF File), 36KB - mhealth_v7i4e12289_app1.pdf]

References

1. World Health Organization. Obesity URL: http://www.euro.who.int/en/health-topics/noncommunicable-diseases/obesity [accessed 2018-09-20] [WebCite Cache ID 72ZdoaWLV]
2. Marques A, Peralta M, Naia A, Loureiro N, de Matos MG. Prevalence of adult overweight and obesity in 20 European countries, 2014. Eur J Public Health 2018 Apr 01;28(2):295-300. [doi: 10.1093/eurpub/ckx143] [Medline: 29036436]
3. Bochud M, Chatalan A, Blanco JM, Beer-Borst S. Anthropometric characteristics and indicators of eating and physical activity behaviors in the Swiss adult population: results from menuCH 2014-2015. Bern, Switzerland: Bern Open Repository and Information System: The Federal Office of Public Health and the Food Safety and Veterinary Office; 2017. URL: https://www.bag.admin.ch/dam/bag/it/dokumente/npp/forschungsberichte/forschungsberichte-e-und-b/schlussbericht-menuch-chuv.pdf.download.pdf/MenuCH1_Report_March2017.pdf [WebCite Cache ID 779pxaWbk]
4. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC Public Health 2009;9:88 [FREE Full text] [doi: 10.1186/1471-2458-9-88] [Medline: 19320986]
5. Hruby A, Manson JE, Qi L, Malik VS, Rimm EB, Sun Q, et al. Determinants and consequences of obesity. Am J Public Health 2016 Sep;106(9):1656-1662. [doi: 10.2105/AJPH.2016.303326] [Medline: 27459460]
6. Wu YK, Berry DC. Impact of weight stigma on physiological and psychological health outcomes for overweight and obese adults: a systematic review. J Adv Nurs 2018 May;74(5):1030-1042. [doi: 10.1111/jan.13511] [Medline: 29171076]
7. Specchia ML, Veneziano MA, Cadeddu C, Ferriero AM, Mancuso A, Ianuale C, et al. Economic impact of adult obesity on health systems: a systematic review. Eur J Public Health 2015 Apr;25(2):255-262. [doi: 10.1093/eurpub/cku170] [Medline: 25320051]

8. Tremmel M, Gerdtham U, Nilsson PM, Saha S. Economic burden of obesity: a systematic literature review. Int J Environ Res Public Health 2017 Dec 19;14(4) [FREE Full text] [doi: 10.3390/ijerph14040435] [Medline: 28422077]
9. Schneider H, Venez W. Cost of Obesity in Switzerland in 2012.: Federal Office of Public Health; 2014 Jun 16. URL: https://www.bag.admin.ch/dam/bag/fr/dokumente/npp/forschungsberichte/forschungsberichte-e-und-b/cost-of-obesity.pdf.download.pdf/cost-of-obesity.pdf [WebCite Cache ID 76YgbOU1B]
10. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. Circulation 2014 Jun 24;129(25 Suppl 2):S102-S138 [FREE Full text] [doi: 10.1161/01.cir.0000437739.71477.ee] [Medline: 24222017]
11. Yumuk V, Tsigos C, Fried M, Schindler K, Busetto L, Micic D, Obesity Management Task Force of the European Association for the Study of Obesity. European guidelines for obesity management in adults. Obes Facts 2015;8(6):402-424 [FREE Full text] [doi: 10.1159/000442721] [Medline: 26641646]
12. Leblanc ES, O'Connor E, Whitlock EP, Patnode CD, Kapka T. Effectiveness of primary care-relevant treatments for obesity in adults: a systematic evidence review for the U.S. Preventive Services Task Force. Ann Intern Med 2011 Oct 04;155(7):434-447. [doi: 10.7326/0003-4819-155-7-201110040-00006] [Medline: 21969342]
13. Webb VL, Wadden TA. Intensive lifestyle intervention for obesity: principles, practices, and results. Gastroenterology 2017 Dec;152(7):1752-1764. [doi: 10.1053/j.gastro.2017.01.045] [Medline: 28192109]
14. Patnode C, Evans C, Senger C, Redmond N, Lin J. Evidence Synthesis, No 152. Rockville (MD): Agency for Healthcare Research and Quality (US); 2017. Behavioral Counseling to Promote a Healthful Diet and Physical Activity for Cardiovascular Disease Prevention in Adults Without Known Cardiovascular Disease Risk Factors: Updated Systematic Review for the U.S. Preventive Services Task Force URL: https://www.ncbi.nlm.nih.gov/books/NBK476368/ [accessed 2019-03-01] [WebCite Cache ID 76YirvJdc]
15. Gupta H. Barriers to and facilitators of long term weight loss maintenance in adult UK people: a thematic analysis. Int J Prev Med 2014 Dec;5(12):1512-1520 [FREE Full text] [Medline: 25709786]
16. Metzgar CJ, Preston AG, Miller DL, Nickols-Richardson SM. Facilitators and barriers to weight loss and weight loss maintenance: a qualitative exploration. J Hum Nutr Diet 2015 Dec;28(6):593-603. [doi: 10.1111/jhn.12273] [Medline: 25231461]
17. Appel LJ, Clark JM, Yeh H, Wang N, Coughlin JW, Daumit G, et al. Comparative effectiveness of weight-loss interventions in clinical practice. N Engl J Med 2011 Nov 24;365(21):1959-1968 [FREE Full text] [doi: 10.1056/NEJMoa1108660] [Medline: 22085317]
18. Ash S, Reeves M, Bauer J, Dover T, Vivanti A, Leong C, et al. A randomised control trial comparing lifestyle groups, individual counselling and written information in the management of weight and health outcomes over 12 months. Int J Obes (Lond) 2006 Oct;30(10):1557-1564. [doi: 10.1038/sj.ijo.0803263] [Medline: 16534529]
19. Radcliff TA, Bobroff LB, Lutes LD, Durning PE, Daniels MJ, Limacher MC, et al. Comparing costs of telephone vs face-to-face extended-care programs for the management of obesity in rural settings. J Acad Nutr Diet 2012 Sep;112(9):1363-1373 [FREE Full text] [doi: 10.1016/j.jand.2012.05.002] [Medline: 22818246]
20. Ramsay LE, Ramsay MH, Hettiarachchi J, Davies DL, Winchester J. Weight reduction in a blood pressure clinic. Br Med J 1978 Jul 22;2(6132):244-245 [FREE Full text] [Medline: 678886]
21. Donaldson EL, Fallows S, Morris M. A text message based weight management intervention for overweight adults. J Hum Nutr Diet 2014 Apr;27(Suppl 2):90-97. [doi: 10.1111/jhn.12096] [Medline: 23738786]
22. Aguilar-Martínez A, Solé-Sedeño JM, Mancebo-Moreno G, Medina FX, Carreras-Collado R, Saigí-Rubió F. Use of mobile phones as a tool for weight loss: a systematic review. J Telemed Telecare 2014 Sep;20(6):339-349. [doi: 10.1177/1357633X14537777] [Medline: 24875928]
23. Fukuoka Y, Kamitani E, Bonnet K, Lindgren T. Real-time social support through a mobile virtual community to improve healthy behavior in overweight and sedentary adults: a focus group analysis. J Med Internet Res 2011;13(3):e49 [FREE Full text] [doi: 10.2196/jmir.1770] [Medline: 21752785]
24. Adolfsson B, Arnold M. Behavioral Approaches to Treating Obesity: Helping Your Patients Make Changes That Last. New York: American Diabetes Association; 2012.
25. Hassan ZA, Schattner P, Mazza D. Doing a pilot study: why is it essential? Malays Fam Physician 2006;1(2-3):70-73. [Medline: 27570591]
26. van Teijlingen E, Hundley V. The importance of pilot studies. Nurs Stand 2002;16(40):33-36. [doi: 10.7748/ns2002.06.16.40.33.c3214] [Medline: 12216297]
27. Paxton AE, Strycker LA, Toobert DJ, Ammerman AS, Glasgow RE. Starting the conversation performance of a brief dietary assessment and intervention tool for health professionals. Am J Prev Med 2011 Jan;40(1):67-71. [doi: 10.1016/j.amepre.2010.10.009] [Medline: 21146770]
28. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [Medline: 8628042]

29. Kolotkin RL, Haaz S, Fontaine KR. Measures of health-related quality of life in obesity eating disorders. In: Allison DB, editor. Handbook of Assessment Methods for Eating Behaviors and Weight-Related Problems. Measures, Theory, and Research. Los Angeles: Sage; 2009:33-78.
30. Brunner E, Domhof S, Langer F. Nonparametric Analysis of Longitudinal Data in Factorial Experiments. New York: Wiley-Interscience; 2001.
31. Haste A, Adamson AJ, McColl E, Araujo-Soares V, Bell R. Web-based weight loss intervention for men with type 2 diabetes: pilot randomized controlled trial. JMIR Diabetes 2017 Jul 07;2(2):e14 [FREE Full text] [doi: 10.2196/diabetes.7430] [Medline: 30291100]
32. Martin CK, Miller AC, Thomas DM, Champagne CM, Han H, Church T. Efficacy of SmartLoss, a smartphone-based weight loss intervention: results from a randomized controlled trial. Obesity (Silver Spring) 2015 May;23(5):935-942. [doi: 10.1002/oby.21063] [Medline: 25919921]
33. Stephens JD, Yager AM, Allen J. Smartphone technology and text messaging for weight loss in young adults: a randomized controlled trial. J Cardiovasc Nurs 2017;32(1):39-46. [doi: 10.1097/JCN.0000000000000307] [Medline: 26646593]
34. Alencar MK, Johnson K, Mullur R, Gray V, Gutierrez E, Korosteleva O. The efficacy of a telemedicine-based weight loss program with video conference health coaching support. J Telemed Telecare 2017 Jan 01 Epub ahead of print(forthcoming). [doi: 10.1177/1357633X17745471] [Medline: 29199544]
35. Hurkmans E, Matthys C, Bogaerts A, Scheys L, Devloo K, Seghers J. Face-to-face versus mobile versus blended weight loss program: randomized clinical trial. JMIR Mhealth Uhealth 2018 Jan 11;6(1):e14 [FREE Full text] [doi: 10.2196/mhealth.7713] [Medline: 29326093]
36. Rees K, Dyakova M, Wilson N, Ward K, Thorogood M, Brunner E. Dietary advice for reducing cardiovascular risk. Cochrane Database Syst Rev 2013 Dec 06;12:1465-1858. [doi: 10.1002/14651858.CD002128.pub5] [Medline: 24318424]
37. Wanner M, Tarnutzer S, Martin BW, Braun J, Rohrmann S, Bopp M, Swiss National Cohort (SNC). Impact of different domains of physical activity on cause-specific mortality: a longitudinal study. Prev Med 2014 May;62:89-95. [doi: 10.1016/j.ypmed.2014.01.025] [Medline: 24513168]
38. Sarwer DB, Moore RH, Diewald LK, Chittams J, Berkowitz RI, Vetter M, POWER-UP Research Group. The impact of a primary care-based weight loss intervention on the quality of life. Int J Obes (Lond) 2013 Aug;37(Suppl 1):S25-S30 [FREE Full text] [doi: 10.1038/ijo.2013.93] [Medline: 23921778]
39. Artinian NT, Fletcher GF, Mozaffarian D, Kris-Etherton P, Van Horn L, Lichtenstein AH, American Heart Association Prevention Committee of the Council on Cardiovascular Nursing. Interventions to promote physical activity and dietary lifestyle changes for cardiovascular risk factor reduction in adults: a scientific statement from the American Heart Association. Circulation 2010 Jul 27;122(4):406-441. [doi: 10.1161/CIR.0b013e3181e8edf1] [Medline: 20625115]
40. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
41. Mitchell LJ, Ball LE, Ross LJ, Barnes KA, Williams LT. Effectiveness of dietetic consultations in primary health care: a systematic review of randomized controlled trials. J Acad Nutr Diet 2017 Dec;117(12):1941-1962. [doi: 10.1016/j.jand.2017.06.364] [Medline: 28826840]
42. Middleton KR, Anton SD, Perri MG. Long-term adherence to health behavior change. Am J Lifestyle Med 2013;7(6):395-404 [FREE Full text] [doi: 10.1177/1559827613488867] [Medline: 27547170]
43. Brandt CJ, Søgaard GI, Clemensen J, Søndergaard J, Nielsen JB. Determinants of successful eHealth coaching for consumer lifestyle changes: qualitative interview study among health care professionals. J Med Internet Res 2018 Jul 05;20(7):e237. [doi: 10.2196/jmir.9791] [Medline: 29980496]
44. Stein K. Remote nutrition counseling: considerations in a new channel for client communication. J Acad Nutr Diet 2015 Oct;115(10):1561-1576. [doi: 10.1016/j.jand.2015.06.375] [Medline: 26271692]

Abbreviations

BMI: body mass index
BP: blood pressure
eHealth: electronic health
GPAQ: Global Physical Activity Questionnaire
HbA_{1c}: hemoglobin A_{1c}
HDL: high-density lipoprotein
ITT: intention-to-treat
M0: baseline or month 0
M3: 3 months
M12: 12 months or 1 year
MCS: mental health component summary score
MET: metabolic equivalent
PCS: physical health component summary score

PP: per-protocol
SMS: short message service

Edited by C Dias; submitted 21.09.18; peer-reviewed by L Mitchell, C Brandt; comments to author 12.11.18; revised version received 10.02.19; accepted 12.02.19; published 11.04.19.

Please cite as:
Haas K, Hayoz S, Maurer-Wiesner S
Effectiveness and Feasibility of a Remote Lifestyle Intervention by Dietitians for Overweight and Obese Adults: Pilot Study
JMIR Mhealth Uhealth 2019;7(4):e12289
URL: https://mhealth.jmir.org/2019/4/e12289/
doi:10.2196/12289
PMID:30973338

Original Paper

Physical Activity and Mobile Phone Apps in the Preschool Age: Perceptions of Teachers and Parents

Anna Ek[1,2], PhD; Johanna Sandborg[1,3], MSc; Christine Delisle Nyström[1], PhD; Anna-Karin Lindqvist[4], PhD; Stina Rutberg[4], PhD; Marie Löf[1,3], PhD

[1]Department of Biosciences and Nutrition, The Innovative Use of Mobile Phones to Promote Physical Activity and Nutrition Across the Lifespan Research Group, Karolinska Institutet, Huddinge, Sweden

[2]Division of Pediatrics, Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Huddinge, Sweden

[3]Division of Community Medicine, Department of Medical and Health Sciences, Linköping University, Linköping, Sweden

[4]Division of Health and Rehabilitation, Department of Health Sciences, Luleå University of Technology, Luleå, Sweden

Corresponding Author:
Anna Ek, PhD
Department of Biosciences and Nutrition
The Innovative Use of Mobile Phones to Promote Physical Activity and Nutrition Across the Lifespan Research Group
Karolinska Institutet
Huddinge,
Sweden
Phone: 46 8 585 866 02
Email: anna.ek@ki.se

Abstract

Background: Physical activity (PA) is already beneficial at the preschool age. In many countries, young children spend most of their days in the preschool setting, making it a common arena for PA interventions. Mobile health tools are becoming increasingly popular to promote PA in different populations; however, little is known about the interest for and how the preschool setting could incorporate such a tool.

Objective: This study aimed to examine how teachers and parents perceive PA in preschool-aged children in general and their perceptions of how a mobile phone app could be used to promote PA in the preschool setting.

Methods: Semistructured interviews were conducted with 15 teachers (93%, [14/15] women, mean age 43.5 years, 47%, [7/15] with a university degree and 10 parents [91%, 9/10] women, mean age 38.9 years, all with a university degree) recruited from 2 urban preschools in central Sweden. The interviews were recorded, fully transcribed, coded, and analyzed using thematic analysis by means of an inductive approach.

Results: The analysis revealed 4 themes: (1) *children are physically active by nature*, (2) *the environment as a facilitator or a barrier*, (3) *prerequisites of the adult world*, and (4) *an app in the preschool setting—challenges and possibilities*. Parents and teachers perceived preschoolers as being spontaneously physically active; however, high-intensity PA was perceived as low. The PA was specifically performed during the day in the preschool. Identified facilitators of PA were access to safe and engaging outdoor environments such as forests, spacious indoor areas, and adult involvement. Adult involvement was considered especially important for children preferring sedentary activities. Identified barriers for PA were restricted indoor and outdoor space, rules for indoor activities, and lack of adult involvement because of time constraints. The teachers perceived that they had limited skills and experiences using apps in general, although they also acknowledged the increasing role of technological tools in the curriculum. Thus, the teachers expressed an interest for an app designed as a support tool for them, especially for situations when PA was limited because of perceived barriers. They suggested the app to include accessible information regarding the health benefits of PA in children linked to a library of activities for different settings and seasons. Parents suggested interactive app features including problem-solving tasks and music and dance, but not video clips as they made children passive.

Conclusions: Vigorous PA was perceived as low in preschool-aged children. Future tailoring of interventions in the preschool setting should work around barriers and support facilitators to PA, especially PA of high intensity. In such work, an app could serve as a source of inspiration for PA in different ages, settings, and seasons and thus reduce environmental and structural inequalities in the preschool setting.

(JMIR Mhealth Uhealth 2019;7(4):e12512) doi:10.2196/12512

KEYWORDS

child, preschool; mHealth; physical activity; parents; school teachers; qualitative research

Introduction

Physical Activity Levels in Preschoolers

Higher physical activity (PA) levels are associated with improved metabolic and psychosocial health as well as motor and cognitive development in preschoolers [1]. In addition, preschool-aged children who spend more time in vigorous PA have a more favorable body composition and better physical fitness (ie, cardiorespiratory fitness, motor fitness, and muscular strength) [2]. Establishing a habit of being physically active during childhood increases the likelihood of being physically active as an adult [3,4]. Globally, data show that many children aged 5 to 17 years do not meet the recommended levels of 60 min of moderate-to-vigorous PA (MVPA) per day [5], and Sweden is no exception [6]. Few studies have investigated the fulfillment of PA in the newly developed 24-hour movement guidelines for the early years, that is, 180 min of total PA, of which at least 60 min is MVPA [7-9]. However, previous Swedish data have shown that only one-third of 4-year-old children reached 60 min per day of MVPA [10]. Correspondingly, another recently published Swedish study showed that children only spent 7 and 12 min per day on average on vigorous PA at 4 and 5 years of age, respectively [2]. Considering the pronounced negative health effects of physical inactivity, actions to improve children's PA levels are required, especially activities of higher intensity.

Traditionally, PA interventions targeting preschoolers use the preschool environment [11,12]. The design of such interventions has primarily been face-to-face education combined with written information and interactive pedagogic materials [13]. This intervention design has proven well accepted, although results vary [14,15]. The advantage of interventions in the preschool setting is that it reaches all children. In Sweden, 84% of all 1- to 5-year-old children spend a significant part of their day at preschool [16]. Furthermore, and contrary to many other settings for interventions, preschools also reach children from different sociodemographic backgrounds. For example, approximately 78% of all 1- to 5-year-old children with a migrant background attend preschool [16]. However, traditional interventions are often difficult to scale-up as face-to-face education require extensive resources, and thus, alternative delivery options should be considered.

Mobile Health to Promote Physical Activity in Preschoolers

In the past decade, interest in the use of mobile phones for delivering lifestyle interventions has increased [17-21]. In comparison with traditional lifestyle interventions, mobile health (mHealth) programs are more flexible and can be tailored to meet individual needs [17]. Moreover, mHealth interventions have the advantages that they can be delivered anywhere and at any time and they can easily be scaled-up [17,18]. However, few mHealth interventions exist for promoting PA in the preschool age group, and so far, they have only targeted parents [22]. In Australia, an intervention aimed to decrease screen time in preschool children aged 2 to 4 years by sending supportive text messages (short message service, SMS) to parents [23]. The study design proved effective with increased PA in the children and was appreciated by the parents [23]. Similarly, another intervention included delivering text messages to parents to support healthy habits as an intervention for children with obesity [24]. The intervention was well accepted, although no effect was seen for child weight status or behavior [24]. In Sweden, the Mobile-based intervention intended to stop obesity in preschoolers (MINISTOP) study used a mobile phone app to promote healthy eating and PA in 4.5-year-old children [25]. Although no significant difference in weight status was found between the intervention and control group, the intervention group improved their composite score based on fruit and vegetable consumption, PA, and fat mass index. Interestingly, this improvement was higher among children with a higher fat mass index [25]. Importantly, all the aforementioned mHealth interventions were well accepted and used by the parents. Taken together, the results from these interventions suggest that an app targeting parents can be used to influence lifestyle behaviors in young children. However, such apps may also target preschool teachers who spend many hours with the children, and tablets are already being used for pedagogical purposes within the Swedish preschool curriculum.

To develop an app for the preschool setting, we need to know more about the interest in such a tool as well as appropriate content and features to include and to whom the app should be designed for. Thus, a first step is to collect such information from potential end users, that is, teachers and parents. Including the views of end users can improve the intervention's compatibility and increase the likelihood that it will be effective and sustained [26-28]. Furthermore, information regarding perceptions of PA in general needs to be collected to ensure the app reflects the users' expectations. For example, this could include the following: what are the types of PA that is conducted currently, what activities are preferred by children and teachers, what role do adults have in promoting PA in preschoolers, and what are the facilitators and barriers of PA. Therefore, the aims of this study were to explore how teachers and parents perceive PA in preschool-aged children and how a mobile phone app may be used to promote PA in the preschool setting.

Methods

Recruitment

A total of 2 preschools representing 2 cities (Norrköping and Linköping) in central Sweden were asked to participate in this study. The preschools were chosen based on previously established contact between the research group and the head of the preschool. The preschools differed in size (48 children in Linköping and 30 children in Norrköping). In Linköping, most of the parents were born in Sweden; however, in Norrköping,

one-third of the children had parents born outside of Sweden. All staff at the 2 preschools were informed about the aims of this study and were invited to participate in semistructured interviews. To recruit parents, posters with information about the study were put up on bulletin boards at the 2 preschools so that interested parents could contact the research team. The teachers' interviews were coordinated between the research staff and the head of the preschool. The parents emailed the researchers so that interviews could be scheduled at an appropriate time. The interviews were conducted between January and April 2017. The study was approved by the regional ethical board in Stockholm (October 10, 2016, dnr: 201612099-32). Informed written consent was obtained from both the parents and teachers participating in the semistructured interviews. The reporting of this study follows the Consolidated criteria for reporting qualitative research checklist; see Multimedia Appendix 1 [29].

Semistructured Interviews

The teachers and parents were interviewed individually in a separate room at the preschool where they worked or had their child enrolled. The interviews were conducted in Swedish face-to-face by the female researchers AE and JS and lasted for 12 to 42 min. For educational purposes, AE, postdoctoral fellow and registered dietician with previous experience in performing semistructured interviews with adults regarding children's lifestyle behaviors, conducted the interviews in Norrköping, whereas JS, nutritionist and master student, observed. JS conducted the interviews in Linköping. All participants were asked the same set of main questions (see Multimedia Appendix 2), with follow-up questions tailored to individual responses following a semistructured design. The main questions were developed and discussed in our research group, which has extensive experience with PA interventions and mHealth for preschoolers. The interviews were recorded (audio) and then fully transcribed by JS.

Before the interviews started, the interviewees were given a short introduction about the aim of the study and the research group. Teachers and parents answered questions regarding their age, sex, country of birth, and education level. In addition, parents gave information about how many children they had and the children's ages.

Data Analysis

To analyze the interviews, thematic analysis was conducted, which is a commonly used method to describe qualitative data [30]. Thematic analysis is a flexible and useful tool to clearly summarize key features and provide a description of the dataset [30]. The transcribed material was read and reread in an active way by AE and JS and then separately coded using an inductive approach (ie, data-driven; without trying to fit data into a pre-existing coding frame or the preconceptions of the researcher). When all of the data had been coded, the process of finding themes was initiated. The codes and the interpretation of the data were discussed between the 2 coders over the phone and during face-to-face meetings as well as in email correspondence. Disagreements between the coders were resolved in discussions before the codes were sorted into potential themes. The themes were identified at a semantic level (ie, it is what the participants explicitly said that is being analyzed and not the potentially underlying meaning) [30].

Descriptive data for sociodemographic variables (means, SDs, and percentages) were analyzed using IBM SPSS statistics version 23 (IBM, Armonk, NY, USA).

Results

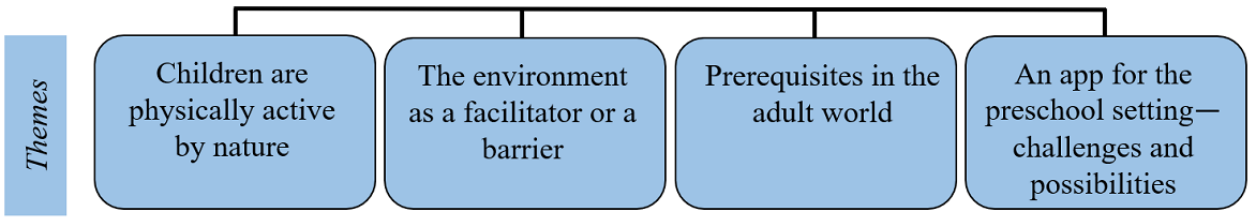
Participants

All but 1 (because of illness) of the consulted teachers agreed to participate in the interviews (N=15). The majority of the teachers were women (93%, 14/15) and born in Sweden (93%, 14/15). Almost half of the teachers had a university degree (47%, 7/15), and the other half had up to a 3-year secondary school degree (53%, 8/15). The mean age of the teachers was 43.5 years (SD 13.9; min-max 20-65 years). All of the participating parents were born in Sweden (N=11), the majority were women (10/11, 91%), and all had a university degree. The mean age of the parents was 39.8 years (SD 5.2; min-max 32-47 years). The parents had, on average, 2.4 children (min-max 1-5) between the ages of 1 and 15 years. The mean age of the children attending preschool (n=15) was 3.7 years (min-max 2-5 years).

Themes From the Thematic Analysis

A total of 4 themes were identified in the thematic analysis (see Figure 1). Each quote has been labeled with the following: P for parent or T for teacher, the participant number, and M for male and F for female.

Figure 1. The themes identified in the thematic analysis.



Children are Physically Active by Nature

The preschool was considered by both teachers and parents as the most important setting for children's PA because of the long hours spent there. Teachers and parents also expressed how preschoolers are spontaneously physically active and move around all the time. However, they also indicated that there was room for more high-intensity PA as it was perceived as low by both teachers and parents. There was a consensus about the need of everyday PA for children's overall physical and psychosocial health and development. Among the positive effects of PA, teachers found active play, preferably outdoors, fundamental for children to be able to concentrate on less active indoor activities during the day. After active play, the environment often became more harmonious with less conflicts and more responsive children. The effect on the children's mood was also experienced by parents who expressed that there were less conflicts and overall better family dynamics when they had spent time outdoors during weekends:

> *On the weekends we always try to spend time outdoors. We have noticed such a big difference in all our children that the days when we get outside in the morning, they are so much better...They (the children) are happier, have more energy, have plenty of ideas of what we can do, are more creative and want to play. You know, more balanced in every way, easier to deal with. If we don't, there is more nagging, fuss and arguments between them and they don't have the energy, and they don't really want to...* [P8F]

Teachers and parents also agreed that children need a variety of PA, both spontaneous and more structured and planned activities. However, the characteristics of the child often decided what type and intensity of PA were preferred, as all *children are different*. Teachers described how the need for physical challenges increased as the children grow older. In addition, teachers suggested active transportation (eg, walking as transportation) as a natural way to increase PA. However, they perceived that today's children were not used to active transportation and that the children found walking both difficult and exhausting. A contributing factor for this perceived change over the decades was suggested to be that parents often drive the children by car or push them in the stroller. Thus, the majority of the teachers thought that children need to practice walking and also be challenged in uneven and hilly environments to strengthen their muscles. This was in some ways supported by the parents, and the lack of active transportation was often explained by a lack of time or practical limitations:

> *I believe children need to be active in hilly terrain to strengthen their muscles. I have noticed that the children have poor muscle strength, they get exhausted and tired right away when you are out for a walk...* [T15F]

The Environment as a Facilitator or a Barrier

According to teachers and parents, PA was typically associated with factors in the outdoor environment such as a spacious yard, hills, open areas (eg, a soccer field), trees to climb on, and paved areas to ride tricycles on. Under these conditions, the teachers expressed that they could engage children in PA all year round; however, in the winter, many layers of clothing could sometimes limit movements for the youngest children. Contrary to the facilitators for PA, a small flat yard with no open space was perceived to have a negative impact with regard to children engaging in PA, which was the situation for 1 of the preschools. Thus, the outdoor environment and the location of the preschool were brought up as important factors when planning for an optimal preschool environment. For example, the forest was mentioned as the perfect place for diverse activities by both teachers and parents and a popular place to visit. Thus, the location of the preschool decided if visits were possible on a regular basis. In addition, parents expressed that a backyard or other safe outdoor areas and playgrounds as well as living in a larger home offered opportunities for the children to play more freely and therefore be more physically active:

> *I wish we had a more encouraging yard...more opportunities for climbing and doing balance exercises...a hill where the children could run up and down for instance, now it's just a small flat yard...* [T3F]

Indoors, teachers described children to be less active; however, this variation in PA was described as important for the development of fine motor skills (eg, drawing and putting together jigsaw puzzles). Teachers and parents perceived limited indoor space as the most important barrier to PA, as the activities often became disturbing, noisy, and even unsafe. To maintain a pleasant and safe indoor environment, calmer types of activities were encouraged, and the children were often divided by teachers into smaller groups engaging in different types of activities. In the preschool situated in the larger facility, the teachers described that the children could be taken to different rooms to perform different activities. For example, to the music room or a larger room where the children could engage in active play. In the other preschool, teachers found it more difficult to encourage indoor PA because of the limited space. Teachers explained that with larger groups of children, they more often encouraged calm activities:

> *It easily gets messy when there are many children indoors and it gets noisy in a different way, sonically, and we don't really have access to larger facilities that are often required.* [T16F]

Prerequisites in the Adult World

The teachers described that the daily operations in the preschool were governed by several factors: the preschool curriculum and learning goals, the applied pedagogical approach, time constraints, the possibility to get temporary staff at times of illness, and the number of children. Low staff density made it difficult for teachers to attend to the PA needs for all children. For instance, trips to the forest or other planned activities could be canceled on these occasions. A further consequence of low staff density was the lack of time and opportunity to plan and come up with new engaging activities for the children.

The 2 preschools investigated had different pedagogical approaches: 1 was more permissive when it came to PA indoors, whereas the other one had stricter rules of how the children

could be active indoors. The latter also had a clear schedule for mornings devoted to calmer indoor activities (ie, Montessori shift) and afternoons spent outdoors (ie, free play). The teachers explained that there were many goals in the preschool curriculum that they were obligated to fulfill, which could also stand in the way for PA. Thus, a way to include PA with a pedagogical approach was for teachers to use rhymes and fairy tales that encouraged the children to be physically active. The guidelines and policies from higher authorities impacted the focus in the preschool, making it easier to implement something if it had precedence from a higher authority. However, a teacher clarified this statement by saying that even though there were many goals to fulfill, the means of how the preschool chose to do it was up to the teachers, thus leaving them with opportunities:

> *I love to go with the children to the forest. It doesn't happen that frequently now since I work with the youngest children and because there are so many children in the group. Therefore, we don't get to go so often, but I really enjoy going to the forest. In the forest, you discover things together. With the older children there is more free play, we build forts and you can see their imagination take over and they climb. I also think it is fun to be in the tobogganing hill, the activity there feels like good quality and you can see the children run up and down. They exhaust themselves a bit and I think that's good.* [T16F]

> *...to see the joy in the children's eyes when they (realize) "I can, I could climb up" or "I can run this far" that is also really, really fun...* [T5F]

The long hours at the preschool made parents and teachers agree on the important role of the teacher to facilitate PA in preschoolers. Teachers described how their own interest, attitude, and awareness regarding PA affected both the duration and types of activities offered in the preschool:

> *I really enjoy going tobogganing with them (the children), and I think it's really fun to bicycle together because then we (the teachers) also go bicycling on the small bikes and you get this close contact with the children...I might not be a big fan of soccer because I'm not good at it, but it's a lot of fun when we play music outdoors and do gymnastics together out in the yard, that's really fun.* [T10F]

Although PA was part of the curriculum, teachers described how it was their individual characteristics that decided to what extent PA was used as a means to meet other curriculum goals. To make PA less dependent on individual characteristics of the staff, they highlighted the importance of planning PA together in the staff group and prioritizing PA in the weekly schedule. Several teachers expressed how it was their responsibility to acknowledge the various needs and characteristics of the children, for example, to encourage sedentary children to engage in more active play and to challenge children who are already physically active:

> *I think most of it (PA) depends on the adults, what attitude you as an adult have to a certain activity. If you yourself think that it is fun, then you make sure that the children get involved. Because there are opportunities everywhere even if you don't have a gym and those kind of things. There are opportunities, if you have imagination, and then it depends on the adults.* [T11F]

Both teachers and parents described how children find activities more fun when adults participate and how, for some children, adult initiatives were crucial to getting them involved in active play. Parents created opportunities for PA by being outdoors with their children and buying special equipment. Signing their child up for organized PA was another support parents offered, such as gymnastics or dance classes. Contrary to this, parents also described themselves as the biggest barrier to their child's PA. Parents' busy work schedules and children spending long hours at the preschool left little or no time, nor energy, for PA during weekdays. The ambition of being outdoors during the weekends was also a challenge as other things tended to get in the way. Similar to the preschool, house rules could also limit the children's ability to be active as parents often found it noisy and disturbing:

> *It is more often that we as parents limit the children; when they want to start moving and play–the kind of games that we often, as adults, limit by saying "no, indoors we walk", we limit them in their play because it takes up too much space or gets too loud and noisy, even though they need it daily.* [P20F]

An App in the Preschool Setting—Challenges and Possibilities

All teachers except 1 stated that they had not yet started working with tablets in their organization. Most of the teachers expressed that they felt uncertain and not comfortable using tablets and apps. The variation in technical skills among the staff affected the implementation of the use of tablets. Teachers explained that it would also take time for the children to get used to the tablets as a component in the preschool and to learn how to use it as intended in the curriculum:

> *I find it a bit difficult to see (the need)...I struggle to find the role for Ipads in the preschool setting...* [T14F]

Even though some could not see the need for an app in this age group, others did come up with ideas on how an app could contribute to the promotion of PA in the preschool setting. Suggested content for the app was information regarding the positive benefits of PA for children, as this was regarded as a motivating factor for teachers to use the app. Practical tips on active play including music and dance and how to create a stimulating environment for PA were also desirable attributes to include in an app. The teachers also stressed that it was necessary that the app could be used in several different settings: indoors, outdoors, in the forest, in the yard, different seasons, for different ages, and various numbers of children:

> *...an app with features for when you are in the forest or when you are in the yard or in a large room to play...or if you have a small room, what you could do there...that would be very interesting...* [T14F]

The parents were positive toward the use of an app in the preschool setting. Furthermore, the parents identified role models such as a popular personality from a children's TV program or other children in the same age group as important elements within an app to engage the children. Some parents also described interactive games involving treasure hunts and reward systems (eg, Pokémon Go) as potential app features:

> *Gathering things to go on treasure hunts are fun and some kind of positive reinforcement like a point system, also, Pokémon Go has a nice design and fun figures–my five-year-old likes it.* [P21F]

Another function that was suggested was the use of videos as a way of demonstrating PA and various exercises. However, video clips in the app revealed some contradiction, as screen time, in general, was a concern for the effect it could cause on the child's imagination, creativity, and own initiatives for other activities. One parent stressed that video clips included in an app promoting PA should not be too captivating as this could instead turn the child into an observer, whereas a teacher expressed that videos can be a source of inspiration:

> *If we put on music and just dance together then we have a really good time, it's a spontaneous activity, but as soon as a video comes on they stop. Even though there's music they become passive observers.* [P8F]

> *An app that can inspire through videos of different activities and in the preschool both how you can do it indoors and outdoors and pictures and maybe also that these videos are with other children so that the children here can relate to themselves and to the preschool setting.* [T19M]

The idea of who should use an app differed between the teachers. Some thought that it should only be used as a tool by the teachers, whereas others thought it should be used by both the teachers and the children. However, all teachers agreed that the children should be involved in some way and that the use of tablets and apps should have a pedagogical value and not be used the same way as it is in the home environment (ie, mostly sedentary).

Discussion

Principal Findings

In this study, we interviewed parents and teachers regarding their perspectives of PA in preschoolers. The responses regarding PA were then followed with questions about what role a mobile phone app could have to support PA in the preschool setting and what features such an app should include. The interviews revealed 4 themes: *children are physically active by nature, the environment as a facilitator or a barrier, prerequisites of the adult world, and an app in the preschool setting—challenges and possibilities.* We found an understanding for the important role of PA for optimal child development in our sample. The long hours children spend in the preschool highlighted the important role of the preschool teachers for sufficient PA. However, both parents and teachers found preschoolers to engage in little vigorous PA and identified environmental and structural facilitators and barriers for PA. To overcome the proposed barriers, creative solutions for how to support preschoolers' PA were welcomed. A mobile phone app promoting PA in the preschool setting was seen as a viable option to support teachers.

Facilitators and Barriers to Preschoolers' Physical Activity

Our results confirm and extend the results from a recently published review by Hesketh et al [28] on qualitative research regarding parents' and care providers' perceptions of facilitators and barriers on preschoolers' PA [28]. First, in our study, the preschool was perceived to be the most important setting for children's PA because of the long hours spent there, which is in line with previous research [28,31]. This is especially important to recognize when performing interventions in countries where children spend most of their days at preschool, such as in Sweden. Furthermore, PA in this young age group was generally regarded as something that occurred naturally and spontaneously and that children engage in PA if opportunities are provided. However, both teachers and parents agreed that higher-intensity PA and more diverse and challenging PA (eg, climbing and balancing) could be offered to children more often. The reason for it not being offered was often because of environmental and structural barriers (eg, small indoor and outdoor facilities, low staff density, and large groups of children) and time constraints (eg, little time to plan for PA). Contrary to our findings, in the review by Hesketh et al [28], the weather (eg, children being affected by too hot or too cold weather or children getting dirty outdoors) and safety (eg, when climbing) were also recognized as barriers to vigorous and challenging PA [28]. All these barriers need recognition, as reports from different countries show that the level of PA in preschools is low and the levels of sedentary behaviors are high [8,10,32] and interventions are warranted. The solutions on how to promote high-intensity PA may actually not be very difficult to find. Positive findings from the mHealth trial, MINISTOP, have shown that relatively small changes in the intensity of PA can lead to improvements in preschoolers' body composition and physical fitness [2]. More specifically, this was observed when 5 min per day of sedentary behavior and low PA were substituted for 5 min of vigorous PA [2]. Increasing vigorous PA by at least 5 min per day could easily be implemented in preschools given the long hours the children spend there. Thus, an app intervention including a short daily exercise program has the potential to help in making this change. In addition, this structured form of an aerobic program tailored for young children was appreciated by the teachers in our sample. Furthermore, such a program would also support teachers as role models for PA and to succeed in what the teachers in our sample perceived as their responsibility, to ensure sufficient daily PA for all preschoolers to complement the low levels of PA observed in the home environment [31,33-35].

Our findings also showed how the teachers may become a barrier to PA, especially in the indoor setting where active play easily became loud and rowdy, which increased the risk of children getting hurt. However, if the facilities had more space or had an indoor playroom where the children were allowed to run around, teachers acknowledged how they would like to

encourage active play more indoors, especially during winter. The preschool's structures (eg, the pedagogical approach, ie, stricter rules for PA indoors) and the comprehensive preschool curriculum (ie, many goals to fulfill) also made the teachers to act as a barrier for PA. The ambiguous role of teachers was also recognized by Hesketh et al [28]. Training of staff has been suggested to facilitate PA in the preschool setting and even out the impact of individual characteristics among the teachers [15,36-38]. Although staff training was mentioned in our sample, the participating teachers especially requested more time for planning activities as well as an increased staff density as ways to promote PA. It is possible that an app could be a support in situations when time and space are limited. Previous research found that text messages were appreciated by parents and were effective in making preschoolers more active [23,24]. Therefore, a text message function with suggestions for daily activities could also be an option in the preschool setting to counteract the environmental and structural barriers.

Teachers found it especially challenging to support the older preschoolers' need for more diverse and challenging PA that requires a larger and more challenging yard (eg, hills and open spaces with trees) or regular trips to parks, fields, and the forest. The difficulties in keeping older preschoolers active were recognized in a previous study showing that preschoolers' PA decreases with every yearly increase in age [33]. The facilitators to PA found in this study are consistent with the existing literature, suggesting that preschools with more supportive environments (eg, portable and fixed play equipment, staff education and training, positive engagement by teachers in the playground, less children per square meter in the yard, and the presence of vegetation and open play areas) increase vigorous and overall PA levels in children and decrease sedentary behaviors to a greater extent compared with preschools with less supportive environments [15,36-38]. If the environmental restrictions bring on a decrease in PA already in this young age, a first step for interventions would be to find accessible creative solutions, preferably low cost, to support factors of greatest importance to PA as perceived by the target population [28]. This was also what our sample recognized as a role for an app in the preschool setting.

An App Promoting Physical Activity in the Preschool Setting

To our knowledge, this is the first study to investigate if and how an app could be used to increase PA in the preschool setting and what specific features to include; thus, no comparable data exist. Below, we summarize suggestions based on our findings for specific app features and highlight some of the possibilities and challenges with regard to implementing and using apps in the preschool setting. Our results show that the idea and use of any apps in the preschool setting were still new to most of the teachers, even though the current Swedish preschool curriculum includes the usage of technological tools [39]. Given the diverse technical skills and lack of time for planning in the preschool, apps targeting the preschool environment need to be pedagogical and user-friendly to maximize usage. Hence, when developing an app for this setting, it is particularly important that the formative work is conducted through close collaboration between teachers and app developers at an early stage to ensure that the specific user requirements are fulfilled. In addition, when developing an app for use in the preschool environment, we need to be cautious so that we do not introduce screen-related issues such as conflicts between children and between children and teachers, reduce the children's own creativity for other activities, and features that could make children inactive (eg, watching video clips). For these reasons, the teachers suggested that the app should primarily be aimed at them. An app would benefit from including accessible information about the health benefits of PA in children linked to a library of activities promoting child development in different ages, for example, activities supporting motor skills and cognitive development such as different types of music and dances with diverse coordination practices. As discussed above, the suggested activities would preferably be independent of environmental and structural barriers to PA. For these features, the teachers should be used as a source of knowledge and creativity as rhymes and interactive storytelling are elements they often use in their daily work. Features that could be directed to the children to get them moving and interacting at the same time were suggested; however, this depends on the child's age. An example of such an app with such features is Pokémon Go [40]. In a qualitative evaluation of what made children appreciate the Pokémon Go app, the strongest and most prevalent answer was the collaboration with others [41]. Thus, gamification elements might be interesting to incorporate into an app. Finally, it is relevant to note that costs for the tablets were not identified as a barrier in our study. Most likely because Sweden, as a very high-income country, already uses technical devices as part of the curriculum; however, the high cost of these devices may be an issue in other countries.

A theoretical framework for guidance in the development of an app was not acknowledged by the parents and teachers in our sample; however, this is recommended for interventions [42]. The social learning and social cognitive theory are often used in interventions targeting PA and children [43-45]. These theories would support the features suggested in the interviews, for example, teachers as role models and children interacting in a dance or a game and being rewarded and getting feedback for achievements in interactive games. In summary, an app promoting PA in the preschool setting is warranted, especially as an easy and ready-to-use support tool for teachers to find solutions for the children to be active when the environment is not ideal.

Strengths and Limitations

A strength of this study is the qualitative approach, which is a tool for developing feasible and sustainable interventions in different populations [46]. Furthermore, the semistructured interviews enabled a more detailed and rich description of the participants' perceptions compared with questionnaires that leave participants with little room to elaborate on their opinions. In addition, the wide age span of teachers allowed us to capture perspectives from those with various amounts of work experience and views and use of apps. Limitations are the convenience sample of 2 preschools: both were situated in affluent areas and participating teachers and parents were well educated. It is possible that this distribution in education had an impact on the results, and it would have been valuable to

broaden the interviews to a more diverse sample. The homogeneity of the parents might have been because of the way they were recruited, as they approached us after reading about the study on the preschool's bulletin board rather than all parents receiving a personal request. Although our sample was homogenous, we obtained a saturation of responses within this group. Furthermore, the factors presented do represent what other similar studies have found in more heterogeneous populations [28]; however, some differences are worth mentioning. First, and somewhat unexpectedly, the weather and clothing aspects were not discussed to a large extent as a barrier or facilitator to children's PA. This factor was addressed in a Canadian study with a sociodemographic diverse sample of child care providers [47]. It is possible that a Swedish preschool situated in a more ethnically diverse area with parents not yet used to the Swedish climate and conditions in different seasons would have addressed clothing as a greater issue. Second, gender differences for PA were not discussed in our sample, but this has been raised in previous qualitative research where boys were perceived as more active than girls [28]. In addition, in a Malawian study, although PA was seen as positive for both boys and girls, some parents thought that older girls should not be too active when growing up as this might not be culturally acceptable [48]. Thus, culture differences regarding PA in boys and girls are important to acknowledge when performing interventions in areas where children represent many cultures so that misconceptions can be avoided. Finally, the skewed gender distribution in our sample is a limitation. However, for the teachers, it is similar to the distribution in Swedish preschools (ie, only 4% of the annual employees constitute of men) [16]. The absence of males working in the preschool setting was also recognized as a gap in the literature by Hesketh et al [28]. As research has shown that fathers play an important role for children's PA in the home environment [49], it is possible that male caregivers could have a positive impact on children's PA in the preschool setting; however, this needs further investigation.

Future Perspectives for an App in the Preschool Setting

There are a lot of reports showing that screen-based behaviors are becoming more prominent in younger ages, with preschoolers spending about 2 hours per day in front of a screen [8,10,50]. Therefore, the creation of an app targeting this age group is somewhat contradicting. However, an app encouraging children to be active may serve as a competitor to apps promoting sedentary behaviors in children, and it may also provide important information and support to adults' promotion of PA in this age group. The preschool environment is an essential arena from where we can work to prevent lifestyle inequalities in young children. This is especially important as the level of childhood obesity in the preschool age group is a global concern [51]. Finding solutions to work against impaired health in young children must be a global priority. Thus, providing preschools with optimal conditions for PA is key. As changing the physical environment of already established preschools is not always feasible, an app could support teachers in the preschools to provide sufficient daily PA of different intensity regardless of physical and structural barriers. However, the development of such app features needs to be investigated further and tested in heterogeneous samples.

Conclusions

This study identified facilitators and barriers for PA in preschoolers and provides suggestions for how an app can be used in the preschool setting to promote PA. The children spent long hours at preschool, and thus, the majority of their PA was performed there. Given opportunities, children were spontaneously active, although the amount of higher-intensity PA was perceived as low by teachers. Small indoor areas and a restricted outdoor space were identified as the main barriers for PA. Facilitators for PA were engaging teachers, high staff density, and scheduled time for planning PA. Future tailoring of interventions in the preschool setting needs to work around these barriers and support the facilitators. In such work, an app could serve as a source of inspiration for PA in different settings and thus reduce environmental and structural inequalities in the preschool setting.

Acknowledgments

The authors would like to thank the teachers and parents who participated in the interviews.

Authors' Contributions

ML designed the study together with AE and JS. AE and JS recruited participants, performed the interviews, coded the interviews, conducted the thematic analysis, and wrote the manuscript. ML, CDN, AKL, and SR critically reviewed the manuscript. All authors approved of the final draft of the manuscript as submitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The consolidated criteria for reporting qualitative research checklist.

[PDF File (Adobe PDF File), 499KB - mhealth_v7i4e12512_app1.pdf]

Multimedia Appendix 2

Interview questions to teachers and parents.

[PDF File (Adobe PDF File), 38KB - mhealth_v7i4e12512_app2.pdf]

References

1. Carson V, Lee EY, Hewitt L, Jennings C, Hunter S, Kuzik N, et al. Systematic review of the relationships between physical activity and health indicators in the early years (0-4 years). BMC Public Health 2017 Nov 20;17(Suppl 5):854 [FREE Full text] [doi: 10.1186/s12889-017-4860-0] [Medline: 29219090]
2. Leppänen MH, Nyström CD, Henriksson P, Pomeroy J, Ruiz JR, Ortega FB, et al. Physical activity intensity, sedentary behavior, body composition and physical fitness in 4-year-old children: results from the ministop trial. Int J Obes (Lond) 2016 Jul;40(7):1126-1133. [doi: 10.1038/ijo.2016.54] [Medline: 27087109]
3. Telama R, Yang X, Leskinen E, Kankaanpää A, Hirvensalo M, Tammelin T, et al. Tracking of physical activity from early childhood through youth into adulthood. Med Sci Sports Exerc 2014;46(5):955-962. [doi: 10.1249/MSS.0000000000000181] [Medline: 24121247]
4. Jones RA, Hinkley T, Okely AD, Salmon J. Tracking physical activity and sedentary behavior in childhood: a systematic review. Am J Prev Med 2013 Jun;44(6):651-658. [doi: 10.1016/j.amepre.2013.03.001] [Medline: 23683983]
5. Aubert S, Barnes JD, Abdeta C, Abi Nader P, Adeniyi AF, Aguilar-Farias N, et al. Global matrix 3.0 physical activity report card grades for children and youth: results and analysis from 49 countries. J Phys Act Health 2018 Nov 1;15(S2):S251-S273. [doi: 10.1123/jpah.2018-0472] [Medline: 30475137]
6. Delisle Nyström C, Larsson C, Alexandrou C, Ehrenblad B, Eriksson U, Friberg M, et al. Results from Sweden's 2018 report card on physical activity for children and youth. J Phys Act Health 2018 Nov 1;15(S2):S413-S414. [doi: 10.1123/jpah.2018-0519] [Medline: 30475149]
7. Chaput JP, Colley RC, Aubert S, Carson V, Janssen I, Roberts KC, et al. Proportion of preschool-aged children meeting the Canadian 24-Hour Movement Guidelines and associations with adiposity: results from the Canadian Health Measures Survey. BMC Public Health 2017 Nov 20;17(Suppl 5):829 [FREE Full text] [doi: 10.1186/s12889-017-4854-y] [Medline: 29219075]
8. De Craemer M, McGregor D, Androutsos O, Manios Y, Cardon G. Compliance with 24-h movement behaviour guidelines among Belgian pre-school children: the ToyBox-study. Int J Environ Res Public Health 2018 Oct 3;15(10):- [FREE Full text] [doi: 10.3390/ijerph15102171] [Medline: 30282898]
9. Cliff DP, McNeill J, Vella SA, Howard SJ, Santos R, Batterham M, et al. Adherence to 24-hour movement guidelines for the early years and associations with social-cognitive development among Australian preschool children. BMC Public Health 2017 Nov 20;17(Suppl 5):857 [FREE Full text] [doi: 10.1186/s12889-017-4858-7] [Medline: 29219104]
10. Berglind D, Hansson L, Tynelius P, Rasmussen F. Levels and patterns of objectively measured physical activity and sedentary time in 4-year-old Swedish children. J Phys Act Health 2017 Feb;14(2):117-122. [doi: 10.1123/jpah.2016-0250] [Medline: 27775484]
11. Temple M, Robinson JC. A systematic review of interventions to promote physical activity in the preschool setting. J Spec Pediatr Nurs 2014 Oct;19(4):274-284. [doi: 10.1111/jspn.12081] [Medline: 24888784]
12. Mehtälä MA, Sääkslahti AK, Inkinen ME, Poskiparta ME. A socio-ecological approach to physical activity interventions in childcare: a systematic review. Int J Behav Nutr Phys Act 2014 Feb 22;11:22 [FREE Full text] [doi: 10.1186/1479-5868-11-22] [Medline: 24559188]
13. Pate RR, Brown WH, Pfeiffer KA, Howie EK, Saunders RP, Addy CL, et al. An intervention to increase physical activity in children: a randomized controlled trial with 4-year-olds in preschools. Am J Prev Med 2016 Jul;51(1):12-22 [FREE Full text] [doi: 10.1016/j.amepre.2015.12.003] [Medline: 26803357]
14. Peirson L, Fitzpatrick-Lewis D, Morrison K, Ciliska D, Kenny M, Usman Ali M, et al. Prevention of overweight and obesity in children and youth: a systematic review and meta-analysis. CMAJ Open 2015 Jan 13;3(1):E23-E33 [FREE Full text] [doi: 10.9778/cmajo.20140053] [Medline: 25844367]
15. Peden ME, Okely AD, Eady MJ, Jones RA. What is the impact of professional learning on physical activity interventions among preschool children? A systematic review. Clin Obes 2018 Aug;8(4):285-299. [doi: 10.1111/cob.12253] [Medline: 29852538]
16. The Swedish National Agency for Education. URL: http://www.skolverket.se/publikationer?id=3777 [accessed 2018-10-10] [WebCite Cache ID 733nGPCu8]
17. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
18. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
19. Moore HJ, Hillier FC, Batterham AM, Ells LJ, Summerbell CD. Technology-based dietary assessment: development of the Synchronised Nutrition and Activity Program (SNAP). J Hum Nutr Diet 2014 Jan;27(Suppl 1):36-42. [doi: 10.1111/j.1365-277X.2012.01295.x] [Medline: 23848949]

20. Joiner KL, Nam S, Whittemore R. Lifestyle interventions based on the diabetes prevention program delivered via eHealth: a systematic review and meta-analysis. Prev Med 2017 Jul;100:194-207 [FREE Full text] [doi: 10.1016/j.ypmed.2017.04.033] [Medline: 28456513]
21. Ludwig K, Arthur R, Sculthorpe N, Fountain H, Buchan DS. Text messaging interventions for improvement in physical activity and sedentary behavior in youth: systematic review. JMIR Mhealth Uhealth 2018 Sep 17;6(9):e10799 [FREE Full text] [doi: 10.2196/10799] [Medline: 30224335]
22. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 7;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
23. Downing KL, Salmon J, Hinkley T, Hnatiuk JA, Hesketh KD. Feasibility and efficacy of a parent-focused, text message-delivered intervention to reduce sedentary behavior in 2- to 4-year-old children (Mini Movers): pilot randomized controlled trial. JMIR Mhealth Uhealth 2018 Feb 9;6(2):e39 [FREE Full text] [doi: 10.2196/mhealth.8573] [Medline: 29426816]
24. Militello L, Melnyk BM, Hekler EB, Small L, Jacobson D. Automated behavioral text messaging and face-to-face intervention for parents of overweight or obese preschool children: results from a pilot study. JMIR Mhealth Uhealth 2016 Mar 14;4(1):e21 [FREE Full text] [doi: 10.2196/mhealth.4398] [Medline: 26976387]
25. Nyström CD, Sandin S, Henriksson P, Henriksson H, Trolle-Lagerros Y, Larsson C, et al. Mobile-based intervention intended to stop obesity in preschool-aged children: the MINISTOP randomized controlled trial. Am J Clin Nutr 2017 Jun;105(6):1327-1335. [doi: 10.3945/ajcn.116.150995] [Medline: 28446496]
26. Rutberg S, Lindqvist AK. Active school transportation is an investment in school health. Health Behav Policy Rev 2018 Mar 1;5(2):88-97. [doi: 10.14485/HBPR.5.2.9]
27. Lindqvist AK, Rutberg S. One step forward: development of a program promoting active school transportation. JMIR Res Protoc 2018 May 8;7(5):e123 [FREE Full text] [doi: 10.2196/resprot.9505] [Medline: 29739733]
28. Hesketh KR, Lakshman R, van Sluijs EM. Barriers and facilitators to young children's physical activity and sedentary behaviour: a systematic review and synthesis of qualitative literature. Obes Rev 2017 Dec;18(9):987-1017 [FREE Full text] [doi: 10.1111/obr.12562] [Medline: 28589678]
29. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Hesketh KR, Griffin SJ, van Sluijs EM. UK Preschool-aged children's physical activity levels in childcare and at home: a cross-sectional exploration. Int J Behav Nutr Phys Act 2015 Sep 26;12:123 [FREE Full text] [doi: 10.1186/s12966-015-0286-1] [Medline: 26410252]
32. Hinkley T, Salmon J, Crawford D, Okely AD, Hesketh KD. Preschool and childcare center characteristics associated with children's physical activity during care hours: an observational study. Int J Behav Nutr Phys Act 2016 Nov 11;13(1):117 [FREE Full text] [doi: 10.1186/s12966-016-0444-0] [Medline: 27836004]
33. Hinkley T, Salmon J, Okely AD, Hesketh K, Crawford D. Correlates of preschool children's physical activity. Am J Prev Med 2012 Aug;43(2):159-167. [doi: 10.1016/j.amepre.2012.04.020] [Medline: 22813680]
34. Määttä S, Konttinen H, Haukkala A, Erkkola M, Roos E. Preschool children's context-specific sedentary behaviours and parental socioeconomic status in Finland: a cross-sectional study. BMJ Open 2017 Nov 3;7(11):e016690 [FREE Full text] [doi: 10.1136/bmjopen-2017-016690] [Medline: 29101133]
35. Berglind D, Tynelius P. Objectively measured physical activity patterns, sedentary time and parent-reported screen-time across the day in four-year-old Swedish children. BMC Public Health 2017 Aug 1;18(1):69 [FREE Full text] [doi: 10.1186/s12889-017-4600-5] [Medline: 28764730]
36. McWilliams C, Ball SC, Benjamin SE, Hales D, Vaughn A, Ward DS. Best-practice guidelines for physical activity at child care. Pediatrics 2009 Dec;124(6):1650-1659 [FREE Full text] [doi: 10.1542/peds.2009-0952] [Medline: 19917582]
37. Bower JK, Hales DP, Tate DF, Rubin DA, Benjamin SE, Ward DS. The childcare environment and children's physical activity. Am J Prev Med 2008 Jan;34(1):23-29. [doi: 10.1016/j.amepre.2007.09.022] [Medline: 18083447]
38. Trost SG, Ward DS, Senso M. Effects of child care policy and environment on physical activity. Med Sci Sports Exerc 2010 Mar;42(3):520-525. [doi: 10.1249/MSS.0b013e3181cea3ef] [Medline: 20068496]
39. The Swedish National Agency for Education. URL: https://www.skolverket.se/download/18.4fc05a3f164131a741815d2/1535537399180/Laroplan_forskolan_SKOLFS_2018_50.pdf [accessed 2018-10-10] [WebCite Cache ID 733oQquAo]
40. Niantic Inc. Pokémon Go URL: https://www.pokemongo.com/en-us/ [accessed 2019-01-09] [WebCite Cache ID 75IjHEaVH]
41. Lindqvist AK, Castelli D, Hallberg J, Rutberg S. The praise and price of Pokémon go: a qualitative study of children's and parents' experiences. JMIR Serious Games 2018 Jan 3;6(1):e1 [FREE Full text] [doi: 10.2196/games.8979] [Medline: 29298750]

42. Rothman AJ. "Is there nothing more practical than a good theory?": Why innovations and advances in health behavior change will arise if interventions are used to test and refine theory. Int J Behav Nutr Phys Act 2004 Jul 27;1(1):11 [FREE Full text] [doi: 10.1186/1479-5868-1-11] [Medline: 15279674]
43. Downing KL, Salmon J, Hinkley T, Hnatiuk JA, Hesketh KD. A mobile technology intervention to reduce sedentary behaviour in 2- to 4-year-old children (Mini Movers): study protocol for a randomised controlled trial. Trials 2017 Mar 3;18(1):97 [FREE Full text] [doi: 10.1186/s13063-017-1841-7] [Medline: 28253904]
44. Delisle C, Sandin S, Forsum E, Henriksson H, Trolle-Lagerros Y, Larsson C, et al. A web- and mobile phone-based intervention to prevent obesity in 4-year-olds (MINISTOP): a population-based randomized controlled trial. BMC Public Health 2015 Feb 7;15:95 [FREE Full text] [doi: 10.1186/s12889-015-1444-8] [Medline: 25886009]
45. Duvinage K, Ibrügger S, Kreichauf S, Wildgruber A, De Craemer M, De Decker E, ToyBox-study group. Developing the intervention material to increase physical activity levels of European preschool children: the ToyBox-study. Obes Rev 2014 Aug;15(Suppl 3):27-39. [doi: 10.1111/obr.12176] [Medline: 25047376]
46. Choo EK, Garro AC, Ranney ML, Meisel ZF, Morrow Guthrie K. Qualitative research in emergency care part I: research principles and common applications. Acad Emerg Med 2015 Sep;22(9):1096-1102 [FREE Full text] [doi: 10.1111/acem.12736] [Medline: 26284696]
47. Copeland KA, Sherman SN, Kendeigh CA, Saelens BE, Kalkwarf HJ. Flip flops, dress clothes, and no coat: clothing barriers to children's physical activity in child-care centers identified from a qualitative study. Int J Behav Nutr Phys Act 2009 Nov 6;6:74 [FREE Full text] [doi: 10.1186/1479-5868-6-74] [Medline: 19895677]
48. Pulakka A, Ashorn P, Gondwe A, Phiri N, Ashorn U. Malawian parents' perceptions of physical activity and child development: a qualitative study. Child Care Health Dev 2015 Nov;41(6):911-919. [doi: 10.1111/cch.12218] [Medline: 25403103]
49. Zahra J, Sebire SJ, Jago R. "He's probably more Mr. sport than me"--a qualitative exploration of mothers' perceptions of fathers' role in their children's physical activity. BMC Pediatr 2015 Aug 26;15:101 [FREE Full text] [doi: 10.1186/s12887-015-0421-9] [Medline: 26306617]
50. Hinkley T, Brown H, Carson V, Teychenne M. Cross sectional associations of screen time and outdoor play with social skills in preschool children. PLoS One 2018;13(4):e0193700 [FREE Full text] [doi: 10.1371/journal.pone.0193700] [Medline: 29617366]
51. Skinner AC, Ravanbakht SN, Skelton JA, Perrin EM, Armstrong SC. Prevalence of obesity and severe obesity in us children, 1999-2016. Pediatrics 2018 Mar;141(3). [doi: 10.1542/peds.2017-3459] [Medline: 29483202]

Abbreviations

MINISTOP: the mobile-based intervention intended to stop obesity in preschoolers
mHealth: mobile health
MVPA: moderate-to-vigorous physical activity
PA: physical activity

Edited by C Dias; submitted 18.10.18; peer-reviewed by A Davoudi, L Poppe, B Chaudhry, M Verloigne, M Lozano-Lozano; comments to author 16.11.18; revised version received 10.01.19; accepted 19.02.19; published 17.04.19.

Please cite as:
Ek A, Sandborg J, Delisle Nyström C, Lindqvist AK, Rutberg S, Löf M
Physical Activity and Mobile Phone Apps in the Preschool Age: Perceptions of Teachers and Parents
JMIR Mhealth Uhealth 2019;7(4):e12512
URL: http://mhealth.jmir.org/2019/4/e12512/
doi:10.2196/12512
PMID:30994465

Review

Application of Mobile Health Technologies Aimed at Salt Reduction: Systematic Review

Shahmir H Ali[1*], BA; Rong Luo[2*], MSc; Yuan Li[2,3], PhD; Xiangjun Liu[4], BSc; Chengyao Tang[5], BA; Puhong Zhang[2,3], PhD

[1]Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, MD, United States

[2]The George Institute for Global Health, Peking University Health Science Center, Beijing, China

[3]Faculty of Medicine, University of New South Wales, Sydney, Australia

[4]School of Health Humanities, Peking University, Beijing, China

[5]Public Health Graduate School of Medicine, Osaka University, Osaka, Japan

[*]these authors contributed equally

Corresponding Author:
Puhong Zhang, PhD
The George Institute for Global Health
Peking University Health Science Center
Level 18, Tower B, Horizon Tower, No 6 Zhichun Rd Haidian District
Beijing,
China
Phone: 86 10 8280 0577 ext 512
Fax: 86 10 8280 0177
Email: zpuhong@georgeinstitute.org.cn

Abstract

Background: High salt consumption has contributed to the rise of noncommunicable diseases around the world. The application of mobile health (mHealth) technologies has witnessed rapid growth in recent years. However, evidence to support mHealth interventions to confront the challenge of salt reduction has not yet been critically reviewed.

Objective: The aim of this study was to identify, characterize, and evaluate mHealth interventions aimed at salt reduction across the world.

Methods: A systematic search of studies in English or Chinese language published from January 1, 1992 to July 31, 2017 was conducted using 4 English databases (PubMed, MEDLINE, Global Health, and Cochrane) and 3 Chinese databases (Wanfang, China Science and Technology Journal, and China National Knowledge of Infrastructure). All studies directly using mobile technologies in health care with a primary or secondary objective of reducing dietary salt consumption were included.

Results: A total of 1609 articles were found using the search strategy, with 11 full articles (8 English and 3 Chinese) being included for data extraction, including 11 interventional studies. Overall, few high-quality interventions were identified. Most interventions were limited by small study population sample sizes, lack of control groups, and short follow-up times, all of which were obstacles in generating long-term scalable approaches. Most interventions employed short message service as a platform for mHealth interventions, whereas some innovative mHealth technologies were also explored. Most interventions had a primary focus of improving awareness of dietary salt consumption. The outcome variables used to measure intervention effectiveness included 24-hour urinary sodium excretion, spot urine sampling, dietary records, and indirect behavior or knowledge indicators targeting salt consumption. Although most interventions displayed positive outcome results, none of them provided reliable evidence to evaluate the effectiveness of salt reduction.

Conclusions: Salt reduction in mHealth initiatives remains relatively unexplored; however, studies that did intervene on salt-reduction show the potential of mHealth as an effective intervention method. We provide 3 recommendations for future mHealth interventions in salt reduction—(1) increased use of new, innovative, and interactive mHealth technologies; (2) development of mHealth interventions with primary prevention measures and goals of salt reduction; and (3) large-scale, rigorously designed, and object-targeted clinical trials of mHealth interventions with appropriate quantitative outcome variables, in particular 24-hour urine sodium.

(JMIR Mhealth Uhealth 2019;7(4):e13250) doi:10.2196/13250

KEYWORDS

mobile health; sodium; diet; cardiovascular diseases; systematic review

Introduction

Salt Consumption and Its Impact on Health

Noncommunicable diseases (NCDs) have become one of the leading causes of disease globally, and their impact on the global health burden has recently surpassed that of infectious diseases [1,2]. In fact, the burden of NCDs in developing countries seems to be worsening, with projections suggesting that by 2020, 7 out of 10 deaths in developing countries will be a result of NCDs [3]. Among the various dietary determinants of detrimental health outcomes, high salt intake has been noted as a key contributing factor for the prevalence of NCDs around the world. High-salt diets are linked to elevated blood pressure, a major risk factor for heart diseases and stroke, which in turn are among the leading causes of death worldwide [4-7]. In fact, salt reduction has been identified as 1 of the 5 priority interventions in response to the global NCD crisis [8]. Developing countries, in particular, have experienced a notable rise of excessive salt intake [9], with Uzbekistan, Turkmenistan, Thailand, China, and Vietnam among the countries falling within the top 20 highest consumers of salt in 2010 [10].

Impact of Salt-Reduction Initiatives on Disease Burden of Noncommunicable Diseases

The millennium development goals have emphasized an urgent need for adequate and equitable health services through the improvement of global health systems [11]. NCDs such as cardiovascular diseases (CVDs) have placed an increasing burden on the developing world, with 80% of all CVD-related deaths occurring in low- to middle-income countries [12]. Insufficiencies in health services are also reflected in the notable significance of health equity challenges in low-income countries [13]. Reducing dietary salt intakes is increasingly being foregrounded as a highly effective and implementable measure of reducing the burden of NCDs globally. Research has shown that a 15% reduction in population salt intake could prevent 8.5 million cardiovascular deaths over 10 years in 23 developing countries and result in major cost savings [14]. Following the 2011 Moscow global ministerial conference, the United Nations identified salt reduction as among its 3 key cost-effective priority actions to reduce the risk of NCDs [15]. In 2013, the World Health Organization (WHO) identified a 30% mean reduction of population salt intake as one of its 9 global targets of reducing NCDs [16].

Salt Reduction Interventions Globally

Globally, more countries are now conducting various interventions and formulating strategies to reduce population-wide salt intake to prevent and control NCDs. The main implementation strategies for salt reduction are food reformulation, consumer education, front-of-pack labeling, interventions in public institution settings (such as schools, hospitals, and the workplace), and taxation [17]. Almost all countries are multifaceted in their approach to dietary salt reduction, and consumer awareness and education activities are often used in conjunction with other salt-reduction interventions. However, different countries may need different salt-reduction strategies and interventions [17]. For example, in most developed countries, salt reduction can be achieved by reduction in the amount of salt added to food by the food industry. In contrast, in some developing countries such as China, where salt added by consumers during cooking or in sauces plays a bigger role in population salt intake, salt-reduction interventions tend to target settings such as schools, workplaces, or hospitals and focus more on public health campaigns to encourage consumers to use less salt [17]. Although many of the traditional approaches to salt reduction have been evaluated and significant reductions in salt intake observed [18], most of these interventions are resource intensive and limited in population reach. Therefore, with the rapid change of technology, mobile health (mHealth) represents a potential platform that can be used to address some of these challenges.

Rapid Growth of Mobile Technologies and Development of Mobile Health

The mobile technologies sector has been a field of great innovation and improvement in recent years. The number of mobile phone users are estimated to have surpassed 5 billion in 2017 [19], and in China alone, there are an estimated 788 million people who access internet via mobile phones (or about 98.3% of the total Chinese internet-using population) as of June 2018 [20]. Given this potential, mobile technologies have played an increasingly significant role in the health sector. mHealth has been defined as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices” [21]. There are several key functions of mHealth, including health education, data collection, electronic data systems, and clinical decision support systems [22-24].

Aims and Objectives

Despite the strong advocacy of mHealth as an innovative tool for improving both quality and access of health care [25-27], there remains a great need of continued systematic academic evaluations of its direct impact on key health outcomes [28-33]. Can mHealth be used as an innovative and effective long-term means of addressing global salt-reduction objectives? Currently there is no systematic review yet to describe the use of mHealth apps for salt-reduction purposes. The specific aims of this systematic review are to (1) characterize mHealth developments aimed at dietary salt reduction across the world, (2) evaluate the impact and effectiveness of mHealth interventions focused on salt reduction, and (3) identify the success factors and gaps in mHealth intervention development and evaluation on salt reduction that need to be addressed in the future.

Methods

Database Search

We followed the methods detailed in a peer-reviewed systematic review protocol that is registered with International Prospective Register of Systematic Reviews (PROSPERO: CRD42017075361). A systematic search of the literature in both Chinese and English published from January 1, 1992 to July 31, 2017 was performed following the preferred reporting items for systematic reviews and meta-analyses guidelines [34,35] using 4 English electronic databases—PubMed, MEDLINE, Global Health, and Cochrane; and 3 Chinese databases—Wanfang Database, Chinese Science and Technology Journal Database, and China National Knowledge of Infrastructure. We also searched for registered trials in the WHO International Clinical Trials Registry Platform and Clinicaltrial.gov, and some gray literature sources (notably a manual search of reports and government documents). The search equation is built around free-text words referring to mHealth and salt reduction. Keywords used in the English database search strategy include the following: mobile phone, cell phone, mHealth, text message, personal digital assistants, salt consumption, salt reduction, salt management, sodium intake, sodium restriction, and sodium education. Keywords used in the Chinese database search strategy include the following: “Shou Ji” (mobile phone, cell phone), “Yi Dong Jiang Kang” (mHealth), “Yi Dong Yi Liao” (mobile medical), “Yan” (salt), and “Na” (sodium). Multimedia Appendix 1 lists the detailed search strategy for each database.

Inclusion and Exclusion Criteria

Articles from all countries and with study populations of all ages, genders, and ethnicities were within the scope of this review. A study (with full-text available) was included under the condition that (1) it must include some form of measurement or incorporation of salt education, testing, or consumption data; (2) the study has an mHealth component, with mHealth being defined as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices” [21]; (3) the types of studies are randomized controlled trials (RCTs), noncontrolled trials, observational studies, qualitative interventional and quasi-experimental (QE) studies, and review studies; (4) it should be written in English or Chinese; and (5) it should be published from 1992 to 2017. A study was excluded if (1) no full-text was available, (2) the full-text was neither in English nor Chinese, (3) it did not include any significant salt-reduction-oriented component, (4) it used exclusively internet of computer or Web-based interventions, (5) it was not conducted on humans, and (6) it focused on biochemistry or microbiology.

Study Selection

Using the criteria, 2 independent reviewers conducted the abstract screening process for the English database results, and 2 independent reviewers conducted the abstract screening process for the Chinese database results. Full-text reviews were then conducted based on the results of the abstract screening process using the exclusion criteria; 2 independent reviewers reviewed the results from the English database abstract reviews and 2 independent reviewers reviewed the results from the Chinese database abstract reviews. The determinations made by the independent reviewers were then reconciled into a collective decision; if the 2 reviewers were unable to come to a consensus on the relevance of a particular study, then a third independent reviewer was asked to provide his or her input.

Data Extraction

A spreadsheet was developed to extract key data points from the filtered studies. The extraction sheet was pilot-tested on 10 randomly selected records and refined. Information extracted from each study consisted of characteristics of study participants (including age, sex, disease, and setting of the study), study design, inclusion and exclusion criteria (including type of mHealth device used), and relevant outcome measurements (salt-consumption behavior, salt intake, urinary sodium, and usability of the select mHealth device). A meta-analytic statistical comparison was unable to be performed because of the small study sample size and diversity of outcome indicators. All reviewers reached a consensus on all key definitions and interpretations of the data extraction form before the extraction process was commenced. Overall, 2 reviewers independently extracted data from the filtered English studies, and 2 reviewers independently extracted data from the filtered Chinese studies. All extractions were cross-reviewed, and any disagreements were deliberated upon following which, a consensus was reached. If they were unable to reach a consensus, a third reviewer was brought in to provide a definitive input.

Quality Assessment

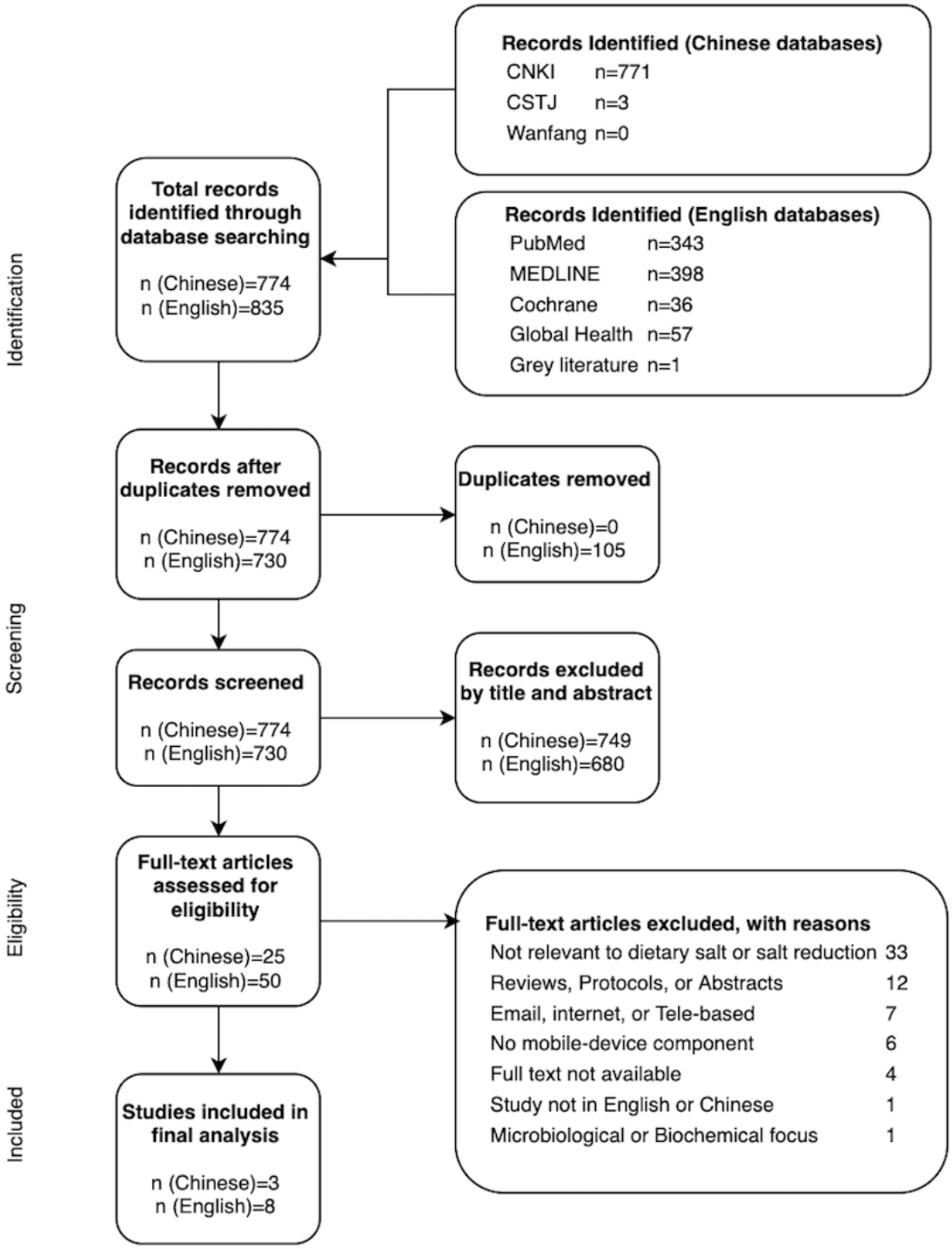
Given the overall low methodological quality of QE studies and lack of mature quality assessment tools, we only conducted quality assessment for RCTs. For RCTs, methodological quality was assessed using the Cochrane Risk of Bias Assessment Tool. Specifically, the criteria used to evaluate each study included the random sequence generation, allocation concealment, blinding of participants, personnel and outcome assessors, incomplete outcome data, selective outcome reporting, and other sources of bias. In addition, 2 reviewers independently conducted a quality assessment for each study, and any discrepancies in judgment were discussed and resolved. If discrepancies remained, a third reviewer was invited to make a final judgment.

Results

Included Studies

We retrieved 1609 articles from the Chinese and English search terms (835 through English search terms and 774 through Chinese search terms), 75 of which were collectively selected for full-text review (50 from the English search and 25 from the Chinese search). The 64 articles were excluded for the reasons specified in Figure 1.

Figure 1. Study flowchart. CNKI: China National Knowledge Infrastructure; CSTJ: China Science and Technology Journal.



Study Characteristics

The 11 included studies and their key characteristics are summarized in Table 1. Of these, 6 were RCTs [36-41] and 5 were QE studies [42-46]. Overall, 64% (n=7/11) studies had small sample sizes (had <100 participants), and 73% (8/11) had short follow-up time (<3 months in duration). Although all interventional studies directly evaluated changes in a salt-intake-related variable following an mHealth intervention, only 3 studies had an explicit primary objective of reducing dietary salt intake [36,38,46].

Table 1. Summary of key characteristics of included studies (n=11).

Study (year)	Country	Study design	Study population	Intervention duration	Primary intervention target
Eyles et al (2017) [36]	New Zealand	RCT[a]	66 cardiovascular disease patients	4 weeks	Use the app to make lower salt choices
Zhang et al (2017) [37]	China	RCT	72 patients with primary hypertension	3 months	Use the app to improve follow-up efficiency
Ipjian (2016) [38]	United States	RCT	33 adults with mobile phones	4 weeks	Use the app to change dietary, including less sodium intake
Golshahi et al (2015) [39]	Iran	RCT	180 hypertensive patients	6 months	Use SMS[b] for hypertension self-care, including vegetable intake and blood pressure
Yu et al (2013) [40]	China	RCT	385 community residents	12 months	Use SMS to improve peoples' knowledge, attitude, practice levels on noncommunicable diseases prevention and treatment
Chen et al (2008) [41]	China	RCT	762 patients with metabolic syndrome	2 years	Use SMS to inform medicine and nutrition information
Lee et al (2017) [42]	South Korea	QE[c]	33 high school students	3 months	Use the app to monitor dietary intake, including various nutrients
Radhakrishnan et al (2016) [43]	United States	QE	27 heart failure patients with minimal cognitive or physical impairment	4 weeks	Use game to improve disease knowledge, self-management, and behavior
Ahn et al (2016) [44]	South Korea	QE	26 patients with diabetes	1 month	Use program or app for nutritional management and dietary change
Huang (2013) [45]	China	QE	165 hypertension patients	3 months	Use SMS and app to improve awareness and control of hypertension
Nundy et al (2013) [46]	United States	QE	15 African American adults with Acute decompensated heart failure	30 days	Use SMS to provide self-care reminders and patient education on diet, symptom recognition, and health care navigation

[a]RCT: randomized controlled trial.

[b]SMS: short message service.

[c]QE: quasi-experimental.

A summary of the mHealth technologies used across the studies is displayed in Table 2. Short message service (SMS) was the most common form of mHealth intervention (n=6) [36,39,40,41,45,46]. Some studies incorporated a multimethod mHealth intervention approach, including combining SMS with QQ (Tencent) - an instant communication software in China similar to Facebook and Twitter [45] - and combining SMS with educational pamphlets [41]. Interactive mobile phone apps were used by 5 studies [36-38,42,44]. Of the intervention studies, most (n=8) had a primary focus of improving awareness and education on dietary salt consumption. Direct nutritional guidance by medical experts or passive dissemination of specific strategies or background information pertaining to salt-reduction was commonly employed [37,39-41,45,46]. Moreover, 1 study explored a more active form of health promotion by using a mobile phone app to allow consumers to scan packaged foods and beverages for nutritional information as well as low-salt alternatives [36]. Overall, 4 studies focused on disease management or the monitoring of diet trends [37,38,42,44], all of which involved a component of written input of daily consumption information. Moreover, 2 of the studies also involved photographic input of daily consumption information [37,44], 2 studies involved an audio component [37,42], and 1 study involved a video component [37]. For the function of mHealth interventions, 7 studies used mHealth for health education and awareness to patients [36,37,39-41,43,46], 3 studies used apps to monitor behaviors [38,42,44], and 1 study used SMS to provide self-care reminders to patients [45].

The majority of the studies were conducted in an urban area (n=9) [36-40,42,43,45,46], whereas the other 2 studies did not mention about rural or urban areas. Overall, 5 studies occurred in community settings [36,38,40,42,44], whereas 6 occurred in hospital settings [37,39,41,43,45,46]. The most common disease focus was hypertension (n=3) [37,39,45], whereas other studies centered around heart failure (n=2) [43,46], CVDs (n=1) [36], diabetes (n=1) [44], metabolic syndrome (n=1) [41], and any NCDs (n=1) [40].

Effectiveness of Mobile Health

The salt-reduction-related outcome variables employed by the studies with RCT or QE study designs are summarized in Table 3.

Table 2. Technologies used in the randomized controlled trials and quasi-experimental studies (n=11). A checkmark indicates that the technology was observed in the study, while an em dash indicates that it was not.

Study	Short message service	App-based	Others (eg, game device)	Web-based	In-person
Eyles et al [36]	✓	✓	—	—	—
Golshahi et al [39]	✓	—	—	—	—
Yu et al [40]	✓	—	—	—	—
Chen et al [41]	✓	—	—	—	✓
Huang [45]	✓	—	—	✓	—
Nundy et al [46]	✓	—	—	—	—
Zhang et al [37]	—	✓	—	—	—
Ipjian and Johnston [38]	—	✓	—	—	—
Lee et al [42]	—	✓	—	—	—
Ahn et al [43]	—	✓	—	✓	—
Radhakrishnan et al [43]	—	—	✓	—	—
Total	6	5	1	2	1

Table 3. Primary outcome and outcome variables employed by studies categorized by study design and if with successful salt reduction (n=11).

Study design	Outcome variables[a]			
	24-hour urine (mg/24 hours)	Spot urine (mg/L)	Dietary record	Behavior or knowledge indicators
Randomized controlled trial: Is salt-reduction or intake a primary outcome?				
Yes (n=2)	0	2 [36,38]	0	1 [36]
No (n=4)	0	0	2 [39,41]	2 [37,40]
Quasi-experimental study: Is salt-reduction or intake a primary outcome?				
Yes (n=0)	0	0	0	0
No (n=5)	0	0	3 [42,44,45]	3 [43,44,46]
Total	0	2	5	6

[a]A study might employ multiple outcome variables to evaluate the effectiveness on salt reduction, so that the total number of outcome variables might be larger than the number of studies in the same row.

Table 4. Summary of effectiveness results (among the 11 randomized controlled trials and quasi-experimental studies).

Outcome variable category	Studies with positive result[a], n (%)
Spot urine	2 (100) [36,38]
Dietary record	4 (80) [41,42,44,45]
Behavior or knowledge indicators	2 (33) [40,46]
Total	8 (73)

[a]A study outcome was regarded as positive if the postintervention study population recorded statistically significant decreases in salt intake or improvements in behavior or knowledge indicators implying the reduction of salt consumption.

The primary categories examined 24-hour urinary sodium excretion, spot urine sampling, dietary records, and indirect behavior or knowledge indicators targeting salt consumption. The dietary record category included any self-reported measurements of daily salt intake. The behavior or knowledge indicators category reflected study-specific measurements of changes in either behavior or knowledge variables that linked with sodium consumption tendencies. The results of the outcome variables are summarized in Table 4.

Overall, a majority of the studies (8/11, 73%) showed a positive result in the salt consumption-related outcome indicators examined.

No RCT or QE study evaluated the effectiveness of salt reduction using the 24-hour urine collection method. Dietary sodium consumption records were measured through 24-hour dietary recalls [42,44] as well as changes in participants reporting daily salt consumption above 10 g per day [41] and below 6 g per day [40,45]. Most studies displayed positive results; self-reported reductions in the frequency of days of high

salt consumption (more than 10 g/day) [41] as observed, along with increased self-reporting of salt consumption of less than 6 g per day [40,45]. Moreover, 1 intervention noted a poststudy daily salt intake decrease of approximately 0.8 g per day [42]. Salt consumption-related behavior or knowledge outcome indicators included Self-Care Heart Failure Index (SCHFI) [46]; survey questions on whether participants ate salty foods [44,39]; salt content of household packaged food purchases [36]; self-identified salt consumption patterns of either "light," "medium," or "heavy" [37]; and self-motivation to reduce salt consumption. Overall, 4 studies found no statistically significant difference in salt consumption behavior post intervention [36,37,39,44]. Among studies with positive results, 1 study showed improvements in the SCHFI index questions on maintenance of low-salt diets and reductions in salt intake [46], whereas another study found that 84% of participants agreed the mHealth game used in the intervention motivated them to restrict their salt intake [43]. Indeed, overall, given the large discrepancies in study methodology and mHealth technologies employed within this study, systematically and conclusively identifying key success factors in mHealth interventions targeting salt consumption remain difficult.

Quality Assessment

Using the Cochrane Risk of Bias Tool, the majority of the bias risk criteria for the included RCTs remained either unclear or low (Multimedia Appendix 2). The quality of 5 QE studies was not assessed.

Discussion

Principal Findings

mHealth is becoming increasingly popular around the world as the information technology, telecommunication, and health sectors continue to experience waves of innovation. With this progress, the development of mHealth interventions and its evaluation have also become increasingly prominent. In this comprehensive systematic review, we examined the application of mHealth technologies aimed at salt reduction in the published studies and registered trials. We particularly focused on interventional studies but also examined studies of other methodologies examining mHealth in this area. Most (7/11, 64%) studies indicated that mHealth interventions indeed had a positive effect on salt reduction; however, these studies also had very small sample sizes and short intervention periods, whereas none of the RCTs or QEs used the 24-hour urinary excretion measurement method (the gold standard of evaluating salt intake). The overall number of published papers (n=11) remained extremely small, 10 of which (91%) were published on or after 2013, indicating that the evaluation of mHealth technologies for the purposes of dietary salt reduction is indeed at an early stage. A vast majority of the studies examined salt reduction as a mere component of a broader investigation of other macronutrients—only 2 studies focused exclusively on salt reduction [36,38], and few studies had clear objectives and goals of salt reduction. Thus, there is need for greater and more specific focused mHealth interventional research with clear objectives and goals of salt reduction.

SMS platforms were identified as the most common form of mHealth intervention in salt reduction. However, this review also identified a number of innovative user-engagement methods of mHealth technology that hold potential for further exploration and research. For example, 1 study involved the design of a mobile-device video game that incorporated salt reduction promotion [43], whereas another study examined a mobile app that allowed users to actively scan food products and receive nutritional information as well as targeted low-salt alternatives [36]. Some studies incorporated mobile phone programs with specific regional relevance; the Chinese social media apps QQ and WeChat were used in 2 studies conducted in China [37,45]. Popular global social media platforms such as Facebook, Twitter, and YouTube, thus, also hold the potential to be used as tools of health, given the abilities of such platforms to disseminate information quickly, reach broader audiences, customize health messages for specific groups, and encourage interaction and engagement [47,48]. The intensity of user engagement across the mHealth technologies and interventions examined remained relatively minimal and was commonly limited to instructive and reminder-oriented forms of health promotion. Indeed, there is a greater need for further research on various mHealth interventions to better gauge the true potential of mHealth technologies to achieve salt-reduction goals.

In this review, most studies (8/11, 73%) focused on populations diagnosed with specific salt-related diseases. Many of the salt-reduction initiatives examined tended to be part of broad treatment plans for those already experiencing diseases such as hypertension and cardiovascular illnesses. Nonetheless, some studies examined much broader subgroups, including healthy adults [38], high school students [42], and community residents [40]. Thus, given that high salt intake is an issue identified across multiple general subpopulations [4,5], this review identifies a strong need for mHealth-based salt reduction interventions aimed at disease prevention among healthy populations as opposed to simply disease management as well as targeted interventions to better engage with a wider spectrum of subpopulations.

With respect to overall effectiveness of mHealth as an interventional tool for dietary salt reduction, given the significant heterogeneity in intervention designs and outcome indicators across the interventional studies, no decisive conclusions pertaining to effectiveness have currently been reached. A majority of the outcome measures focused on qualitative or survey-based indicators. Ultimately, the emphasis on qualitative indicators in the included studies indicates a very strong need for studies that use more quantitative and standardizable measurement indicators. The 24-hour urinary sodium excretion method is considered the gold standard in the measurement of sodium intake [49]. Single 24-hour urinary sodium readings are often subject to high levels of discrepancies between individuals in a population; thus, consecutive 24-hour urinary sodium collection readings over several days can reduce the impact of intraindividual differences. However, 24-hour urinary sodium collection carries a high participant burden of compliance; thus, many studies have to rely on alternative sodium intake measurement options—such as spot urine sampling. None of

the interventional studies in the final study sample employed 24-hour urinary sodium measurements as an outcome indicator, whereas only 2 studies utilized spot urine sampling [36,38], signaling a clear need for further mHealth research evaluating changes in salt consumption using high-quality sodium intake measurements.

It is estimated that a total of 75 countries now have a national salt reduction strategy, with the majority of programs including interventions in public institutions [17,50]. However, across the included studies, China and the United States, the 2 engines of global internet development, are clearly leading the way in mHealth-related interventional research on salt-reduction promotion. We believe that more countries, especially those countries with high dietary salt consumption, will utilize mHealth as an intervention measure to reduce salt consumption for hypertension and CVD prevention and control in the coming years. Likewise, mHealth interventions that can target consumers of salt may particularly benefit countries (such as China) in which the use of large quantities of salt in cooking plays a larger role in population salt intake than food processing and preservation by food industries (which is a more significant problem in countries such as the United States and Japan) [51].

Quality of the Evidence

Our review found that most interventional studies were limited by small sample sizes, a lack of control groups and reliable salt-intake measurements, and short follow-up times; the external and internal validities of many studies remain low. Thus, the low-quality nature of this study cannot provide substantive evidence for more scaled-up implementations. Therefore, well-designed health interventions and rigorous evaluation of salt-reduction programs will be vital for future scale-up and policy making.

Limitations

There were several limitations in our review. First, we were unable to conduct a meta-analysis to evaluate effectiveness of mHealth interventions and conduct further analysis of success factors in mHealth intervention development and evaluation on salt reduction because of the significant heterogeneity of outcome indicators and health issues reported from the very limited number of RCTs. Second, only English and Chinese language articles were included, and only 4 English and 3 Chinese databases were examined. Third, this review focused on academic studies in the literature; mHealth is growing in prominence within the commercial world, and not all of this potential health impact has been studied exclusively in academic contexts.

Conclusions

In summary, despite the growth of mHealth and its potential to achieve salt-reduction goals, evidence for a positive effect on health is still limited. Salt-reduction as a primary focus in mHealth initiatives remains relatively unexplored, and the emphasis on salt reduction is not very strong. However, among the included studies, the use of mHealth interventions has generally led to improvements in salt-reduction-related outcomes (notably changes in sodium from spot-urine sampling and dietary salt consumption records). In general, the methodological quality of the studies included in this systematic review is low. However, there are numerous opportunities for research and implementation of mHealth in salt reduction. Moving forward, we have identified 3 priority areas to improve the mHealth research agenda—(1) increased use of new, innovative, and interactive mHealth technologies (such as mobile phone apps) in the context of salt reduction; (2) development of mHealth interventions with primary prevention measures and goals of salt reduction; and (3) large-scale, rigorously designed, and object-targeted clinical trials of mHealth interventions aimed at salt reduction with appropriate quantitative outcome variables, in particular, 24-hour urinary sodium.

Acknowledgments

This research was supported by the UK National Institute for Health Research (NIHR) Global Health Research Unit Action on Salt China Program under award number 16/136/77. The content is solely the responsibility of the authors and does not necessarily represent the official views of the sponsors. The sponsors had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author (PZ) had the final responsibility for decision of publication.

Authors' Contributions

RL and YL designed the study, and RL and SHA designed the search strategies. SHA, RL, XL, and CT conducted the literature search, screening, and data extraction. SHA and RL conducted data collection, data analysis, and contributed equally to data interpretation and writing. PZ and YL contributed to data interpretation and manuscript revision.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PNG File, 698KB - mhealth_v7i4e13250_app1.png]

Multimedia Appendix 2

Quality assessment of included randomized controlled trials (using Cochrane Risk of Bias Tool).

[PNG File, 279KB - mhealth_v7i4e13250_app2.png]

References

1. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2095-2128. [doi: 10.1016/S0140-6736(12)61728-0] [Medline: 23245604]
2. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2197-2223. [doi: 10.1016/S0140-6736(12)61689-4] [Medline: 23245608]
3. Boutayeb A. The double burden of communicable and non-communicable diseases in developing countries. Trans R Soc Trop Med Hyg 2006 Mar;100(3):191-199. [doi: 10.1016/j.trstmh.2005.07.021] [Medline: 16274715]
4. Mozaffarian D, Fahimi S, Singh GM, Micha R, Khatibzadeh S, Engell RE, Global Burden of Diseases Nutrition and Chronic Diseases Expert Group. Global sodium consumption and death from cardiovascular causes. N Engl J Med 2014 Aug 14;371(7):624-634. [doi: 10.1056/NEJMoa1304127] [Medline: 25119608]
5. Lava SA, Bianchetti MG, Simonetti GD. Salt intake in children and its consequences on blood pressure. Pediatr Nephrol 2015 Sep;30(9):1389-1396. [doi: 10.1007/s00467-014-2931-3] [Medline: 25127918]
6. Strazzullo P, D'Elia L, Kandala N, Cappuccio FP. Salt intake, stroke, and cardiovascular disease: meta-analysis of prospective studies. Br Med J 2009 Nov 24;339:b4567 [FREE Full text] [doi: 10.1136/bmj.b4567] [Medline: 19934192]
7. Aburto NJ, Ziolkovska A, Hooper L, Elliott P, Cappuccio FP, Meerpohl JJ. Effect of lower sodium intake on health: systematic review and meta-analyses. Br Med J 2013;346:f1326 [FREE Full text] [doi: 10.1136/bmj.f1326] [Medline: 23558163]
8. Beaglehole R, Bonita R, Horton R, Adams C, Alleyne G, Asaria P, Lancet NCD Action Group, NCD Alliance. Priority actions for the non-communicable disease crisis. Lancet 2011 Apr 23;377(9775):1438-1447. [doi: 10.1016/S0140-6736(11)60393-0] [Medline: 21474174]
9. A dietary intervention trial for nutritional management of cardiovascular risk factors. Nutr Rev 1997 Feb;55(2):54-56. [doi: 10.1111/j.1753-4887.1997.tb01596.x] [Medline: 9155219]
10. Powles J, Fahimi S, Micha R, Khatibzadeh S, Shi P, Ezzati M, Global Burden of Diseases Nutrition and Chronic Diseases Expert Group (NutriCoDE). Global, regional and national sodium intakes in 1990 and 2010: a systematic analysis of 24 h urinary sodium excretion and dietary surveys worldwide. BMJ Open 2013;3(12):e003733 [FREE Full text] [doi: 10.1136/bmjopen-2013-003733] [Medline: 24366578]
11. Willis-Shattuck M, Bidwell P, Thomas S, Wyness L, Blaauw D, Ditlopo P. Motivation and retention of health workers in developing countries: a systematic review. BMC Health Serv Res 2008 Dec 4;8:247 [FREE Full text] [doi: 10.1186/1472-6963-8-247] [Medline: 19055827]
12. Bovet P, Paccaud F. Cardiovascular disease and the changing face of global public health: a focus on low and middle income countries. Public Health Rev 2011 Dec 11;33(2):397-415. [doi: 10.1007/BF03391643]
13. Orach CG. Health equity: challenges in low income countries. Afr Health Sci 2009 Oct;9(Suppl 2):S49-S51 [FREE Full text] [Medline: 20589106]
14. Asaria P, Chisholm D, Mathers C, Ezzati M, Beaglehole R. Chronic disease prevention: health effects and financial costs of strategies to reduce salt intake and control tobacco use. Lancet 2007 Dec 15;370(9604):2044-2053. [doi: 10.1016/S0140-6736(07)61698-5] [Medline: 18063027]
15. World Health Organization. 2011. First global ministerial conference on healthy lifestyles and NCDs control URL: http://www.who.int/nmh/events/moscow_ncds_2011/en/ [accessed 2017-11-12] [WebCite Cache ID 74giqkVew]
16. World Health Organization. 2013. NCD Global Monitoring Framework URL: http://www.who.int/nmh/global_monitoring_framework/en/ [accessed 2017-11-13] [WebCite Cache ID 74gj7LpAg]
17. Trieu K, Neal B, Hawkes C, Dunford E, Campbell N, Rodriguez-Fernandez R, et al. Salt reduction initiatives around the world–a systematic review of progress towards the global target. PLoS One 2015;10(7):e0130247 [FREE Full text] [doi: 10.1371/journal.pone.0130247] [Medline: 26201031]
18. Christoforou A, Trieu K, Land M, Bolam B, Webster J. State-level and community-level salt reduction initiatives: a systematic review of global programmes and their impact. J Epidemiol Community Health 2016 Dec;70(11):1140-1150. [doi: 10.1136/jech-2015-206997] [Medline: 27222501]
19. Global System for Mobile Communications. 2017. The Mobile Economy 2017 URL: https://www.gsmaintelligence.com/research/?file=9e927fd6896724e7b26f33f61db5b9d5&download [accessed 2017-12-10] [WebCite Cache ID 74gjRkqCD]
20. China Internet Network Information Center. 2018. Statistical Report on the Development of China's Internet URL: https://cnnic.com.cn/IDR/ReportDownloads/201807/P020180711391069195909.pdf [accessed 2017-12-11] [WebCite Cache ID 74gjb1C8e]

21. World Health Organization. 2011. mHealth New horizons for health through mobile technologies URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2017-11-10] [WebCite Cache ID 74gjp32fz]
22. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171 [FREE Full text] [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]
23. Blaya JA, Fraser HS, Holt B. E-health technologies show promise in developing countries. Health Aff (Millwood) 2010 Feb;29(2):244-251 [FREE Full text] [doi: 10.1377/hlthaff.2009.0894] [Medline: 20348068]
24. Lindquist AM, Johansson PE, Petersson GI, Saveman B, Nilsson GC. The use of the Personal Digital Assistant (PDA) among personnel and students in health care: a review. J Med Internet Res 2008;10(4):e31 [FREE Full text] [doi: 10.2196/jmir.1038] [Medline: 18957381]
25. Kaplan WA. Can the ubiquitous power of mobile phones be used to improve health outcomes in developing countries? Global Health 2006;2:9 [FREE Full text] [doi: 10.1186/1744-8603-2-9] [Medline: 16719925]
26. Lester R, Karanja S. Mobile phones: exceptional tools for HIV/AIDS, health, and crisis management. Lancet Infect Dis 2008 Dec;8(12):738-739. [doi: 10.1016/S1473-3099(08)70265-2] [Medline: 19022188]
27. Vodafone. 2005. Africa: The impact of mobile phones URL: https://www.vodafone.com/content/dam/vodafone/about/public_policy/policy_papers/public_policy_series_2.pdf [accessed 2017-12-09] [WebCite Cache ID 74gkNFFpq]
28. Catwell L, Sheikh A. Evaluating eHealth interventions: the need for continuous systemic evaluation. PLoS Med 2009 Aug;6(8):e1000126 [FREE Full text] [doi: 10.1371/journal.pmed.1000126] [Medline: 19688038]
29. Greenhalgh T, Russell J. Why do evaluations of eHealth programs fail? An alternative set of guiding principles. PLoS Med 2010;7(11):e1000360 [FREE Full text] [doi: 10.1371/journal.pmed.1000360] [Medline: 21072245]
30. 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]
31. Kahn JG, Yang JS, Kahn JS. 'Mobile' health needs and opportunities in developing countries. Health Aff (Millwood) 2010 Feb;29(2):252-258 [FREE Full text] [doi: 10.1377/hlthaff.2009.0965] [Medline: 20348069]
32. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
33. Pop-Eleches C, Thirumurthy H, Habyarimana JP, Zivin JG, Goldstein MP, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. AIDS 2011 Mar 27;25(6):825-834 [FREE Full text] [doi: 10.1097/QAD.0b013e32834380c1] [Medline: 21252632]
34. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
35. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
36. Eyles H, McLean R, Neal B, Jiang Y, Doughty RN, McLean R, et al. A salt-reduction smartphone app supports lower-salt food purchases for people with cardiovascular disease: findings from the SaltSwitch randomised controlled trial. Eur J Prev Cardiol 2017 Dec;24(13):1435-1444. [doi: 10.1177/2047487317715713] [Medline: 28631933]
37. Zhang J, Ma Q, Li R, Li Y. [Application of mobile phone WeChat software in the follow-up of patients with essential hypertension]. West China Med J 2017;32(1):42-45 http://kns.cnki.net/kcms/detail/detail.aspx?filename=DDZT201801021&dbcode=CJFQ&dbname=CJFDTEMP&v=.
38. Ipjian ML, Johnston CS. Smartphone technology facilitates dietary change in healthy adults. Nutrition 2017 Jan;33:343-347. [doi: 10.1016/j.nut.2016.08.003] [Medline: 27742102]
39. Golshahi J, Ahmadzadeh H, Sadeghi M, Mohammadifard N, Pourmoghaddas A. Effect of self-care education on lifestyle modification, medication adherence and blood pressure in hypertensive adults: randomized controlled clinical trial. Adv Biomed Res 2015;4:204 [FREE Full text] [doi: 10.4103/2277-9175.166140] [Medline: 26601092]
40. Yu J, Wang X, Zhou L, Gu C. [The effect of chronic disease health education enhanced by SMS on the knowledge and behavior of community residents]. Chin J Prev Contr Chron Dis 2013;21(2):165-168 http://wuxizazhi.cnki.net/Search/ZMXB201302015.html.
41. Chen Y, Lin M, Pu X, Wu C, Liu Y. [Management of patients with metabolic syndrome]. Chinese J Practic Intern Med 2008;28(2):29-32.
42. Lee J, Song S, Ahn JS, Kim Y, Lee JE. Use of a mobile application for self-monitoring dietary intake: feasibility test and an intervention study. Nutrients 2017 Jul 13;9(7) [FREE Full text] [doi: 10.3390/nu9070748] [Medline: 28703758]
43. Radhakrishnan K, Toprac P, O'Hair M, Bias R, Kim M, Bradley P, et al. Interactive digital e-health game for heart failure self-management: a feasibility study. Games Health J 2016 Dec;5(6):366-374 [FREE Full text] [doi: 10.1089/g4h.2016.0038] [Medline: 27976955]

44. Ahn Y, Bae J, Kim H. The development of a mobile u-Health program and evaluation for self-diet management for diabetic patients. Nutr Res Pract 2016 Jun;10(3):342-351 [FREE Full text] [doi: 10.4162/nrp.2016.10.3.342] [Medline: 27247732]
45. Huang M. [Investigation on Disease Cognitive Behavior and Nursing Intervention in Patients with Hypertension in Special Needs]. West China Med J 2013;28(2):289-291.
46. Nundy S, Razi RR, Dick JJ, Smith B, Mayo A, O'Connor A, et al. A text messaging intervention to improve heart failure self-management after hospital discharge in a largely African-American population: before-after study. J Med Internet Res 2013;15(3):e53 [FREE Full text] [doi: 10.2196/jmir.2317] [Medline: 23478028]
47. Centers for Disease Control and Prevention. 2011. The Health Communicator's Social Media Toolkit URL: https://www.cdc.gov/healthcommunication/ToolsTemplates/SocialMediaToolkit_BM.pdf [accessed 2017-11-11] [WebCite Cache ID 74gm99XYr]
48. Hesse BW. Enhancing consumer involvement in health care. In: Parker JC, Thorson E, editors. Health Communication in the New Media Landscape. New York: Springer Publishing; 2009:119-150.
49. World Health Organization. Reducing salt intake in populations: Report of a WHO forum and technical meeting URL: https://www.who.int/dietphysicalactivity/SaltReport-VC-April07.pdf [accessed 2018-12-15] [WebCite Cache ID 74gmIobC8]
50. Webster JL, Dunford EK, Hawkes C, Neal BC. Salt reduction initiatives around the world. J Hypertens 2011 Jun;29(6):1043-1050. [doi: 10.1097/HJH.0b013e328345ed83] [Medline: 21546876]
51. Anderson CA, Appel LJ, Okuda N, Brown IJ, Chan Q, Zhao L, et al. Dietary sources of sodium in China, Japan, the United Kingdom, and the United States, women and men aged 40 to 59 years: the INTERMAP study. J Am Diet Assoc 2010 May;110(5):736-745 [FREE Full text] [doi: 10.1016/j.jada.2010.02.007] [Medline: 20430135]

Abbreviations

CVDs: cardiovascular diseases
mHealth: mobile health
NCDs: noncommunicable diseases
QE: quasi-experimental
RCT: randomized controlled trial
SCHFI: Self-Care Heart Failure Index
SMS: short message service
WHO: World Health Organization

Edited by G Eysenbach; submitted 28.12.18; peer-reviewed by W Na, C Grimes, C Gonçalves; comments to author 13.02.19; revised version received 27.02.19; accepted 03.03.19; published 17.04.19.

Please cite as:
Ali SH, Luo R, Li Y, Liu X, Tang C, Zhang P
Application of Mobile Health Technologies Aimed at Salt Reduction: Systematic Review
JMIR Mhealth Uhealth 2019;7(4):e13250
URL: http://mhealth.jmir.org/2019/4/e13250/
doi:10.2196/13250
PMID:30994467

Original Paper

Effectiveness of Low Glycemic Index Diet Consultations Through a Diet Glycemic Assessment App Tool on Maternal and Neonatal Insulin Resistance: A Randomized Controlled Trial

Yi Zhang[1*], MD, MSc; Liping Wang[2*], MD, MSc; Wenhong Yang[3*], MD; Dayan Niu[1], MSc; Chunying Li[3], MD; Liling Wang[2], MD; Ping Gu[3], MD; Yingqian Xia[2], MD; Ying Shen[2], MD; Juhua Yan[3], MD; Qian Zhao[1], MSc; Kai Mu[1], MSc; Weili Yan[1], MD, PhD

[1]Children's Hospital of Fudan University, Shanghai, China

[2]International Peace Maternity & Child Health Hospital of China Welfare Institute, Shanghai, China

[3]Kunshan Maternity and Child Care Center, Kunshan, China

[*]these authors contributed equally

Corresponding Author:
Weili Yan, MD, PhD
Children's Hospital of Fudan University
399 Wanyuan Road
Shanghai, 201102
China
Phone: 86 21064931215
Fax: 86 2164931215
Email: yanwl@fudan.edu.cn

Abstract

Background: Low glycemic index (LGI) diet has shown to be effective in reducing maternal and neonatal complications in high-risk pregnancies.

Objective: This trial aimed to examine the effectiveness of individualized LGI diet consultations based on the accurate diet glycemic load (GL) assessment tool on maternal and neonatal insulin resistance levels and diet behavior changes in overweight and obese pregnant women.

Methods: Overweight and obese pregnant women were recruited before 16 weeks of gestation and randomized to the LGI diet arm or the control arm. All participants received standard dietary education according to the Chinese Dietary Guide for Pregnant Women. In the intervention arm, additional individualized dietary GL assessments were performed using an app and instructions of lowering diet glycemic index (GI) to achieve LGI diet were provided by a clinical dietitian at early, middle, and late gestation. Primary outcomes were serum insulin at late gestation, incidence of gestational diabetes mellitus (GDM) for mothers, and cord blood C-peptide level of neonates.

Results: In total, 400 subjects were randomized and received different interventions. There were no significant differences in maternal serum insulin levels (13.2 [9.3−13.2] uU/mL vs 12.4 [10.5−12.4] uU/mL), incidence of GDM (45 [22.5%] vs 43 [21.5%]), or cord blood C-peptide levels (mean 0.9ng/mL [SD 0.7] vs mean 0.8ng/mL [SD 0.6]) in the intervention group compared with the controls. The diet GI at late gestation was similar (mean 63.2 [SD 10.4] vs mean 64.3 [SD 10.4]), whereas greater diet fiber intake was observed in the intervention group (mean 11.6 grams [SD 8.0] vs mean 9.0 grams [SD 5.6]; *P*=.006). Adherence measurements did not significantly differ between 2 groups.

Conclusions: Individualized LGI diet consultations for overweight and obese pregnant women failed to make a significant difference in maternal or neonatal insulin resistance compared with the standard gestational diet consultation.

Trial Registration: ClinicalTrials.gov NCT01628835; http://clinicaltrials.gov/ct2/show/NCT01628835 (Archived by WebCite at http://www.webcitation.org/77LHgWP0k)

(JMIR Mhealth Uhealth 2019;7(4):e12081) doi:10.2196/12081

KEYWORDS

glycemic index; overweight; pregnancy; insulin resistance; randomized controlled trial

Introduction

Background

The prevalence of overweight and obesity is increasing globally [1,2]. Data from China Chronic Disease and Risk Factor Surveillance survey indicate that 32.2% women (the majority was at childbearing age) were overweight and obese [3]. Overweight and obesity during pregnancy is associated with an increased risk of a number of adverse consequences for both mother and baby [4-6]. Prepregnancy body mass index (BMI) and weight gain during gestation are 2 of the most important risk factors for gestational diabetes mellitus (GDM) [7-9]. However, attempts at reducing complications during pregnancy among overweight or obese pregnant women have generally been unsuccessful. An alternative strategy is to take a low glycemic index (LGI) diet. The concept of glycemic index (GI) was developed by Jenkins et al in 1981 as a method of ranking the postprandial glycemic response to equivalent portions of carbohydrate in different foods [10], GI ≤55 was considered as low GI [11]. LGI foods produce lower postprandial increases in blood glucose and reduce diurnal postprandial glucose and insulin responses compared with high GI foods [11,12]. Studies involving women with GDM have shown that an LGI diet reduces postprandial glucose values and the need of insulin [13,14]. However, trials in overweight pregnant women have mostly small sample size and tend to focus on weight change of mothers and babies [15]. Insulin resistance has been accepted as a common underlying basis and indicator of cardiovascular risk. Evidence of effectiveness of the LGI diet during gestation on insulin resistance is needed.

Objective

The current trial aimed to test the hypothesis that individualized LGI dietary consultation started from the first antenatal visit would reduce the level of insulin resistance on overweight pregnant women and their babies compared with standard diet consultation according to national nutrition recommendations for pregnancy.

Methods

This study was a randomized, single-blinded controlled intervention trial, with a 1:1 allocation ratio, and approved by the Ethics Committee of Children's Hospital of Fudan University (approval NO.071-2012). Written informed consent was obtained from every subject before baseline data collection. The trial was registered at the ClinicalTrials.gov registry (NCT01628835).

Subjects

Overweight or obese pregnant women were assessed for eligibility and recruited from primary antenatal care settings of Kunshan Maternity and Child Care Center (Kunshan city, Jiangsu province, China) and the International Peace Maternity and Child Health Hospital of China Welfare Institute (Shanghai, China) from June 2012 to October 2015. Women were eligible if they met all of the following criteria: first antenatal visit ≤16 weeks of gestation (changed from ≤14 weeks in registration), aged 18 to 45 years, with BMI ≥24 kg/m^2, and would take routine prenatal examinations. Exclusion criteria were artificial impregnation, a history of hypertension, diabetes, and coronary heart disease, or subjects with mental disorder or special dietary needs (eg, vegetarianism). Subjects were recruited from the early pregnancy clinic in the hospitals where health care records of pregnant women were established. One obstetrician was trained for each center and assigned to judge the eligibility, complete the informed consent process, and implement random allocation according to the randomization plan. Subjects were referred to either one of the 2 dietitians to receive different diet consultations but were blinded to the content of the intervention.

Randomization and Dietary Intervention

The randomization sequence was generated, and block randomization (block size=4) was performed in the center by a statistician, using Microsoft Office Excel. A 1:1 randomized allocation plan for 200 subjects was provided to each center followed by a training of standard manipulation to the obstetrician who would be in charge of the allocation.

All the participants in the 2 arms received a standard nutrition and physical activity consultation according to the national recommendations of the Chinese Nutrition Society [16], as well as advice to keep gestational weight gain (GWG) according to the 2009 Institute of Medicine guidelines [17]. The dietary consultation consisted of individualized diet assessment, followed by diet planning. Participants were first asked to recall food consumption in 24 hours of the nearest working day for diet assessment, on the basis of which daily intake of conventional nutrition such as total energy, protein, fat, and carbohydrate intake was provided. Then, the dietitian worked with the participant to make an individualized diet plan to the meet the standard goals. The difference of consultation in the experimental group was the additional diet GI and glycemic load (GL) calculations in the diet assessment, and a diet plan was made to achieve an LGI goal with consideration of individual food preference. A mobile phone app DietGI (Children's Hospital of Fudan University, Multimedia Appendix 1) was equipped with the function of selecting food types and amount of intake for every meal; the GI and GL were output for a single meal or 3 meals a day. In addition, by using the tool, the dietitian demonstrated how to adapt foods for 3 meals to achieve an LGI diet, including replacing high GI foods with LGI foods that they preferred with the preferred portion size. The knowledge of combining foods in meals and cooking techniques (eg, lessen cooking time) to achieve lower GI was also provided. We developed a customized excel worksheet as the tool for quantitative calculation of diet GI and GL, which was designed with algorithms following the international rules and the published GI information of Chinese [8] and international foods [18-20]. The 24-hour diet records and nutrition assessments at 3 visits were saved in the worksheet for every subject. To blind participants to dietary assignment

and avoid contamination, dietary consultation was arranged on a different day and by separate dietitians.

In total, 3 diet consultation interviews were incorporated in the routine antenatal care to avoid loss of follow-up at first antenatal visit, middle gestation visit (24th to 28th week) with the routine 75 g oral glucose-tolerance test (OGTT), and late gestation visit (34th to 36th week) with routine liver and kidney function test. Project nurses contacted and made an appointment with participants before the due visit to increase attendance or answer any dietary queries. Between visits participants were followed up by a telephone interview at least once a month to prompt compliance.

Data Collection and Outcome Measurements

At the first antenatal visit, demographic characteristics and clinical and anthropometric measures were collected after subjects signed the informed consent. BMI was calculated as weight (kg)/(height [m])^2 and categorized according to Chinese categories (underweight <18.5 kg/m^2, normal weight 18.5 to 23.9 kg/m^2, overweight 24 to 27.9 kg/m^2, and obese ≥28 kg/m^2) [21]. The gestational age was estimated from self-reported last menstrual period and corrected by the first routine ultrasound examinations around the 14th week. Maternal GWG was defined as the weight gain from the first antenatal visit to delivery. Routine examinations of plasma glucose, serum insulin, glycosylated hemoglobin (HbA_{1c}), and blood pressure (BP) were conducted at the prenatal care setting of each center. The homeostasis model assessment of insulin resistance (HOMA-IR) was calculated as follows [22]: (fasting plasma glucose [mmol/L])×(fasting insulin [mIU/L])/22.5. GDM was defined according to international standards [23,24] based on the routine 75 g OGTT screening between 24 and 28 weeks of gestation. The routine lab examination results were extracted from the hospital information system. Cord blood was collected at birth, and serum was separated and stored in the freezer (−20 degree) for no more than 1 night and transferred to a −80-degree freezer according to standard protocols. The stored cord blood serum from 2 centers was transported to the central lab and were examined for C-peptide levels instead of insulin levels as the index of neonatal beta cell function [25,26].

The primary outcomes included incidence of GDM, maternal insulin levels before delivery, and cord blood C-peptide levels as indicators of maternal and neonatal insulin resistance levels. Maternal secondary outcome measures included GWG, incidence of gestational hypertension, and a cesarean. Key secondary outcomes for the fetus or neonate included birth weight, preterm birth before 37 weeks of gestation, and incidence of macrosomia. Gestational hypertension was defined as a systolic BP of 140 mmHg or more or a diastolic BP of 90 mmHg or more on at least 2 occasions at least 4 hours apart in a patient who was normotensive before 20 weeks of gestation. Low birth weight and macrosomia were considered as birth weight <2500 g or birth weight ≥4000 g, respectively.

Power

The sample size was calculated based on one of the primary aims: GDM. Considering that there was no analogous research at the time, relative risk was assumed as 3.57 in overweight pregnant women with a baseline risk of 3.6% according to the effect of exercises intervention compared with none [27]. In total, 400 overweight pregnant women in a 1:1 ratio were needed to achieve 90% power to detect the effect size at a 5% significance level, allowing for an expected withdrawal of 10%. For continuous outcomes, a mean difference of 0.3 SD will be powered (power ≥0.85) by this sample size at an alpha of .05.

Statistical Analysis

Continuous data are reported as mean (SD) or median (interquartile range), and categorical data are reported as percentages. Statistical analyses of primary outcomes were in accordance to the intention-to-treat strategy. Comparisons between the 2 arms were carried out by using independent samples of *t* tests for continuous variables and chi-square tests for categorical variables. Stata version 15.0 was used for all statistical analyses. Results were considered significant when the *P* value was <.05. Spearman correlation analyses were performed to test the relationship between change values of selective outcomes.

Results

Baseline Characteristics of Study Subjects

The general characteristics between the intervention group and the control group did not significantly differ (Table 1). The flow of participants is shown in Figure 1. In total, 400 women who signed the informed consent forms were randomly assigned to 2 arms with 200 subjects for each; all received the first diet consultation interview. About 10% of the subjects missed the second diet consultation interview and almost half missed the third visit. Finally, 183 in the control group and 186 subjects in the intervention group remained at the end point with complete birth data and successful biosample collection. The rate of loss to follow-up in the control group was similar to that in the intervention group (8.5% vs 7.0%, respectively; *P*=.58). General characteristics of participants at baseline are presented in Table 1, showing no significant between-group differences.

Table 1. Baseline characteristics of the study participants (N=200).

Characteristics	Control group	Intervention group	*P* value
Age (years), mean (SD)	28.0 (3.7)	28.1 (3.6)	.86
Gestational week (week), mean (SD)	12.4 (2.2)	12.2 (2.2)	.49
Height (cm), mean (SD)	162.4 (5.7)	161.7 (6.1)	.26
Weight (kg), mean (SD)	73.9 (9.5)	74.2 (9.0)	.69
Body mass index (kg/m^2), mean (SD)	28.0 (3.0)	28.4 (3.0)	.19
Systolic blood pressure (mmHg), mean (SD)	120.7 (12.4)	121.1 (11.2)	.70
Diastolic blood pressure (mmHg), mean (SD)	74.6 (9.9)	74.6 (10.1)	.98
Fasting plasma glucose (mmol/L), mean (SD)	4.7 (0.6)	4.7 (0.4)	.56
Insulin (uU/mL), median (IQR[a])	9.5 (6.8-13.5)	10.0 (6.6-13.9)	.57
HOMA-IR[b], median (IQR)	2.0 (1.3-2.7)	2.0 (1.4-3.0)	.59
HbA_{1c}[c] (%), mean (SD)	5.3 (0.8)	5.2 (0.4)	.44

[a]IQR: interquartile range.

[b]HOMA-IR: homeostasis model assessment of insulin resistance.

[c]HbA_{1c}: glycosylated hemoglobin.

Figure 1. CONSORT flow diagram.

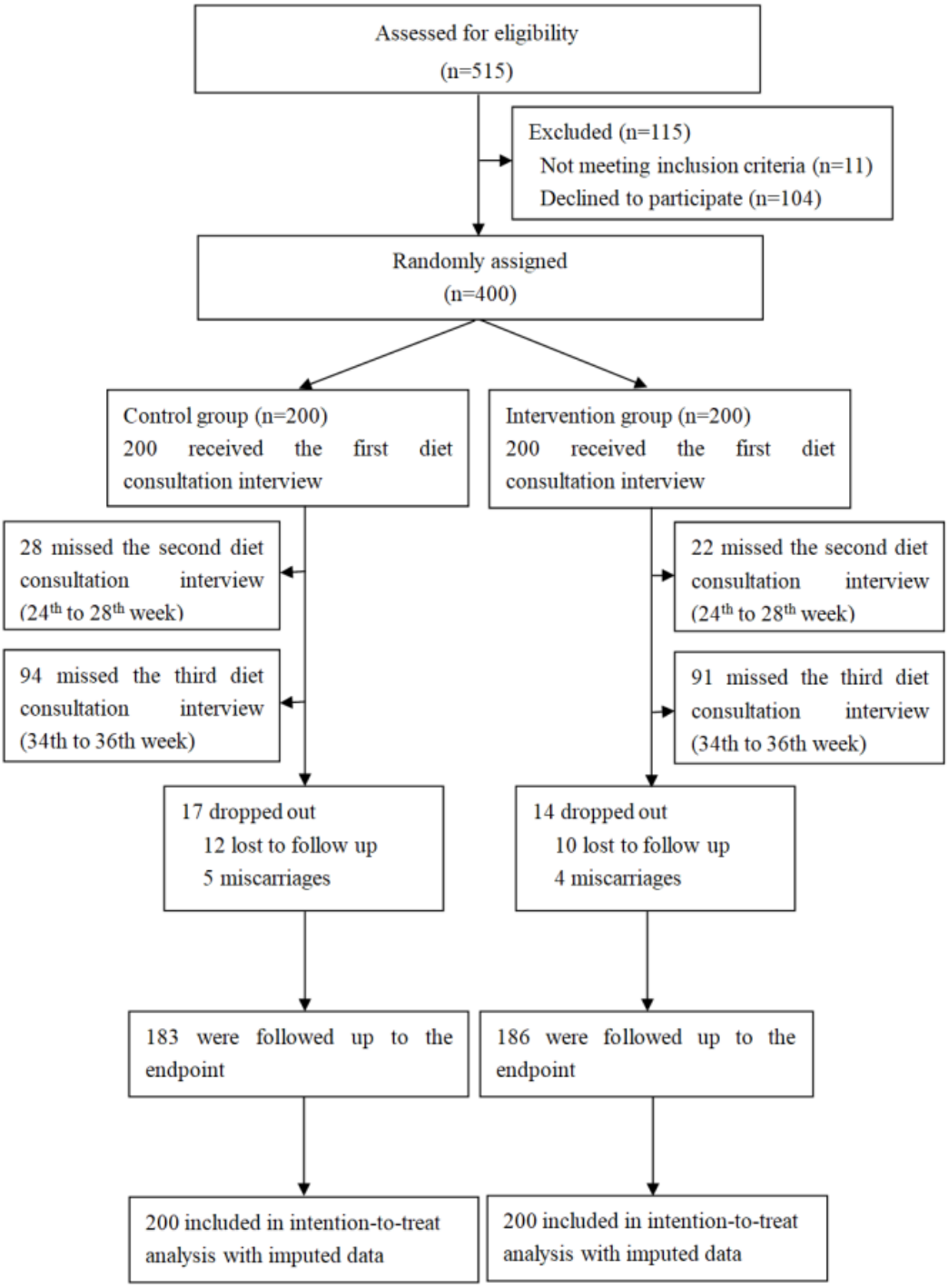


Table 2. Maternal and fetal outcomes (N=200).

Outcomes	Control group	Low glycemic index diet group	*P* value
Maternal outcomes at delivery			
Insulin (uU/mL), median (interquartile range)	12.8 (10.5-12.8)	12.8 (9.3-13.2)	.18
Insulin[a,b] (uU/mL), mean (SD)	1.5 (7.8)	1.9 (6.6)	.56
Gestational diabetes mellitus, n (%)	43 (21.5)	45 (22.5)	.33[b]
Gestational hypertension, n (%)	32 (16.2)	45 (22.8)	.11[b]
Weight (kg), mean (SD)	85.3 (10.3)	84.0 (10.6)	.23[b]
Gestational weight gain (kg), mean (SD)	11.2 (6.3)	9.6 (7.4)	.02
Gestational age (week), mean (SD)	39.8 (1.7)	39.7 (1.2)	.33
Caesarean, mean (SD)	107 (58.5)	124 (67.0)	.09
HbA_{1c} (%), mean (SD)	5.4 (0.4)	5.4 (0.3)	.79
HbA_{1c}[a,b,c] (%), mean (SD)	0.10 (0.75)	0.2 (0.4)	.35
Neonatal outcomes			
Cord blood C-peptide (ng/mL), mean (SD)	0.85 (0.61)	0.86 (0.67)	.84
Birth weight (g), mean (SD)	3452.5 (527.1)	3513.86 (522.4)	.26
Macrosomia, n (%)	21 (19.0)	31 (23.0)	.15
Preterm, n (%)	23 (11.7)	18 (8.7)	.11

[a]Median (interquartile range), *P* values based on log-transformed values.

[b]Levels in third trimester minus levels at baseline.

[c]HbA_{1c}: glycosylated hemoglobin.

Outcomes

As shown in Table 2, participants in the intervention group gained less body weight than those in the control group (9.6 [SD 7.4] vs 11.2 [SD 6.3], respectively; *P*=.03). We did not observe significant differences in maternal insulin levels at late gestation, change from baseline, and cord blood C-peptide, which were examined as markers of insulin resistance.

Diet Assessments

The dietary intakes of nutrition at 3 visits and changes across 3 visits during gestation are shown in Table 3. At the baseline, the second visit, and the last visit, no significant differences in total intake of energy, protein, fat, carbohydrate, diet GL, or GI were observed, whereas at the third visit, greater fiber intakes were observed in the LGI diet group compared with the control group (11.6 (SD 8.0) g vs 8.9 (SD 5.6) g, respectively; *P*=.006). The overall increment of diet fiber intakes in the LGI diet group significantly differs from the overall decrease in the control group across the intervention period (*P*=.006). Although we did not find significantly different responses of diet GI between the 2 intervention groups, we found that steady reductions of diet GL in both groups (15.2 (SD 76.8); *P*=.004), which includes both GI and amount of foods in calculation, are in line with the observed reduction of diet carbohydrate intake (−25.0 [SD 107.1]; *P*=.007). The posthoc analysis revealed that change of diet GI (levels at late gestation minus levels at baseline) was weakly correlated with changes of maternal insulin levels (r=0.19; *P*=.02), the significance remained after adjustment of maternal age and GWG.

The situation of adherence to interventions and missing data were similar between the 2 groups (Table 4), indicating a minor chance of confounding bias.

Table 3. Dietary intake of glycemic load (GL), glycemic index (GI), and nutrients (per day) at 3 visits and changes.

Diet assessments	Control group[a], mean (SD)	Low glycemic index (LGI) diet group[b], mean (SD)	*P* value
GL			
Baseline	130.6 (59.6)	132.4 (53.5)	.75
Second trimester	124.1 (55.4)	125.1 (49.4)	.85
Third trimester	111.6 (59.4)	112.4 (56.0)	.91
Change[c]	−14.4 (79.1)	−16.0 (74.8)	.88
GI			
Baseline	64.1 (9.7)	64.5 (8.7)	.69
Second trimester	62.8 (9.8)	64.8 (9.4)	.05
Third trimester	64.3 (10.4)	63.2 (10.4)	.42
Change	0.25 (13.7)	−1.1 (12.6)	.46
Energy (kcal)			
Baseline	1488.8 (522.3)	1522.8 (472.3)	.49
Second trimester	1539.2 (561.9)	1515.2 (543.9)	.69
Third trimester	1337.0 (495.0)	1373.4 (508.6)	.60
Change	−125.9 (660.0)	−120.8 (682.2)	.96
Carbohydrate (g)			
Baseline	203.1 (87.2)	203.6 (72.0)	.95
Second trimester	195.8 (79.7)	191.9 (70.6)	.63
Third trimester	171.4 (83.5)	173.3 (75.0)	.86
Change	−25.0 (107.0)	−24.7 (112.6)	.97
Protein (g)			
Baseline	60.93 (25.33)	60.27 (24.45)	.79
Second trimester	66.88 (28.71)	63.61 (25.81)	.26
Third trimester	63.00 (26.41)	61.89 (26.44)	.76
Change	−0.41 (34.4)	−1.0 (32.2)	.88
Fat (g)			
Baseline	48.1 (27.8)	52.1 (26.0)	.14
Second trimester	53.8 (33.1)	52.70 (30.5)	.75
Third trimester	44.8 (24.4)	47.7 (33.3)	.47
Change[c]	−2.5 (30.9)	−2.3 (42.4)	.97
Fiber (g)			
Baseline	11.1 (8.0)	10.2 (6.7)	.19
Second trimester	11.8 (8.1)	11.1 (6.8)	.41
Third trimester	8.9 (5.6)	11.6 (8.0)	.006
Change[c]	−2.4 (10.6)	1.5 (9.8)	.005

[a]Control group: n=200, 172, and 106 at baseline, second trimester, and third trimester, respectively.

[b]LGI diet group: n=200, 178, and 109 at baseline, second trimester, and third trimester, respectively.

[c]Change: levels in third trimester to levels at baseline.

Table 4. Comparisons of adherence to dietary interventions and missing data in the low glycemic index (LGI) diet group and the control group (N=200).

Descriptions	Control group, n (%)	Low glycemic index diet group, n (%)	*P* value
Adherence to dietary interventions			
At least 2 times	181 (90.5)	186 (93.0)	.36
All 3 times	106 (53.0)	109 (54.5)	.76
Missing data at end point			
Maternal outcomes			
Weight	17 (8.5)	14 (7.0)	.58
Insulin	85 (42.5)	77 (38.5)	.42
HbA_{1c}[a]	93 (46.5)	84 (42.0)	.37
Neonatal outcomes			
Birth weight	17 (8.5)	15 (7.5)	.71
Cord blood C-peptide	107 (53.5)	100 (50.0)	.48

[a]HbA_{1c}: glycosylated hemoglobin.

Discussion

Principal Findings

This trial was conducted in real-world clinical practice. Starting from early gestation, 3 individualized LGI diet consultations provided by a clinical dietician by using a diet GI and GL calculator failed to make a significant difference in maternal or neonatal insulin levels in overweigh and obese pregnant women compared with standard diet counseling. However, some changes in diet habits in the intervention group were observed that were different from those in the control group, including more fiber intake and less carbohydrate intake but comparable total energy intake, which are favorable to achieving lower diet GI.

Limitations

Our study has certain limitations. First, compliance to intervention in the 2 groups is lower than expected. The attendance to diet intervention interview remained over 90% at the second visit but dropped to 53.8% at the third visit. About 5% of the missed subjects were because of miscarriage and 5% because of moving to a different place for delivery; the overall rate is under assumption in sample size planning. The proportions of dropout are similar in the 2 arms; under the assumption of missing at random, missing values of primary outcomes are not likely to bias the comparisons, whereas this may lower the statistical power. A second explanation to poor compliance may be the fact that changing the diet habit is even harder for pregnant Chinese women, whose diet and nutrition status receive more attention from families. More intensive interactions with dietitians or convenient diet management tools, for example, a diet GI calculator, may help improve the overall compliance. In this trial, the diet GI calculated is performed by a researcher. If the tool was available to each subject and could be used every day, the compliance to a healthy LGI diet would be greatly improved. Second, nutrition intakes may be underestimated in this trial. Total energy intake was assessed by a 24-hour food record, which is lower than the recommendations for pregnant women by the Chinese Nutrition Society (1500 vs 2300 kcal/day) [16]. This may be due to systematic underestimation of the amounts of food intake by the 24-hour diet recall; however, this situation is equally distributed in the 2 arms and will not likely bias the comparisons. Moreover, to make recruitment easier, the gestational week for enrollment was extended to 16 weeks instead of ≤14 weeks in the trial registry. We think this change will not bias the main findings of the study. Finally, physical activity of subjects is not recorded.

Comparison With Previous Work

This study is one of a few trials to examine the effect of LGI diet intervention on insulin resistance of overweight pregnant women and their newborns. The intervention for the active intervention group is 3 individualized diet GL assessments and consultations using a mobile phone app starting at early gestation without providing any foods. The primary aim is to assess additional effectiveness of LGI diet intervention on maternal and neonatal insulin resistance levels compared with the conventional standard diet consultation according to the national nutrition recommendations for pregnant women.

We did not observe significant differences in maternal insulin levels at late gestation, change from baseline, or cord blood C-peptide, which is examined as markers of neonatal insulin resistance. These findings are contrary to our hypothesis, indicating that the standard diet consultation plus LGI diet intervention does not make significant additional improvement in maternal and neonatal insulin resistance. Insulin resistance level changes during gestation are not commonly measured as primary outcomes in previous LGI intervention trials. One study by Walsh et al conducted in a subgroup of low glycemic index diet in pregnancy to prevent macrosomia The Randomised cOntrol trial of Low (ROLO) study [28] reported some benefits of LGI diet consultations initiated from early gestation to maternal insulin resistance [29]. Participants in the LGI intervention group also gained less weight by 1.3 kg than those in the control group [28]; the effect size is similar to what we observed in our study (1.6 kg). However, controls of the ROLO study received routine antenatal care, involving no formal

dietary advice about GWG. In our study, the standard diet consultation received in the active comparative control arm may be too effective and make it hard to make a significant difference from the LGI group. This can be an important explanation to the negative findings in most of the outcomes. This is supported by the finding that the maternal insulin-level changes of the control group of our study are very close to the changes in the intervention group of the previous study (1.5 (SD 7.8) vs 1.79 (SD −2.6 to −4. 6), respectively) [28]. A number of studies have investigated the effect of LGI diet in different subjects, such as in pregnant women with gestational hyperglycemia [30], with GDM [13,31], women at high risk of GDM [32], or in healthy women [33]; clinical outcomes such as birth weight, incidence of GDM, insulin medication use, and GWG are examined. Evidence on the effects of LGI diet in overweight and obese pregnancies is limited. In total, 1 relevant trial is conducted in 46 overweight or obese pregnant women to compare the effect of an LGI diet with a low-fat diet [15]. Except for diet counseling, the interventions also provided LGI foods, such as carbohydrate-rich foods, fats, and snacks. However, they did not find any significant differences in insulin, HbA_{1c}, or HOMA-IR in the 2 groups. The results are compatible with our findings.

One strength of our study is the quantitative diet GI and GL estimation and records along with each visit of diet intervention, which allows us to observe the diet behavior changes. Effectiveness of the LGI intervention is expected to be achieved through change of diet habits, which is measured by food intake of different contents of nutrients. Our findings in diet fiber, carbohydrate, and GL support that these overweight and obese pregnant women have clear but very variable favorable responses to the diet intervention.

Although previous randomized trials show effectiveness and feasibility of using email, internet, or mobile phones in changing prediabetic individuals in changing their lifestyles and biological index [34], changing the lifestyle of overweight subjects by consultation is never an easy goal to achieve; it is even harder with pregnant women. Pregnant Chinese women traditionally are encouraged to eat more food than they need to ensure sufficient nutrition intake to the fetus. A study shows that gaining too much weight is a more critical issue in the Chinese population than other races [35]. We also find that a change of diet GI was weakly correlated with changes of maternal insulin levels, even after adjustment of maternal age and GWG. This result enhances our belief of the long-term healthy effectiveness of adopting the LGI diet for reducing metabolic risk in a high-risk population. The findings and manipulation experiences of the intervention (the longitudinal quantitative assessment of diet GI and low glycemic diet consultations with pregnant women) from this otherwise negative trial may be of interest to patients and clinicians in this field.

Conclusions

Compared with standard nutrition consultations with overweight or obese pregnant women, 3 individualized LGI diet consultations starting from early gestation do not make a significant difference in maternal or neonatal insulin levels. However, the LGI intervention, which is equipped with accurate diet GI and GL calculation, makes some differences in diet habits among overweight pregnant women, including more fiber intake and less carbohydrate intake, which are known to be favorable to achieve a lower diet GI. The pregnancy period is limited; we encourage to provide LGI diet education and an electronic diet GI calculation tool to acquire more significant effectiveness in reducing insulin resistance among a high-risk population.

Acknowledgments

The authors thank the nurses and technicians who are from the Kunshan Maternity and Child Care Center and the International Peace Maternity and Child Health Hospital of China Welfare Institute but not included in the author list for collecting the data and samples. We also thank all participants.

This work was financially supported by the National Science Foundation of China (Grant No. 81273168) and the National Key Scientific and Development Projects (Grant No. 2016YFC1000506).

Authors' Contributions

DN performed the data analysis, made statistical charts, and drafted the paper. YZ assisted with analysis of data and reviewed the paper. LPW and WHY took charge of the field work of the 2 centers and assisted with acquisition of routine clinical data. QZ, KM, and YZ contributed to acquisition and analysis of data. WY was the principal investigator of the study and contributed to the conception, study design, and critical approval of the draft and gave input at all stages of the study. All authors approved the final paper to be submitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Introduction to Diet GI app (in Chinese).

[PDF File (Adobe PDF File), 755KB - mhealth_v7i4e12081_app1.pdf]

References

1. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014 Aug 30;384(9945):766-781 [FREE Full text] [doi: 10.1016/S0140-6736(14)60460-8] [Medline: 24880830]
2. WHO. World Health Organization. 2015. Fact sheet: Obesity and overweight URL: https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2019-03-06] [WebCite Cache ID 76frwAxec]
3. Li JH, Wang LM, Li YC, Zhang M, Wang LH. Prevalence of major cardiovascular risk factors and cardiovascular disease in women in China: surveillance efforts. Biomed Environ Sci 2016 Mar;29(3):205-211 [FREE Full text] [doi: 10.3967/bes2015.025] [Medline: 27109131]
4. Catalano PM, Ehrenberg HM. The short- and long-term implications of maternal obesity on the mother and her offspring. BJOG 2006 Oct;113(10):1126-1133 [FREE Full text] [doi: 10.1111/j.1471-0528.2006.00989.x] [Medline: 16827826]
5. Johansson S, Villamor E, Altman M, Bonamy A, Granath F, Cnattingius S. Maternal overweight and obesity in early pregnancy and risk of infant mortality: a population based cohort study in Sweden. Br Med J 2014;349:a. [doi: 10.1111/j.1471-0528.2006.00989.x]
6. Ding X, Xu S, Hao J, Huang K, Su P, Tao F. Maternal pre-pregnancy BMI and adverse pregnancy outcomes among Chinese women: results from the C-ABCS. J Obstet Gynaecol 2016;36(3):328-332. [doi: 10.3109/01443615.2015.1050652] [Medline: 26492517]
7. Leng J, Shao P, Zhang C, Tian H, Zhang F, Zhang S, et al. Prevalence of gestational diabetes mellitus and its risk factors in Chinese pregnant women: a prospective population-based study in Tianjin, China. PLoS One 2015;10(3):e0121029 [FREE Full text] [doi: 10.1371/journal.pone.0121029] [Medline: 25799433]
8. Yang H, Wei Y, Gao X, Xu X, Fan L, He J, China National GDM Survey Working Group. Risk factors for gestational diabetes mellitus in Chinese women: a prospective study of 16,286 pregnant women in China. Diabet Med 2009 Nov;26(11):1099-1104. [doi: 10.1111/j.1464-5491.2009.02845.x] [Medline: 19929987]
9. Sebire N, Jolly M, Harris J, Wadsworth J, Joffe M, Beard R, et al. Maternal obesity and pregnancy outcome: a study of 287,213 pregnancies in London. Int J Obes Relat Metab Disord 2001 Aug;25(8):1175-1182. [doi: 10.1038/sj.ijo.0801670] [Medline: 11477502]
10. Jenkins D, Wolever T, Taylor R, Barker H, Fielden H, Baldwin J, et al. Glycemic index of foods: a physiological basis for carbohydrate exchange. Am J Clin Nutr 1981 Mar;34(3):362-366. [doi: 10.1093/ajcn/34.3.362] [Medline: 6259925]
11. Brand-Miller J, Stockmann K, Atkinson F, Petocz P, Denyer G. Glycemic index, postprandial glycemia, and the shape of the curve in healthy subjects: analysis of a database of more than 1,000 foods. Am J Clin Nutr 2009 Jan;89(1):97-105. [doi: 10.3945/ajcn.2008.26354] [Medline: 19056599]
12. Solomon TP, Haus JM, Kelly KR, Cook MD, Filion J, Rocco M, et al. A low-glycemic index diet combined with exercise reduces insulin resistance, postprandial hyperinsulinemia, and glucose-dependent insulinotropic polypeptide responses in obese, prediabetic humans. Am J Clin Nutr 2010 Dec;92(6):1359-1368 [FREE Full text] [doi: 10.3945/ajcn.2010.29771] [Medline: 20980494]
13. Hu Z, Tan R, Jin D, Li W, Zhou XY. A low glycemic index staple diet reduces postprandial glucose values in Asian women with gestational diabetes mellitus. J Investig Med 2014 Dec;62(8):975-979. [doi: 10.1097/JIM.0000000000000108] [Medline: 25203150]
14. Moses R, Barker M, Winter M, Petocz P, Brand-Miller J. Can a low-glycemic index diet reduce the need for insulin in gestational diabetes mellitus? A randomized trial. Diabetes Care 2009 Jun;32(6):996-1000 [FREE Full text] [doi: 10.2337/dc09-0007] [Medline: 19279301]
15. Rhodes E, Pawlak D, Takoudes T, Ebbeling C, Feldman H, Lovesky M, et al. Effects of a low-glycemic load diet in overweight and obese pregnant women: a pilot randomized controlled trial. Am J Clin Nutr 2010 Dec;92(6):1306-1315 [FREE Full text] [doi: 10.3945/ajcn.2010.30130] [Medline: 20962162]
16. Chinese-Nutrition-Society. Chinese dietary reference intakes. Acta Nutrimenta Sinica 2001;23:193-196.
17. Institute of Medicine (US) and National Research Council (US) Committee to Reexamine IOM Pregnancy Weight Guidelines, Rasmussen KM, Yaktine AL. Weight Gain During Pregnancy: Reexamining the Guidelines. Washington DC: National Academy of Sciences; 2009.
18. Neuhouser ML, Tinker LF, Thomson C, Caan B, Horn LV, Snetselaar L, et al. Development of a glycemic index database for food frequency questionnaires used in epidemiologic studies. J Nutr 2006 Jun;136(6):1604-1609. [doi: 10.1093/jn/136.6.1604] [Medline: 16702328]
19. Atkinson FS, Foster-Powell K, Brand-Miller J. International tables of glycemic index and glycemic load values: 2008. Diabetes Care 2008 Dec;31(12):2281-2283 [FREE Full text] [doi: 10.2337/dc08-1239] [Medline: 18835944]
20. Cook T, Rutishauser I, Seelig M. Comparable data on food and nutrient intake and physical measurements from the 1983, 1985 and 1995 national nutrition surveys. In: Australian Food and Nutrition Monitoring Unit. Canberra: Department of Health; Oct 10, 2001:2001 Fact sheet: Obesity and overweight URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-food-pdf-comp-data-cnt

21. Zhou B, Cooperative Meta-Analysis Group of the Working Group on Obesity in China. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults--study on optimal cut-off points of body mass index and waist circumference in Chinese adults. Biomed Environ Sci 2002 Mar;15(1):83-96. [Medline: 12046553]
22. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985 Jul;28(7):412-419. [Medline: 3899825]
23. Coustan D, Lowe L, Metzger B, Dyer A, International Association of DiabetesPregnancy Study Groups. The Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study: paving the way for new diagnostic criteria for gestational diabetes mellitus. Am J Obstet Gynecol 2010 Jun;202(6):654.e1-654.e6 [FREE Full text] [doi: 10.1016/j.ajog.2010.04.006] [Medline: 20510967]
24. Freathy RM, Hayes MG, Urbanek M, Lowe LP, Lee H, Ackerman C, HAPO Study Cooperative Research Group. Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study: common genetic variants in GCK and TCF7L2 are associated with fasting and postchallenge glucose levels in pregnancy and with the new consensus definition of gestational diabetes mellitus from the International Association of Diabetes and Pregnancy Study Groups. Diabetes 2010 Oct;59(10):2682-2689 [FREE Full text] [doi: 10.2337/db10-0177] [Medline: 20682688]
25. HAPO Study Cooperative Research Group. Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study: associations with neonatal anthropometrics. Diabetes 2009 Feb;58(2):453-459 [FREE Full text] [doi: 10.2337/db08-1112] [Medline: 19011170]
26. O'Rahilly SB, Burnett MA, Smith RF, Darley JH, Turner RC. Haemolysis affects insulin but not C-peptide immunoassay. Diabetologia 1987 Jun;30(6):394-396. [Medline: 3315796]
27. van der Ploeg HV, van Poppel MNM, Chey T, Bauman AE, Brown WJ. The role of pre-pregnancy physical activity and sedentary behaviour in the development of gestational diabetes mellitus. J Sci Med Sport 2011 Mar;14(2):149-152. [doi: 10.1016/j.jsams.2010.09.002] [Medline: 21030304]
28. Walsh JM, Mahony RM, Culliton M, Foley ME, McAuliffe FM. Impact of a low glycemic index diet in pregnancy on markers of maternal and fetal metabolism and inflammation. Reprod Sci 2014 Nov;21(11):1378-1381 [FREE Full text] [doi: 10.1177/1933719114525275] [Medline: 24642719]
29. Walsh JM, McGowan CA, Mahony R, Foley ME, McAuliffe FM. Low glycaemic index diet in pregnancy to prevent macrosomia (ROLO study): randomised control trial. Br Med J 2012 Aug 30;345:e5605 [FREE Full text] [doi: 10.1136/bmj.e5605] [Medline: 22936795]
30. Grant SM, Wolever TM, O'Connor DL, Nisenbaum R, Josse RG. Effect of a low glycaemic index diet on blood glucose in women with gestational hyperglycaemia. Diabetes Res Clin Pract 2011 Jan;91(1):15-22. [doi: 10.1016/j.diabres.2010.09.002] [Medline: 21094553]
31. Louie JC, Markovic TP, Perera N, Foote D, Petocz P, Ross GP, et al. A randomized controlled trial investigating the effects of a low-glycemic index diet on pregnancy outcomes in gestational diabetes mellitus. Diabetes Care 2011 Nov;34(11):2341-2346 [FREE Full text] [doi: 10.2337/dc11-0985] [Medline: 21900148]
32. Markovic TP, Muirhead R, Overs S, Ross GP, Louie JC, Kizirian N, et al. Randomized controlled trial investigating the effects of a low-glycemic index diet on pregnancy outcomes in women at high risk of gestational diabetes mellitus: the GI baby 3 study. Diabetes Care 2016 Jan;39(1):31-38. [doi: 10.2337/dc15-0572] [Medline: 26185283]
33. Moses RG, Luebcke M, Davis WS, Coleman KJ, Tapsell LC, Petocz P, et al. Effect of a low-glycemic-index diet during pregnancy on obstetric outcomes. Am J Clin Nutr 2006 Oct;84(4):807-812. [doi: 10.1093/ajcn/84.4.807] [Medline: 17023707]
34. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons with prediabetes. J Med Internet Res 2015 Oct 23;17(10):e240 [FREE Full text] [doi: 10.2196/jmir.4897] [Medline: 26499966]
35. Cheikh Ismail L, Bishop D, Pang R, Ohuma E, Kac G, Abrams B, et al. Gestational weight gain standards based on women enrolled in the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project: a prospective longitudinal cohort study. Br Med J 2016 Feb 29;352:i555 [FREE Full text] [doi: 10.1136/bmj.i555] [Medline: 26926301]

Abbreviations

BMI: body mass index
BP: blood pressure
GDM: gestational diabetes mellitus
GI: glycemic index
GL: glycemic load
GWG: gestational weight gain
HbA $_{1c}$: glycosylated hemoglobin
HOMA-IR: homeostasis model assessment of insulin resistance

LGI: low glycemic index
OGTT: oral glucose-tolerance test
ROLO: Randomised cOntrol trial of LOw

Edited by G Eysenbach; submitted 02.09.18; peer-reviewed by H Liang, G Liu; comments to author 07.11.18; revised version received 08.11.18; accepted 17.02.19; published 18.04.19.

Please cite as:
Zhang Y, Wang L, Yang W, Niu D, Li C, Wang L, Gu P, Xia Y, Shen Y, Yan J, Zhao Q, Mu K, Yan W
Effectiveness of Low Glycemic Index Diet Consultations Through a Diet Glycemic Assessment App Tool on Maternal and Neonatal Insulin Resistance: A Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(4):e12081
URL: http://mhealth.jmir.org/2019/4/e12081/
doi:10.2196/12081
PMID:30998227

Original Paper

Postvaccination Fever Response Rates in Children Derived Using the Fever Coach Mobile App: A Retrospective Observational Study

Sang Hyun Ahn[1*], MD; Jooho Zhiang[2*], BS; Hyery Kim[3], MD, PhD; Seyun Chang[4], BS; Jaewon Shin[4], MD; Myeongchan Kim[4], MD; Yura Lee[5], MD, PhD; Jae-Ho Lee[5,6], MD, PhD; Yu Rang Park[7], PhD

[1]Korea Human Resource Development Institute for Health and Welfare, Cheongju, Republic of Korea

[2]Yonsei University College of Medicine, Seoul, Republic of Korea

[3]Department of Pediatrics, University of Ulsan College of Medicine, Asan Medical Center Children's Hospital, Seoul, Republic of Korea

[4]Mobile Doctor Co, Ltd, Seoul, Republic of Korea

[5]Department of Biomedical Informatics, Asan Medical Center, Seoul, Republic of Korea

[6]Department of Emergency Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea

[7]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Yu Rang Park, PhD
Department of Biomedical Systems Informatics
Yonsei University College of Medicine
50-1 Yonsei-ro, Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2228 2493
Email: yurangpark@yuhs.ac

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2020/5/e18921/

Abstract

Background: Postvaccination fever is a mild adverse event that naturally improves without complications, but is highly prevalent and can be accompanied by febrile convulsions in some cases. These adverse effects may cause parents to delay or avoid vaccinating their children.

Objective: This study aimed to identify postvaccination fever patterns and the ability of antipyretics to affect changes in these patterns from data collected from a mobile app named Fever Coach.

Methods: Data provided by parents of feverish children derived from a mobile app, Fever Coach, were used to identify postvaccination fever patterns according to vaccinations and the use of antipyretic drugs. We selected single vaccination records that contained five or more body temperature readings performed within 48 hours of vaccination, and we analyzed postvaccination fever onset, offset, duration, and maximum body temperature. Through observing the postvaccination fever response to vaccination, we identified the effects of antipyretic drugs on postvaccination fever onset, offset, and duration times; the extent of fever; and the rate of decline. We also performed logistic regression analysis to determine demographic variables (age, weight, and sex) involved in relatively high fevers (body temperature ≥39°C).

Results: The total number of Fever Coach users was 25,037, with 3834 users having entered single vaccination records, including 4448 vaccinations and 55,783 body temperature records. Most records were obtained from children receiving the following vaccinations: pneumococcus (n=2069); Japanese encephalitis (n=911); influenza (n=669); diphtheria, tetanus, and pertussis (n=403); and hepatitis A (n=252). According to the 4448 vaccination records, 3427 (77.05%) children had taken antipyretic drugs, and 3238 (89.15%) children took antibiotics at body temperatures above 38°C. The number of children taking antipyretics at a body temperature of 38°C was more than four times that of those taking antipyretics at 37.9°C (307 vs 67 cases). The number of instances in which this temperature threshold was reached was more than four times greater than the number when the temperature was 37.9°C. A comparative analysis of antipyretic and nonantipyretic cases showed there was no difference in onset time; however,

offset and duration times were significantly shorter in nonantipyretic cases than in antipyretic cases (P<.001). In nonantipyretic cases, offset times and duration times were 9.9 and 10.1 hours shorter, respectively, than in antipyretic cases. Body temperatures also decreased faster in nonantipyretic cases. Influenza vaccine-associated fevers lasted relatively longer, whereas pneumococcus vaccine-associated fevers were relatively short-lived.

Conclusions: These findings suggest that postvaccination fever has its own fever pattern, which is dependent on vaccine type and the presence of antipyretic drugs, and that postvaccination temperature monitoring may ease fever phobia and reduce the unnecessary use of antipyretics in medical care.

(JMIR Mhealth Uhealth 2019;7(4):e12223) doi:10.2196/12223

KEYWORDS

patient-generated health data; vaccination; postvaccination fever; digital health care; mobile app

Introduction

The World Health Organization and the Korean Centers for Disease Control and Prevention recommend at least 10 and 14 vaccines, respectively, for routine immunization of children [1]. Despite benefits in preventing serious infectious diseases, vaccinations are also associated with a risk of adverse events in patients. Fever is the most commonly reported adverse event [2,3]. The immune system's lymphocyte and polymorphonuclear leukocyte functions improve at body temperatures between 38°C and 39°C. Therefore, the presence of postvaccination fever indicates that the immune system is functioning, but not due to a pathologic reaction [4-6]. There is no agreed definition of a postvaccination fever, but the gold standard definition of a fever is a rectal temperature of 38°C or above. Several studies report varying timeframes for postvaccination fever, ranging from within 32 hours, 48 hours, and 4 days of vaccination [7-9].

Postvaccination fever is a mild adverse event that is usually self-limiting over a few days without any specific complications. However, it is highly prevalent and, in some cases, accompanied with febrile convulsions. Furthermore, postvaccination fever may serve as an indicator of infectious disease. For these reasons, postvaccination fever can cause excessive anxiety in parents and caregivers [10-12]. Previous studies have indicated that parent perceptions and fear of fever have not significantly changed over the past 20 years and are still common in the Korean population [13-15]. Postvaccination fever phobia could lead to unnecessary testing, treatment (including the overuse of antipyretics), and emergency department visits, which increase both medical costs and the possibility of side effects [7,16].

Although fever can occur after every vaccination, a higher incidence of fever has been observed in patients receiving pneumococcal and DTaP (diphtheria, tetanus, and pertussis) vaccines than in those receiving other vaccines [17-19], indicating that fever patterns appear to be vaccine-specific. Many febrile diseases have specific fever patterns and progression (continuous, intermittent, remittent) that aid in understanding the pathophysiology of each disease and help to make diagnoses and therapeutic judgments [20,21]. Therefore, knowing the pattern and progress of postvaccination fever is likely to help address issues due to misperceptions regarding vaccination. There have been studies reporting on postvaccination fever frequency, but relatively few studies have reported postvaccination fever patterns [3,20-24]. However, many typical fever patterns can be changed through ingestion of antipyretics such as acetaminophen and steroidal anti-inflammatory drugs. The principal action of these drugs is inhibition of the enzyme cyclooxygenase and reduction of the levels of prostaglandin E_2 within the hypothalamus [25]. However, data regarding postvaccination fever patterns and progress are difficult to obtain, with most postvaccination fever-related medical care occurring in outpatient settings where accurate body temperature recordings are not easily obtainable.

Patient-generated health data (PGHD) acquired through high-mobility mediums such as mobile phones, Internet of Things, wearable devices, and mobile apps have recently emerged as alternative methods for collecting data [26-31]. Some studies have shown that these PGHD have improved patient treatment, and studies are underway regarding how PGHD can be linked to routine treatment [32,33]. With immunization, however, obtaining longitudinal and continuous data is difficult. This study aimed to investigate the fever patterns of postvaccination fever through retrospectively analyzing PGHD obtained using Fever Coach, a fever management mobile app, and to analyze changes in fever patterns with the use of antipyretics.

Methods

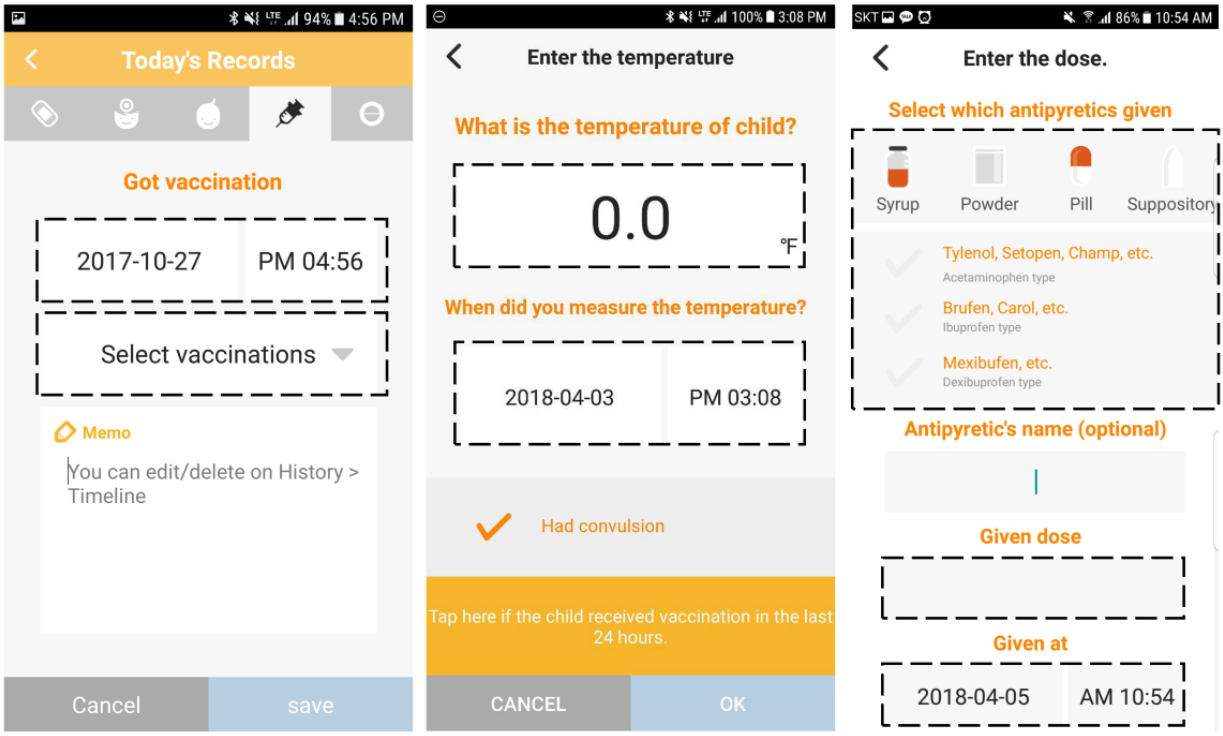
Mobile App Description

Fever Coach is a mobile health care app developed by Mobile Doctor for parents with a feverish child. The app is based on pediatric thermal standards; it assesses a child's condition based on user input and provides guidelines for antipyretic use. The app provides services that support parents' effective and accurate control of common fever symptoms. Fever Coach provides several data services related to managing fever in children, such as microdust concentration status, body temperature information depending on geographical area, disease epidemic alerts, and pediatric health information. The app was made available as a free download from the Google Play Store and Apple App Store. As of June 31, 2017, 197,555 people had registered their child with the app.

We collected vaccination records and records of subsequent postvaccination fever responses and the antipyretics administered. Figure 1 shows the detail screens of the Fever Coach app and the types of data users can enter. All screens of Fever Coach are in Multimedia Appendix 1. The "Today's Records" function provides information for vaccination records. The "Enter the Temperature" function allows the user to provide

data concerning the fever response. The "Enter the Dose" function allows the user to provide records of antipyretic use.

Figure 1. Screenshots of vaccination and fever response in the Fever Coach app. The left screen ("Today's Records"), middle screen ("Enter the temperature"), and right screen ("Enter the dose") have areas for user-input data. The functions corresponding to vaccination data, fever response, and antipyretic data in the three screens are indicated with dotted boxes. The original app showed Korean menu names; for international use, they have been translated into English.



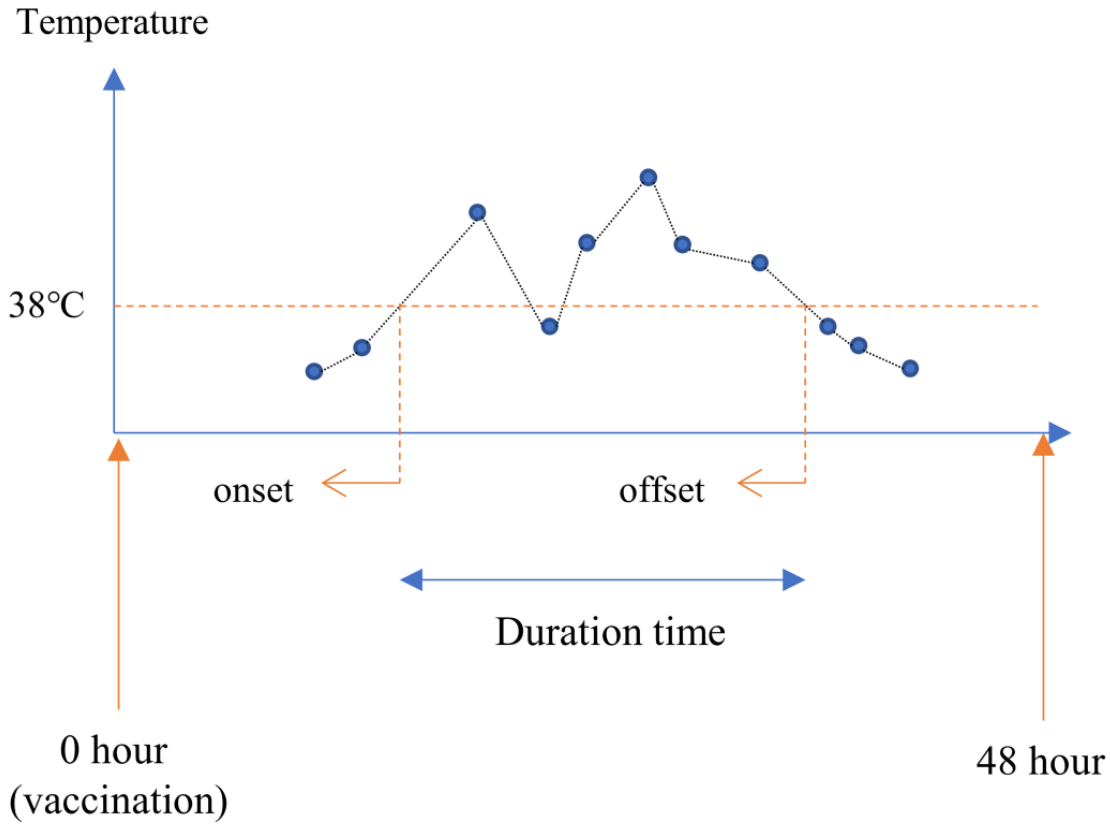
Study Design

To identify vaccination and antipyretic effects on postvaccination fever response patterns, we analyzed the logs of all users who signed up and registered their child between July 2015 and June 2017. Postvaccination fever usually lasts less than 48 hours [34]; therefore, we examined fever response and antipyretic administration data for 48 hours after the vaccination date entered by the user. For postvaccination fever response analyses, we defined the onset of fever as the point when the body temperature exceeded 38.0°C, and the offset as the point when the body temperature fell below 38.0°C [35,36].

Onset time was defined as the time between vaccination and the point at which body temperature exceeded 38.0°C. If the first record was above 38.0°C, the time from vaccination to the registration of the data was defined as the onset time. Similarly, offset time was defined as the time between vaccination and the last point at which the body temperature fell below 38.0°C. The duration time was defined as the time elapsed between onset and offset points (Figure 2). Body temperature values were obtained using linear imputation techniques when they were missing between two actual body temperatures in the neighbors (Figure 2). We used this linear imputation technique on the assumption that the fever progression would show linear characteristics.

This study was approved by the Institutional Review Board of the Asan Medical Center (IRB no. 2018-0179; Seoul, South Korea). The need for informed consent was waived by the Ethics Committee because this study used routinely collected log data that were anonymously managed at all stages, including during data cleaning and statistical analyses.

Figure 2. Onset, offset, and duration time definitions. The x-axis represents the time since vaccination and the y-axis represents the body temperature. The blue dot represents the actual body temperature. The black line between the blue dots is the imputated body temperature.



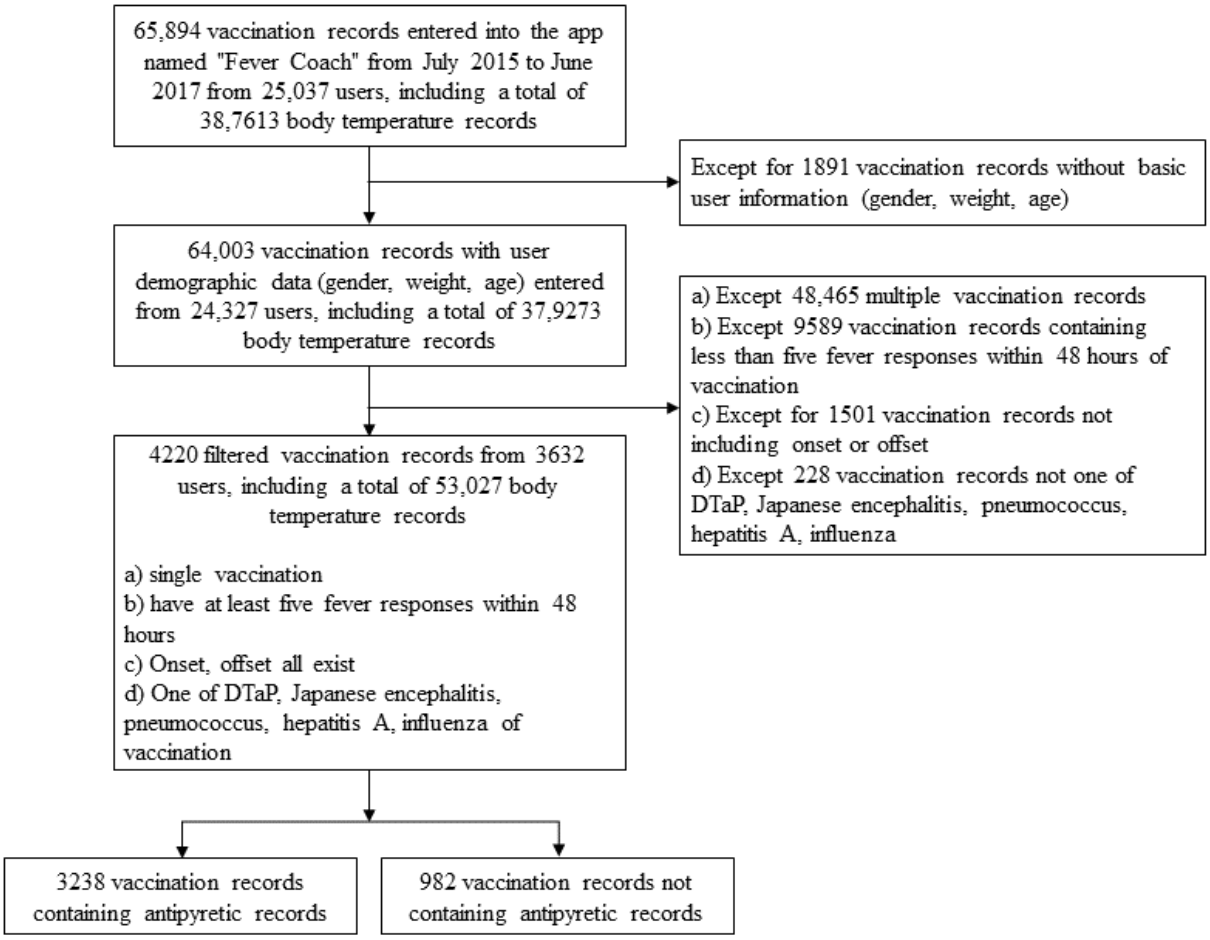
Data Collection and Analysis

The app allows users to enter data regarding vaccinations, body temperatures, and antipyretic drug administration history. The vaccination record consists of the type and time of vaccine given to a particular child. The body temperature record consists of body temperature and measurement time for a specific child. The antipyretic drug administration record consists of antipyretic agent type, dose, and time. Figure 3 shows the target population selection flow of the study. From July 2015 to June 2017, 65,894 vaccination records were entered, of which 64,003 were vaccination records with basic user information (sex, weight, and age). Among them, the number of single vaccination records was 15,538. To obtain a sufficient number of body temperature recordings and significant onset and offset values, these data were filtered into 5949 cases with at least five body temperatures recorded within 48 hours after vaccination. Of these, 4448 comprised data-containing points corresponding to onset and offset times. Among these data, bacillus Calmette-Guérin, hepatitis B, measles-mumps-rubella (MMR), *Haemophilusinfluenzae* type B (Hib), polio, rotavirus, and chickenpox were excluded from the analysis because the number of vaccination records was less than 100; therefore, only data regarding DTaP, Japanese encephalitis, pneumococcus, hepatitis A, and influenza were analyzed. Although DTaP and MMR vaccines are combination vaccines, we considered these combination vaccines to be single vaccinations because they are single shots that can be distinguished from multiple vaccinations that involve multiple shots and that vaccines contain a variety of ingredients that can cause side effects [37].

The analysis was undertaken by grouping the data into 3238 cases with antipyretic records and 982 cases with no antipyretic records. For each vaccination record, the sex and age of the child were also collected.

To confirm the postvaccination fever differences between vaccination and antipyretic drug administration, we performed statistical analysis. To compare differences in onset times, offset times, duration times, and maximum temperatures, an independent sample *t* test was used to determine the degree of difference, while the *P* value for a two-sided test was used to test for significance. To observe the process of fever for each vaccine record, the onset time was defined as reference time 0, and body temperature values recorded over the previous 3 hours and the following 24 hours were obtained. Cases were grouped according to whether they were given antipyretic drugs and were reclassified according to vaccine type. Additionally, an ANOVA test was performed to compare maximum temperature and fever duration among the vaccine types. We used a Dunnett T3 post hoc test because the *P* value of the Levene test was less than .001, which indicated a violation of the assumption of homogeneity of variance. Lastly, binary logistic regression analysis was performed to determine demographic variables (age, weight, and sex) involved in relatively high fevers (body temperature ≥39°C). We performed multiple regression tests to identify multicollinearity between age and weight. If the variance inflation factor (VIF) was less than 10, it was considered that there was no multicollinearity for those variables. Data were processed and analyzed using R version 3.5.0, SPSS 21.0, and Python 3.6 (including packages of Pandas 0.22.0, NumPy 1.14.3, and Jupyter 1.0.0).

Figure 3. Data collection flowchart.



Results

Overall Characteristics

During the 24 months of the Fever Coach operation, a total of 25,037 users recorded 65,894 vaccinations involving 387,613 body temperature records for 25,608 children. Of the 64,003 vaccination records for which a child's basic information was available, 15,538 (24.27%) were single vaccination records. Of the total number of enrolled children, there were 3834 (14.97%) children with five or more body temperature records at onset and offset, with 4448 vaccination records and 55,783 body temperature records (Table 1). The age at vaccination was significantly different in relation to the vaccine types. The proportion of males in this study was 60.12% (2193/3648). The proportion of vaccination records with antipyretic drugs was 77.05% (3427/4448). The majority of records were from children receiving the following vaccinations: pneumococcus (n=2069), Japanese encephalitis (n=911), influenza (n=669), DTaP (n=403), and hepatitis A (n=252). Each of the remaining single vaccinations had less than 100 records each.

Table 1. Basic characteristics according to vaccine type and the presence of antipyretic drugs.

Type of vaccination (number of records, number of children)	Age (months)		Sex (male), n (%)	Weight (kg)		Vaccination records with antipyretics, n (%)	Body temperature records, mean (SD)	Body temperature (°C)	
	Median (IQR[a])	Mean (SD[b])		Median (IQR)	Mean (SD)			Median (IQR)	Mean (SD)
BCG[c] (16, 13)	6.6 (21.5)	13.1 (13.3)	8 (61.5)	7.7 (15.0)	8.6 (3.7)	10 (62.5)	10.3 (5.0)	37.7 (0.9)	37.6 (0.8)
Chickenpox (31, 28)	14.3 (6.2)	18.6 (12.3)	19 (67.9)	10.1 (35.0)	11.3 (4.7)	29 (93.5)	14.0 (7.7)	38.1 (1.0)	38.1 (0.8)
DTaP[d] (394, 352)	13.3 (14.8)	14.9 (13.0)	212 (60.2)	9.5 (27.8)	9.5 (3.1)	318 (80.7)	12.7 (7.2)	37.9 (0.8)	37.9 (0.7)
Hepatitis B (61, 55)	10.7 (7.3)	14.5 (14.4)	37 (67.3)	9 (22.3)	9.7 (3.7)	54 (88.5)	12.1 (9.2)	37.9 (0.9)	37.9 (0.7)
Hepatitis A (247, 223)	19.8 (9.1)	20.4 (6.2)	130 (58.3)	11 (16.0)	11 (1.6)	229 (92.7)	14.3 (9.8)	38 (1.0)	38 (0.7)
Hib[e] (47, 40)	15.3 (9.1)	14.4 (6.0)	20 (50.0)	9.7 (15.0)	9.4 (2.0)	43 (91.5)	12.1 (5.9)	37.9 (0.9)	37.9 (0.7)
Influenza (655, 587)	17.2 (7.9)	28 (20.8)	325 (55.4)	11.5 (39.0)	12.5 (4.3)	576 (87.9)	13.3 (8.1)	38 (0.9)	38 (0.8)
Japanese encephalitis (890, 793)	17.2 (7.9)	19.1 (8.1)	422 (53.2)	10.4 (23.5)	10.7 (1.8)	792 (89.0)	12.0 (7.5)	37.9 (0.8)	37.9 (0.7)
MMR[f] (31, 28)	16.1 (10.0)	23.4 (15.8)	15 (53.6)	11 (19.0)	11.9 (2.8)	29 (93.5)	13.3 (6.2)	38.1 (1.1)	38.1 (0.8)
Pneumococcus (2034, 1771)	7.1 (11.0)	9.3 (8.1)	981 (55.5)	8 (32.0)	8.2 (2.5)	1323 (65.0)	12.3 (7.7)	37.7 (0.7)	37.8 (0.6)
Polio (19, 17)	12.3 (11.2)	12.1 (9.2)	11 (64.7)	8.8 (14.0)	9.3 (2.7)	11 (57.9)	11.5 (6.5)	37.7 (0.7)	37.7 (0.7)
Rotavirus (23, 22)	5.2 (4.4)	5.5 (3.1)	13 (59.1)	7 (11.5)	7.1 (1.7)	13 (56.5)	9.6 (5.1)	37.7 (0.9)	37.7 (0.5)
Total (4448, 3648)	13.5 (13.0)	15.4 (13.3)	2193 (60.1)	9.7 (3.5)	9.7 (3.2)	3427 (77.0)	12.5 (7.8)	37.7 (0.7)	37.9 (0.7)

[a]IQR: interquartile range.

[b]SD: standard deviation.

[c]BCG: bacille Calmette-Guérin.

[d]DTaP: diphtheria, tetanus, pertussis.

[e]Hib: *Haemophilusinfluenzae* type b.

[f]MMR: measles, mumps, and rubella.

Antipyretic Drug Administration Pattern

After vaccination, 14.66% (475/3238) of children were administered the first antipyretic drug within 1 hour, and more than 50% of children were administered antipyretic drugs within 10 hours. By 10 hours, the number of children treated with antipyretic drugs increased gradually, but decreased from 11 hours to less than 10% at 41 hours. A total of 2887 (89.16%) children used antipyretics when the postvaccination body temperature was 38.0°C or above (Figure 4). The number of children taking antipyretics at a body temperature of 38°C was more than four times that of those taking antipyretics at 37.9°C (307 vs 67 cases). The percentages of children who received antipyretic drugs were 0.71% (23/3238) and 0.49% (16/3238) at body temperatures lower than 37°C and above 40°C, respectively.

Comparison of Onset, Offset, and Duration Times Among Vaccination Records With and Without Antipyretic Administration

There were significant differences in offset times, duration times, and maximum temperatures between groups that had taken antipyretics and those that had not, but there was no significant difference in onset times (Figure 5). Differences between groups were marked, especially maximum temperatures that affected offset time and duration. Children vaccinated against hepatitis A showed the greatest difference in maximum temperature (mean 39.0ºC, SD 0.6ºC vs mean 38.4ºC, SD 0.4ºC in children who had been administered antipyretic drugs compared to children who had not been administered antipyretic drugs). The postvaccination fever duration was three times longer in children who were administered antipyretic drugs than in children who were not administered antipyretic drugs (mean 17.4, SD 11.8 hours vs mean 4.9, SD 7.5 hours, with antipyretic drugs compared to without antipyretic drugs).

Postvaccination fever differences between onset, offset, and duration times associated with each vaccine showed statistically significant differences in offset times and duration between groups (P<.001) (Multimedia Appendix 2). Offset and duration times for children who were not administered antipyretics were significantly shorter than for those who were administered antipyretics. There was no significant difference in onset time between the groups.

Figure 4. Time of first administration of antipyretic after vaccination and body temperature at the time of antipyretic use.

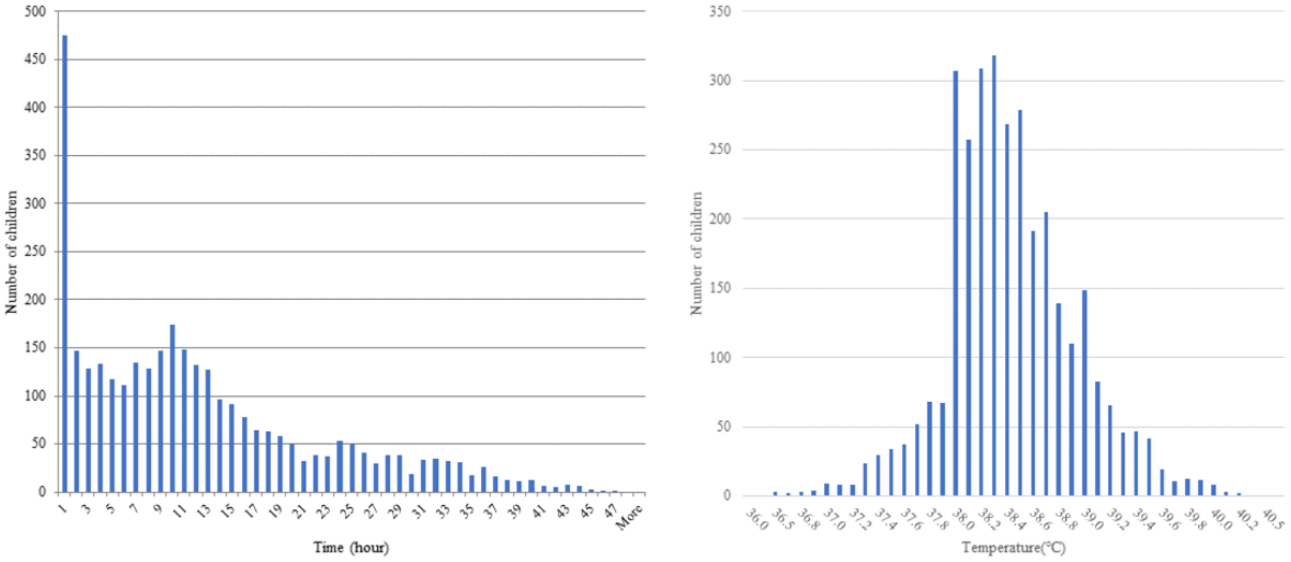
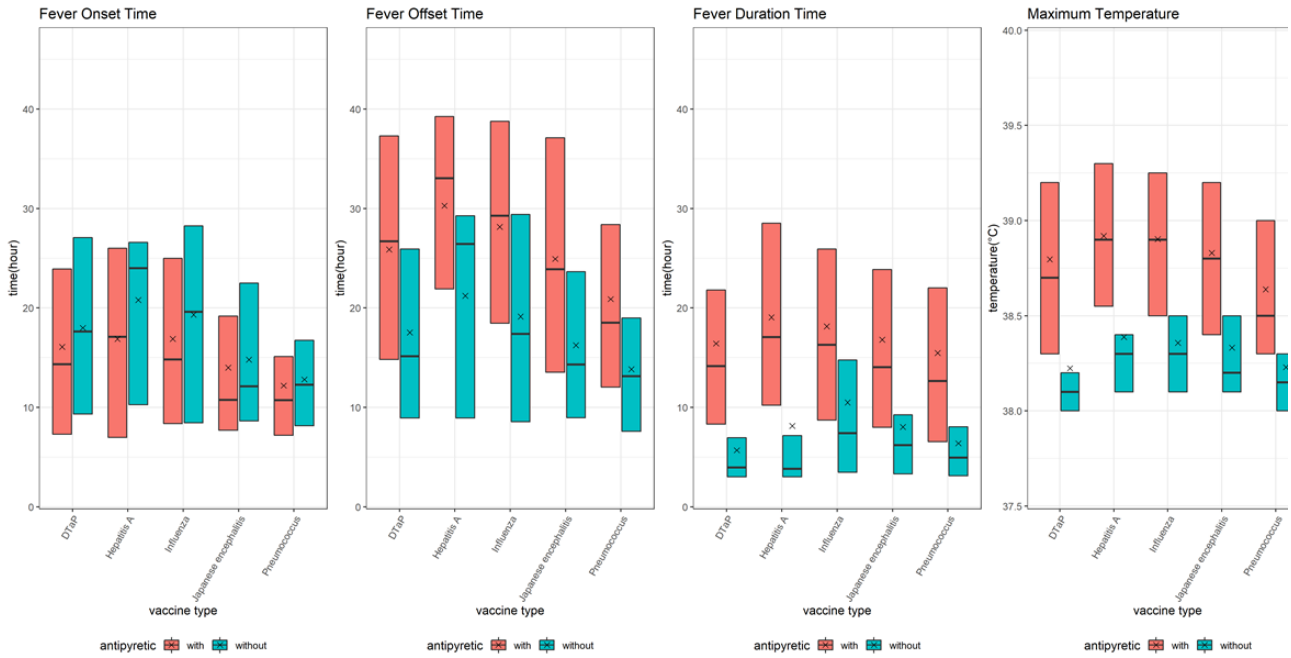


Figure 5. Comparison of onset, offset, duration times, and maximum temperatures among vaccine types, and the effects of antipyretics on postvaccination fever response. From left to right: box plots of onset, offset, duration times, and maximum body temperature are depicted. The bar indicates the median; x indicates the mean.



We found statistically significant differences in maximum temperature and fever duration among vaccines (both $P<.001$). In the postanalysis grouping, maximum temperature and fever duration decreased in the following order: hepatitis A, influenza, Japanese encephalitis, DTaP, and pneumococcus (Table 2).

We used logistic regression to determine the demographic variables involved in relatively high fevers (Table 3). There was no multicollinearity between age and weight (age VIF=4.38, weight VIF=4.38). Sex was not a significant variable in all groups, and age was not significant except in children who were administered the pneumococcus vaccine. Weight was statistically significant in all cases except in those receiving the hepatitis A vaccine.

Table 2. Comparison of maximum temperature and fever duration among vaccine types.

Vaccine type	Mean (SD)	$F_{4,4215}$	*P* value[a]
Maximum temperature (°C)		94.7	<.001
Pneumococcus	38.5 (0.5)		
DTaP	38.7 (0.6)		
Hepatitis A	39.0 (0.6)		
Influenza	38.9 (0.6)		
Japanese encephalitis	38.8 (0.5)		
Fever duration (hours)		62.8	<.001
Pneumococcus	8.5 (10.5)		
DTaP	11.3 (11.1)		
Hepatitis A	16.4 (12.0)		
Influenza	14.6 (12.3)		
Japanese encephalitis	12.8 (11.7)		

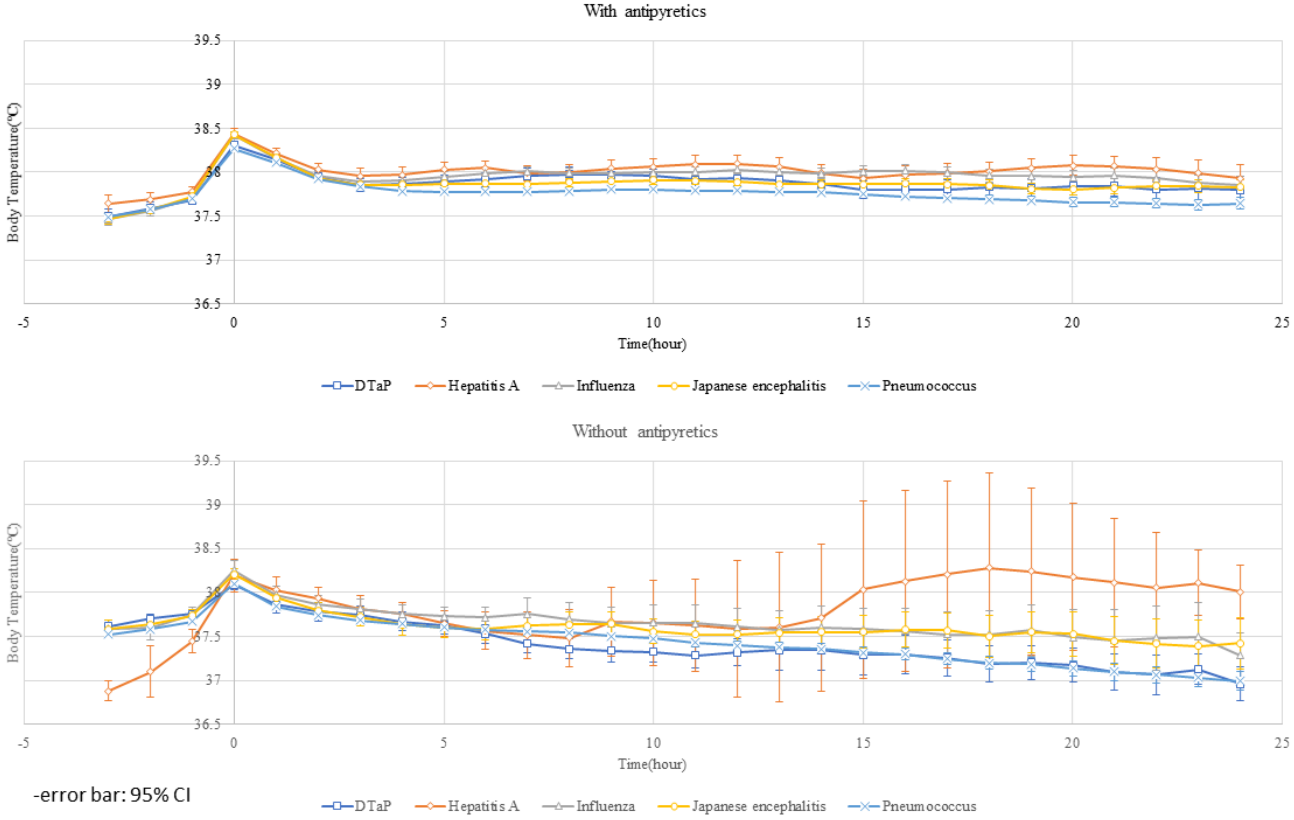
[a]As a result of the post hoc test (Dunnett T3), maximum temperature and fever duration showed the following: Hepatitis A, influenza > Japanese encephalitis, DTaP > pneumococcus.

Table 3. Variables involved in relatively high postvaccination fever (≥39°C).[a]

Vaccine type and variable	B (SE)	*P* value	Adjusted odds ratio (95% CI)
Pneumococcus			
Age	0.056 (0.014)	<.001	1.057 (1.029-1.086)
Weight	0.210 (0.042)	<.001	1.233 (1.135-1.340)
Sex	–0.096 (0.121)	.42	0.908 (0.716-1.152)
DTaP			
Age	–0.004 (0.018)	.82	0.996 (0.960-1.033)
Weight	0.234 (0.083)	.005	1.263 (1.075-1.485)
Sex	–0.331 (0.231)	.15	0.718 (0.457-1.129)
Japanese encephalitis			
Age	0.006 (0.012)	.59	1.006 (0.984-1.029)
Weight	0.108 (0.052)	.03	1.114 (1.005-1.234)
Sex	–0.070 (0.142)	.62	0.933 (0.706-1.233)
Hepatitis A			
Age	0.036 (0.024)	.13	1.037 (0.989-1.088)
Weight	0.011 (0.094)	.90	1.011 (0.841-1.217)
Sex	0.158 (0.264)	.55	1.171 (0.698-1.966)
Influenza			
Age	–0.016 (0.010)	.09	0.984 (0.966-1.003)
Weight	0.104 (0.047)	.02	1.109 (1.102-1.216)
Sex	0.002 (0.160)	.98	1.002 (0.732-1.372)
All			
Age	–0.004 (0.006)	.43	0.996 (0.985-1.007)
Weight	0.226 (0.025)	<.001	1.254 (1.195-1.316)
Sex	–0.106 (0.071)	.13	0.899 (0.782-1.034)

[a]CI: confidence interval; DTaP: diphtheria, tetanus, pertussis; SE: standard error.

Figure 6. Body temperature graph over time for each vaccine showing the effects of antipyretic administration. The empty circle indicates the mean; the error bar represents a 95% confidence interval.



Comparison of Fever Response

Figure 6 shows the result of plotting body temperature over time for each vaccine type and the effects of antipyretic drug administration. In cases where antipyretics were not administered, especially concerning DTaP and pneumococcus, the slope of the graph was steeper than in cases that included antipyretics, indicating that body temperature dropped more rapidly in children who were not administered antipyretic drugs (DTaP R^2: .84 vs .00; pneumococcus R^2: .83 vs .14). In the antipyretic use group, the mean body temperature was 37.5°C or above even after 24 hours, but in the nonantipyretic use group, the mean body temperature dropped below 37.5°C before 20 hours for all vaccines except hepatitis A. In comparing vaccines, we found that postvaccination fever associated with the influenza vaccine tended to be relatively long-lasting, and that the pneumococcus vaccine showed a relatively rapid decline.

Comparison of the Effects of Antipyretics on Postvaccination Fever Response

Multimedia Appendix 2 shows the comparison of onset, offset, duration times, maximum temperatures among vaccine types, and the effects of antipyretics on postvaccination fever response. Postvaccination fever in children who were administered antipyretics exhibited a mean onset time of between 9.0 and 13.7 hours, a mean offset time of between 20.4 and 31.1 hours, a mean duration time of between 11.5 and 17.4 hours, and a maximum temperature of between 38.7°C and 39°C. The mean postvaccination fever onset time was not significantly different between children who were administered antipyretics and those who were not (mean 10.9, SD hours 9.9 vs mean 10.4, SD hours, 9.3 P=.12), but the mean offset time and duration were significantly different (offset time: mean 24.3, SD hours 13.4 vs mean 13.6, SD 9.9 hours, P<.001; duration: mean 13.4, SD 11.9 hours vs mean 3.2, SD 5.0 hours, P<.001). The maximum body temperatures at onset time were mean 38.8°C (SD 0.6°C) in the antipyretic group and mean 38.3°C (SD 0.5°C) in the nonantipyretic group, with a statistically significant difference between the two groups (P<.001).

Discussion

Principal Findings

We identified different postvaccination fever responses for each vaccine using data collected through the mobile app Fever Coach. For example, in relation to the hepatitis A vaccine, the postvaccination fever maximum temperature was statistically significantly higher and was of a longer duration than that for other vaccines (P<.001). We were also able to clearly identify differences among postvaccination fevers with and without antipyretic administration; offset, duration times, and maximum temperature were significantly different between groups (P<.001). Additionally, we presented new evidence from large-scale PGHD concerning fever duration and maximum temperatures reached after a single vaccination. The postvaccination fever duration was generally 48 hours but, for all single vaccinations, the longest postvaccination fever duration was mean 16.4 (SD 12.0) hours, and the highest maximum temperature was mean 39.0°C (SD 0.6°C). In terms of informatics, an additional contribution of this study is that

the majority of users verified the pediatric fever level based on the PGHD. This study is original compared with previous studies for the following reasons: we used actual postvaccination fever response data without any clinical intervention; we analyzed large-scale postvaccination fever responses based on real-world data rather than relying on data from a small, clinical-based group; and our study used data derived from anyone using a mobile phone and not just from a specific hospital or specific patient area, which allowed us to analyze individual data covering a larger area than from data from a single institution or specific area.

Characteristics of Postvaccination Fever and Its Natural Course According to Vaccine Type

Because most previous studies have been conducted to determine vaccine safety, they primarily discussed the frequency and severity of postvaccination fever and did not provide medical direction to parents or caregivers. This study showed that postvaccination fever associated with each vaccine has a unique fever pattern in terms of maximum temperature and duration.

Among all vaccines included in the study, the postvaccination fever in the nonantipyretic use group showed a mean onset time of 9.4 hours (pneumococcal vaccine) to 15.3 hours (hepatitis A vaccine), a mean offset time of 12.5 hours (pneumococcal vaccine) to 20.1 hours (hepatitis A vaccine), and a mean duration time of 2.5 hours (DTaP vaccine) to 6.1 hours (influenza vaccine) (details in Multimedia Appendix 2). These values are considered consistent with associated fever and vaccination administration if the fever occurs within 24 hours of vaccination and if the temperature falls within 24 hours after onset in general practice [38]. The characteristics of postvaccination fever according to vaccine type in all data were as follows: the maximum temperature was highest for children vaccinated against hepatitis A and lowest for children receiving the pneumococcal vaccine. Children receiving the pneumococcal vaccine are known to have a high incidence of postvaccination fever; therefore, it is not uncommon for clinical practitioners to give a warning before and after vaccination and recommend that parents administer antipyretic agents when fever occurs [19,39]. However, the results of this study showed that the mean temperature following pneumococcal vaccination was significantly lower than for other vaccinations, regardless of the use of antipyretics. Fever duration was longest for hepatitis A and influenza, followed by Japanese encephalitis, pneumococcus, and DTaP vaccine. Postvaccination fever associated with pneumococcal vaccine had a lower maximum temperature and a shorter duration than postvaccination fever associated with hepatitis A vaccine but fell below the fever onset temperature (38°C) within 24 hours after fever occurred. If parents were more aware of this natural course of postvaccination fever, they may be less anxious and reduce the use of antipyretics, thereby reducing unnecessary medical care [40,41]. We would expect parents to have a more positive outlook regarding vaccinations once their fears concerning fever had been more fully addressed [15].

Our results showed that, as a child's weight increased, there was a high probability that a relatively high fever with a body temperature of less than 39°C would occur, regardless of age or sex (Table 3). It is known that the incidence of fever (especially a high fever with a body temperature of 39°C) tends to increase in children receiving the pneumococcal booster vaccination after 1 year of age, which is thought to be due to the booster effect, as the immune function matures and the number of vaccinations increase with age [24]. Here, however, only weight was significantly associated with the risk of high fever. Generally, a child's age and weight are linearly correlated. The immune response to vaccination is known to be higher for girls than for boys [42,43], but there was no difference in the frequency of relatively high fevers between the sexes.

Postvaccination Fever and Antipyretics

The duration of postvaccination fever in children who were administered antipyretics tended to be more prolonged than for those who were not administered antipyretics in this study. However, the actual body temperature at onset time and the maximum temperature were significantly higher in children who were administered antipyretics than in those who were not administered antipyretics (mean 38.8, SD 0.6 vs mean 38.3, SD 0.5, respectively, $P<.001$). Therefore, it is possible that if the period of decrease to the fever onset temperature (38°C) is prolonged, antipyretics could be responsible for inhibiting the immune response and extending the duration. It appears that the use of antipyretic drugs influences the offset time rather than the onset time, shortening the duration to offset time. Antipyretics were associated with a trend toward prolonged duration of illness in a group infected with *Shigellasonnei* and the influenza virus [44,45]. A systematic review indicated that the use of antipyretics in cases of malaria or viral diseases could shorten the duration of fever without prolonging the course of disease [38,39]. As noted, the effect of antipyretics on the course of postvaccination fever remains inconclusive, and the effect of antipyretics on postvaccination fever remains unknown. Based on the results of this study, it is possible that antipyretic use may prolong fever duration in children with postvaccination fever.

According to surveys on the use of antipyretics for postvaccination fever prevention and management by parents and caregivers, 11% of parents administered antipyretics for prophylaxis, and 64% of parents administered antipyretics for prevention or management within 48 hours of vaccination [46]. Here, after vaccination, 14.66% of children were administered their first antipyretic drugs within 1 hour and 16.2% of children were administered antipyretics for fever with a body temperature of less than 38°C (Figure 4). This value was higher than in previous studies and may be due to (1) differences in data used in the analysis, (2) cultural differences in the perception of fever, and (3) relative underestimation due to recall bias arising from retrospective survey methods used in prior studies. Also, many parents and caregivers were administering antipyretics for management of postvaccination fever when body temperatures reached 38°C, most likely because the app instructions advised parents not to administer antipyretics below a body temperature of 38°C but rather to administer antipyretics when the body temperature was above 38°C depending on the child's condition. Various studies have been conducted on the risks and benefits of using antipyretics for postvaccination fever. Febrile

convulsions are the most worrying situation for parents and caregivers when fever is present. To prevent convulsions, prophylactic antipyretics and routine antipyretics have been used during fevers, but with no effect on reducing febrile convulsions confirmed [23,39,47]. Additionally, several studies have suggested that antipyretics may be associated with decreased immunogenicity, and routine antipyretic administration is no longer recommended for postvaccination fever in countries such as Canada and New Zealand [22,24,48-51]. Alternatively, a systematic review has shown the difficulty of concluding that the reduction in immunogenicity due to antipyretic use does not fall below the seroprotective level and, thus, does not represent actual vaccine failure [48]. A high proportion of antipyretic use for postvaccination fever is believed to be due to fever phobia and a parent's expectation that administering antipyretic drugs will ease a child's discomfort rather than being based on persuasive scientific evidence [52]. Although antipyretic drugs had a temporary effect in lowering body temperature for a mean of 4.4 hours, it is not known whether their use reduced postvaccination fever-related discomfort. To guide the use of antipyretics for alleviating postvaccination fever-related discomfort based on scientific evidence, the effects of antipyretics on fever patterns and child discomfort need to be clarified through double-blinded randomized clinical trials.

Limitations and Future Work

In our study, fever was defined as a body temperature above 38°C regardless of measurement method or age [53]. Although in cases where rectal temperature measurements of 38°C or above, or 1°C or more above basal body temperature, may be defined as a fever [35,36], rectal temperature measurements are not a readily applicable method of measuring body temperature, and infrared tympanic thermometers, noncontact infrared-forehead thermometers, axillary thermometers, and oral thermometers are most commonly used in practice [54,55] with inconsistent results [56-58]. Body temperature may also vary depending on the age and biological factors of an individual. The dataset used in this study did not contain any information concerning the temperature measurement method or site, and it was difficult to analyze all relevant factors including age and biological factors in relation to temperature measurement. Therefore, we defined a conservative standard for fever as a body temperature greater than 38°C. This criterion was consistent with the body temperature at which the child was administered the antipyretic drug. A more accurate analysis would be possible using a definition of fever as a body temperature of 1°C above the basal body temperature or above 38°C, as well as obtaining the measurement method and site, and acquiring mandatory data on basal body temperature.

Data used here were obtained by app users directly entering their child's body temperature. Since a uniform standard for measuring body temperatures was not applied (including measurement device and site), questions can be raised concerning the accuracy of the body temperature data. Furthermore, the app is dependent on the user entering the data correctly and consistently.

Despite these limitations, the five vaccines included in this study were more frequently recorded compared to the vaccines that we excluded. It is possible that postvaccination fever is actually more likely to occur due to these vaccines than others, and parents may also have believed that postvaccination fever in relation to these vaccines is more common than with others. Therefore, they may have been more likely to perform body temperature measurements. These limitations could be overcome through wearing thermometers and body temperature recording applications that continuously measure body temperatures before and after vaccination.

In practice, multiple vaccinations are recommended rather than single vaccinations, and more people are having multiple vaccinations, since there are no challenges in obtaining immunogenicity, no increase in side effects, and immunization schedules can be simplified [59]. However, multiple vaccinations may have an effect on postvaccination fever patterns depending on vaccine types. This study focused on single vaccinations; subsequent studies are underway concerning multiple vaccinations.

In our study, an analysis of the use of antipyretic drugs was only conducted after vaccination. However, the postvaccination fever pattern can vary depending on when an antipyretic drug is administered. More detailed research on postvaccination fever patterns is needed to determine when an antipyretic drug is to be administered.

Through creating a model that can predict future progress, based on the natural course of postvaccination fever (Figure 6), postvaccination temperature monitoring can provide useful information to parents, caregivers, and health care professionals, which is likely to reduce unnecessary tests and treatments, and consequently contribute to improvements in children's health.

Conclusion

Postvaccination fever has its own fever pattern depending on the type of vaccine administered. The pattern of postvaccination fever can be altered using antipyretic drugs, making the diagnosis of postvaccination fever difficult. This study showed that antipyretic drugs may prolong the duration of postvaccination fever, due to routine use or overuse. Postvaccination body temperature observation and comparison with postvaccination fever patterns indicated here may reduce the unnecessary use of antipyretics.

Acknowledgments

This work was supported by Technology Innovation Program (20002289) funded by the Ministry of Trade, Industry, and Energy (MOTIE, KOREA) and the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2017R1D1A1B03035762).

Authors' Contributions

Study concept and design: all authors; data acquisition: SC, JS, and MK; data preprocessing: SHA and JZ; statistical analysis and interpretation: SHA, JZ, HK, YL, JHL, YRP; interpretation of the results: SHA, JS, JHL, YRP; crafting of the manuscript: SHA, JZ, YRP; and study supervision: JS, JHL, YRP.

Conflicts of Interest

Seyun Chang, Jaewon Shin, and Myeongchan Kim are employees of Mobile Doctor Co, Ltd. All other authors declare no conflicts of interest.

Multimedia Appendix 1

Screenshots of Fever Coach application translated into English.

[PDF File (Adobe PDF File), 1MB - mhealth_v7i4e12223_app1.pdf]

Multimedia Appendix 2

Comparison of onset, offset, and duration times and maximum temperatures among vaccine types and effects of antipyretics on postvaccination fever response.

[DOCX File, 15KB - mhealth_v7i4e12223_app2.docx]

References

1. World Health Organization. WHO recommendations for routine immunization-summary tables URL: https://www.who.int/immunization/policy/immunization_tables/en/ [accessed 2019-03-26] [WebCite Cache ID 77Au79Meh]
2. Kelso JM, Greenhawt MJ, Li JT, Nicklas RA, Bernstein DI, Blessing-Moore J, et al. Adverse reactions to vaccines practice parameter 2012 update. J Allergy Clin Immunol 2012 Jul;130(1):25-43. [doi: 10.1016/j.jaci.2012.04.003] [Medline: 22608573]
3. Zhou W, Pool V, Iskander JK, English-Bullard R, Ball R, Wise RP, et al. Surveillance for safety after immunization: Vaccine Adverse Event Reporting System (VAERS)--United States, 1991-2001. MMWR Surveill Summ 2003 Jan 24;52(1):1-24 [FREE Full text] [Medline: 12825543]
4. Jampel HD, Duff GW, Gershon RK, Atkins E, Durum SK. Fever and immunoregulation. III. Hyperthermia augments the primary in vitro humoral immune response. J Exp Med 1983 Apr 01;157(4):1229-1238 [FREE Full text] [Medline: 6220108]
5. Kluger MJ, Kozak W, Leon LR, Soszynski D, Conn CA. Cytokines and fever. Neuroimmunomodulation 1995 Mar;2(4):216-223. [doi: 10.1159/000097199] [Medline: 8963750]
6. Evans SS, Repasky EA, Fisher DT. Fever and the thermal regulation of immunity: the immune system feels the heat. Nat Rev Immunol 2015 Jun;15(6):335-349 [FREE Full text] [doi: 10.1038/nri3843] [Medline: 25976513]
7. Jackson LA, Peterson D, Dunn J, Hambidge SJ, Dunstan M, Starkovich P, et al. A randomized placebo-controlled trial of acetaminophen for prevention of post-vaccination fever in infants. PLoS One 2011 Jun;6(6):e20102 [FREE Full text] [doi: 10.1371/journal.pone.0020102] [Medline: 21698100]
8. Tichmann-Schumann I, Soemantri P, Behre U, Disselhoff J, Mahler H, Maechler G, et al. Immunogenicity and reactogenicity of four doses of diphtheria-tetanus-three-component acellular pertussis-hepatitis B-inactivated polio virus-Haemophilus influenzae type b vaccine coadministered with 7-valent pneumococcal conjugate Vaccine. Pediatr Infect Dis J 2005 Jan;24(1):70-77. [Medline: 15665713]
9. Black S, Shinefield H, Fireman B, Lewis E, Ray P, Hansen JR, et al. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group. Pediatr Infect Dis J 2000 Mar;19(3):187-195. [Medline: 10749457]
10. Jackson LA, Carste BA, Malais D, Froeschle J. Retrospective population-based assessment of medically attended injection site reactions, seizures, allergic responses and febrile episodes after acellular pertussis vaccine combined with diphtheria and tetanus toxoids. Pediatr Infect Dis J 2002 Aug;21(8):781-786. [doi: 10.1097/01.inf.0000024003.77163.3d] [Medline: 12192169]
11. de Martino M, Chiarugi A. Recent advances in pediatric use of oral paracetamol in fever and pain management. Pain Ther 2015 Dec;4(2):149-168 [FREE Full text] [doi: 10.1007/s40122-015-0040-z] [Medline: 26518691]
12. Tapiainen T, Heininger U. Fever following immunization. Expert Rev Vaccines 2005 Jun;4(3):419-427. [doi: 10.1586/14760584.4.3.419] [Medline: 16026253]
13. Crocetti M, Moghbeli N, Serwint J. Fever phobia revisited: have parental misconceptions about fever changed in 20 years? Pediatrics 2001 Jun 01;107(6):1241-1246. [doi: 10.1542/peds.107.6.1241]
14. Gunduz S, Usak E, Koksal T, Canbal M. Why fever phobia is still common? Iran Red Crescent Med J 2016 Aug;18(8):e23827 [FREE Full text] [doi: 10.5812/ircmj.23827] [Medline: 27781110]

15. Kwak YH, Kim DK, Jang HY, Kim JJ, Ryu J, Oh SB, et al. Fever phobia in Korean caregivers and its clinical implications. J Korean Med Sci 2013 Nov;28(11):1639-1644 [FREE Full text] [doi: 10.3346/jkms.2013.28.11.1639] [Medline: 24265528]
16. Smith LE, Webster RK, Weinman J, Amlôt R, Yiend J, Rubin GJ. Psychological factors associated with uptake of the childhood influenza vaccine and perception of post-vaccination side-effects: A cross-sectional survey in England. Vaccine 2017 Dec 04;35(15):1936-1945 [FREE Full text] [doi: 10.1016/j.vaccine.2017.02.031] [Medline: 28284680]
17. Zhang L, Prietsch SO, Axelsson I, Halperin SA. Acellular vaccines for preventing whooping cough in children. Cochrane Database Syst Rev 2014 Sep 17(9):CD001478. [doi: 10.1002/14651858.CD001478.pub6] [Medline: 25228233]
18. Stockwell MS, Broder K, LaRussa P, Lewis P, Fernandez N, Sharma D, et al. Risk of fever after pediatric trivalent inactivated influenza vaccine and 13-valent pneumococcal conjugate vaccine. JAMA Pediatr 2014 Mar;168(3):211-219. [doi: 10.1001/jamapediatrics.2013.4469] [Medline: 24395025]
19. Bryant KA, Block SL, Baker SA, Gruber WC, Scott DA, PCV13 Infant Study Group. Safety and immunogenicity of a 13-valent pneumococcal conjugate vaccine. Pediatrics 2010 May;125(5):866-875. [doi: 10.1542/peds.2009-1405] [Medline: 20435707]
20. Cunha BA. The clinical significance of fever patterns. Infect Dis Clin North Am 1996 Mar;10(1):33-44. [Medline: 8698993]
21. Ogoina D. Fever, fever patterns and diseases called 'fever'--a review. J Infect Public Health 2011 Aug;4(3):108-124 [FREE Full text] [doi: 10.1016/j.jiph.2011.05.002] [Medline: 21843857]
22. Kirubakaran R, Viswanathan A, Kompithra R. Prophylactic paracetamol for the prevention of fever in children receiving vaccination as part of a standard childhood immunization schedule. Cochrane Database Syst Rev 2017 May 05(5):CD012655. [doi: 10.1002/14651858.CD012655]
23. Monfries N, Goldman R. Prophylactic antipyretics for prevention of febrile seizures following vaccination. Can Fam Physician 2017 Feb;63(2):128-130 [FREE Full text] [Medline: 28209678]
24. Yeh SH, Gurtman A, Hurley DC, Block SL, Schwartz RH, Patterson S, 004 Study Group. Immunogenicity and safety of 13-valent pneumococcal conjugate vaccine in infants and toddlers. Pediatrics 2010 Sep;126(3):e493-e505. [doi: 10.1542/peds.2009-3027] [Medline: 20732948]
25. Aronoff DM, Neilson EG. Antipyretics: mechanisms of action and clinical use in fever suppression. Am J Med 2001 Sep;111(4):304-315. [Medline: 11566461]
26. Kumar RB, Goren ND, Stark DE, Wall DP, Longhurst CA. Automated integration of continuous glucose monitor data in the electronic health record using consumer technology. J Am Med Inform Assoc 2016 May;23(3):532-537 [FREE Full text] [doi: 10.1093/jamia/ocv206] [Medline: 27018263]
27. Jang I, Kim H, Lee E, Jung H, Park H, Cheon S, et al. Impact of a wearable device-based walking programs in rural older adults on physical activity and health outcomes: cohort study. JMIR Mhealth Uhealth 2018 Nov 21;6(11):e11335 [FREE Full text] [doi: 10.2196/11335] [Medline: 30463838]
28. Kim J, Kam HJ, Park YR, Yoo S, Oh JS, Kim Y, et al. Enchanted life space: adding value to smart health by integrating human desires. Healthc Inform Res 2018 Jan;24(1):3-11 [FREE Full text] [doi: 10.4258/hir.2018.24.1.3] [Medline: 29503747]
29. Park YR, Lee Y, Kim JY, Kim J, Kim HR, Kim Y, et al. Managing patient-generated health data through mobile personal health records: analysis of usage data. JMIR Mhealth Uhealth 2018 Apr 09;6(4):e89 [FREE Full text] [doi: 10.2196/mhealth.9620] [Medline: 29631989]
30. Walker DM, Sieck CJ, Menser T, Huerta TR, Scheck MA. Information technology to support patient engagement: where do we stand and where can we go? J Am Med Inform Assoc 2017 Nov 01;24(6):1088-1094. [doi: 10.1093/jamia/ocx043] [Medline: 28460042]
31. Torkamani A, Andersen KG, Steinhubl SR, Topol EJ. High-definition medicine. Cell 2017 Aug 24;170(5):828-843 [FREE Full text] [doi: 10.1016/j.cell.2017.08.007] [Medline: 28841416]
32. Cohen D, Keller S, Hayes G, Dorr D, Ash J, Sittig D. Integrating patient-generated health data into clinical care settings or clinical decision-making: lessons learned from project health and design. JMIR Hum Factors 2016 Oct 19;3(2):e26 [FREE Full text] [doi: 10.2196/humanfactors.5919] [Medline: 27760726]
33. Accenture Federal Services. Conceptualizing a Data Infrastructure for the Capture, Use, and Sharing of Patient-Generated Health Data in Care Delivery and Research through 2024: Draft White Paper for a PGHD Policy Framework. Washington, DC: Office of the National Coordinator for Health Information Technology; 2018 Jan. URL: https://www.healthit.gov/sites/default/files/onc_pghd_practical_guide.pdf [accessed 2019-04-01] [WebCite Cache ID 77JCtKSMs]
34. World Health Organization. Supplementary Information on Vaccine Safety. Part 2: Background Rates of Adverse Events Following Immunization. Geneva: World Health Organization; 2000. URL: https://apps.who.int/iris/handle/10665/66675 [accessed 2019-03-26] [WebCite Cache ID 77Ay9nYu1]
35. National Institute for Health and Care Excellence. 2013. Fever in under 5s: assessment and initial management URL: https://www.nice.org.uk/guidance/CG160 [accessed 2019-03-26] [WebCite Cache ID 77AyHFksp]
36. Craig JV, Lancaster GA, Taylor S, Williamson PR, Smyth RL. Infrared ear thermometry compared with rectal thermometry in children: a systematic review. Lancet 2002 Aug 24;360(9333):603-609. [doi: 10.1016/S0140-6736(02)09783-0] [Medline: 12241932]
37. Centers for Disease Control and Prevention. 2018. Ingredients of vaccines-fact sheet URL: https://www.cdc.gov/vaccines/vac-gen/additives.htm [accessed 2019-03-26] [WebCite Cache ID 77AyNhktg]

38. Harris G, Nolan T, Hartman L. Clinical signs associated with Triple Antigen (DTP) vaccination in infants. J Paediatr Child Health 1995 Jun;31(3):228-232. [Medline: 7669385]
39. Prymula R, Siegrist C, Chlibek R, Zemlickova H, Vackova M, Smetana J, et al. Effect of prophylactic paracetamol administration at time of vaccination on febrile reactions and antibody responses in children: two open-label, randomised controlled trials. Lancet 2009 Oct 17;374(9698):1339-1350. [doi: 10.1016/S0140-6736(09)61208-3] [Medline: 19837254]
40. Allotey P, Reidpath DD, Elisha D. "Social medication" and the control of children: a qualitative study of over-the-counter medication among Australian children. Pediatrics 2004 Sep;114(3):e378-e383. [doi: 10.1542/peds.2004-0759] [Medline: 15342901]
41. Enarson M, Ali S, Vandermeer B, Wright R, Klassen T, Spiers J. Beliefs and expectations of Canadian parents who bring febrile children for medical care. Pediatrics 2012 Oct;130(4):e905-e912. [doi: 10.1542/peds.2011-2140] [Medline: 22966028]
42. Voysey M, Barker CI, Snape MD, Kelly DF, Trück J, Pollard AJ. Sex-dependent immune responses to infant vaccination: an individual participant data meta-analysis of antibody and memory B cells. Vaccine 2016 Mar 29;34(14):1657-1664. [doi: 10.1016/j.vaccine.2016.02.036] [Medline: 26920472]
43. Jørgensen TN. Sex disparities in the immune response. Cell Immunol 2015 Apr;294(2):61-62. [doi: 10.1016/j.cellimm.2015.02.001] [Medline: 25670393]
44. Jefferies S, Weatherall M, Young P, Eyers S, Perrin KG, Beasley CR. The effect of antipyretic medications on mortality in critically ill patients with infection: a systematic review and meta-analysis. Crit Care Resusc 2011 Jun;13(2):125-131. [Medline: 21627583]
45. Hudgings L, Kelsberg G, Safranek S, Neher JO. Clinical inquiries. Do antipyretics prolong febrile illness? J Fam Pract 2004 Jan;53(1):57-58, 61. [Medline: 14709270]
46. Saleh E, Swamy GK, Moody MA, Walter EB. Parental approach to the prevention and management of fever and pain following childhood immunizations: a survey study. Clin Pediatr (Phila) 2017 May;56(5):435-442 [FREE Full text] [doi: 10.1177/0009922816675116] [Medline: 27798399]
47. Rosenbloom E, Finkelstein Y, Adams-Webber T, Kozer E. Do antipyretics prevent the recurrence of febrile seizures in children? A systematic review of randomized controlled trials and meta-analysis. Eur J Paediatr Neurol 2013 Nov;17(6):585-588. [doi: 10.1016/j.ejpn.2013.04.008] [Medline: 23702315]
48. Das RR, Panigrahi I, Naik SS. The effect of prophylactic antipyretic administration on post-vaccination adverse reactions and antibody response in children: a systematic review. PLoS One 2014 Sep;9(9):e106629 [FREE Full text] [doi: 10.1371/journal.pone.0106629] [Medline: 25180516]
49. Falup-Pecurariu O, Man SC, Neamtu ML, Chicin G, Baciu G, Pitic C, et al. Effects of prophylactic ibuprofen and paracetamol administration on the immunogenicity and reactogenicity of the 10-valent pneumococcal non-typeable Haemophilus influenzae protein D conjugated vaccine (PHiD-CV) co-administered with DTPa-combined vaccines in children: An open-label, randomized, controlled, non-inferiority trial. Hum Vaccin Immunother 2017 Dec 04;13(3):649-660 [FREE Full text] [doi: 10.1080/21645515.2016.1223001] [Medline: 27541270]
50. Government of Canada, Public Health Agency of Canada. 2018. Canadian immunization guide URL: https://www.canada.ca/en/public-health/services/canadian-immunization-guide.html [accessed 2019-03-26] [WebCite Cache ID 77Ayl8KF9]
51. Ministry of Health. 2015. Tips following immunisation URL: https://www.health.govt.nz/your-health/healthy-living/immunisation/tips-following-immunisation [accessed 2019-03-26] [WebCite Cache ID 77AzWqUtO]
52. Section on Clinical Pharmacology and Therapeutics, Committee on Drugs, Sullivan JE, Farrar HC. Fever and antipyretic use in children. Pediatrics 2011 Mar;127(3):580-587. [doi: 10.1542/peds.2010-3852] [Medline: 21357332]
53. Mackowiak PA, Worden G. Carl Reinhold August Wunderlich and the evolution of clinical thermometry. Clin Infect Dis 1994 Mar;18(3):458-467. [Medline: 8011836]
54. Craig JV, Lancaster GA, Williamson PR, Smyth RL. Temperature measured at the axilla compared with rectum in children and young people: systematic review. BMJ 2000 Apr 29;320(7243):1174-1178 [FREE Full text] [Medline: 10784539]
55. Leick-Rude MK, Bloom LF. A comparison of temperature-taking methods in neonates. Neonatal Netw 1998 Aug;17(5):21-37. [Medline: 9791443]
56. Berksoy E, Anıl M, Bıcılıoğlu Y, Gökalp G, Bal A. Comparison of infrared tympanic, non-contact infrared skin, and axillary thermometer to rectal temperature measurements in a pediatric emergency observation unit. Int J Clin Exp Med 2018;11(2):567-573 [FREE Full text]
57. Teller J, Ragazzi M, Simonetti GD, Lava SA. Accuracy of tympanic and forehead thermometers in private paediatric practice. Acta Paediatr 2014 Feb;103(2):e80-e83. [doi: 10.1111/apa.12464] [Medline: 24127699]
58. Brown RD, Kearns G, Eichler VF, Wilson JT. A probability nomogram to predict rectal temperature in children. Clin Pediatr (Phila) 1992 Sep;31(9):523-531. [doi: 10.1177/000992289203100902] [Medline: 1468168]
59. Centers for Disease Control and Prevention. 2018. Multiple vaccines and the immune system URL: https://www.cdc.gov/vaccinesafety/concerns/multiple-vaccines-immunity.html [accessed 2019-03-26] [WebCite Cache ID 77Av3jAdS]

Abbreviations

BCG: bacillus Calmette-Guérin

CI: confidence interval
DTaP: diphtheria, tetanus, and pertussis
HIB: Haemophilus influenzae type B
IQR: interquartile range
MMR: measles-mumps-rubella
PGHD: patient-generated health data
VIF: variance inflation factor

Edited by G Eysenbach; submitted 15.09.18; peer-reviewed by D Walker, IY Jang, TH Kwan; comments to author 11.11.18; revised version received 05.01.19; accepted 11.02.19; published 22.04.19.

Please cite as:
Ahn SH, Zhiang J, Kim H, Chang S, Shin J, Kim M, Lee Y, Lee JH, Park YR
Postvaccination Fever Response Rates in Children Derived Using the Fever Coach Mobile App: A Retrospective Observational Study
JMIR Mhealth Uhealth 2019;7(4):e12223
URL: http://mhealth.jmir.org/2019/4/e12223/
doi:10.2196/12223
PMID:31008712

Original Paper

Image-Based Mobile System for Dietary Management in an American Cardiology Population: Pilot Randomized Controlled Trial to Assess the Efficacy of Dietary Coaching Delivered via a Smartphone App Versus Traditional Counseling

Brian G Choi[1], MD, MBA, FACC; Tania Dhawan[1], MD; Kelli Metzger[1], MS, RDN, CDE; Lorraine Marshall[1], RN, BSN; Awdah Akbar[1], MBBS; Tushina Jain[1], MD; Heather A Young[2], PhD, MPH; Richard J Katz[1], MD

[1]Department of Medicine, School of Medicine & Health Sciences, The George Washington University, Washington, DC, United States

[2]Department of Epidemiology and Biostatistics, Milken Institute School of Public Health, The George Washington University, Washington, DC, United States

Corresponding Author:
Brian G Choi, MD, MBA, FACC
Department of Medicine
School of Medicine & Health Sciences
The George Washington University
2150 Pennsylvania Ave NW
Suite 4-417
Washington, DC, 20037
United States
Phone: 1 2027412323
Fax: 1 2027412324
Email: bchoi@mfa.gwu.edu

Abstract

Background: Randomized controlled trials conducted in Mediterranean countries have shown that the Mediterranean diet lowers adverse cardiovascular events. In the American population, diet remains the biggest uncontrolled risk factor for cardiovascular disease.

Objective: This study aimed to test the hypothesis that asynchronous dietary counseling supplied through a custom smartphone app results in better adherence to a Mediterranean diet in a non-Mediterranean population than traditional standard-of-care (SOC) counseling.

Methods: In total, 100 patients presenting to the cardiology clinic of an academic medical center were randomized to either the SOC or smartphone app-based experimental (EXP) Mediterranean diet intervention after informed consent and 1 hour of individual face-to-face dietary counseling with a registered dietitian. Participants in EXP received a custom smartphone app that reinforced the Mediterranean diet, whereas participants in SOC received 2 additional sessions of in-person dietary counseling with the registered dietitian—30 min at 1 month and 30 min at 3 months. Preexisting knowledge of a Mediterranean diet was measured by the validated Mediterranean Diet Score (MDS) instrument. Baseline height, weight, blood pressure (BP), and laboratory biomarkers were collected. At 1, 3, and 6 months, participants presented for a follow-up appointment to assess compliance to the Mediterranean diet using the MDS as well as a patient satisfaction survey, BP, and weight. Repeat laboratory biomarkers were performed at 3 and 6 months.

Results: Enrolled participants had a mean age with SE of 56.6 (SD 1.7) for SOC and 57.2 (SD 1.8) for EXP; 65.3% of SOC and 56.9% of EXP were male, and 20.4% of SOC and 35.3% of EXP had coronary artery disease. There were no significant differences between EXP and SOC with regard to BP, lipid parameters, hemoglobin A_{1c}, or C-reactive protein (CRP). Participants in EXP achieved a significantly greater weight loss on average of 3.3 pounds versus 3.1 pounds for participants in SOC, P=.04. Adherence to the Mediterranean diet increased significantly over time for both groups (P<.001), but there was no significant difference between groups (P=.69). Similarly, there was no significant difference in diet satisfaction between EXP and SOC, although diet satisfaction increased significantly over time for both groups. The proportion of participants with high Mediterranean

diet compliance (defined as the MDS ≥9) increased significantly over time (P<.001)—from 18.4% to 57.1% for SOC and 27.5% to 64.7% for EXP; however, there was no significant difference between the groups.

Conclusions: Both traditional SOC counseling and smartphone-based counseling were effective in getting participants to adhere to a Mediterranean diet, and these dietary changes persisted even after counseling had ended. However, neither method was more effective than the other. This pilot study demonstrates that patients can change to and maintain a Mediterranean diet with either traditional or smartphone app-based nutrition counseling.

Trial Registration: ClinicalTrials.gov NCT03897426;https://clinicaltrials.gov/ct2/show/NCT03897426

(JMIR Mhealth Uhealth 2019;7(4):e10755) doi:10.2196/10755

KEYWORDS

Mediterranean diet; telemedicine; cardiovascular disease; randomized controlled trial

Introduction

Cardiovascular disease (CVD) has been the leading cause of death every year since 1918 in the United States and currently accounts for 1 in every 3 deaths [1]. Diet is the most uncontrolled risk factor for cardiovascular health among US adults; per the 2015-2020 Dietary Guidelines for Americans, approximately 75% of the population follows an unhealthy eating pattern and less than 2% of Americans have an ideal diet [2]. These suboptimal dietary habits are the leading cause of mortality and disability-adjusted life years lost, greater than smoking, obesity, physical inactivity, high cholesterol, hypertension, or diabetes [3].

Randomized controlled trials (RCTs) have shown superiority of the Mediterranean diet in both high-risk primary prevention and secondary prevention populations at risk for CVD [4]. The Mediterranean diet is typified by olive oil, nuts, vegetables, legumes, fish, white meat, and wine, and the Prevención con Dieta Mediterránea (PREDIMED) study showed a 30% reduction in major adverse cardiovascular events for patients at high risk of developing atherosclerotic disease when they received a Mediterranean diet enriched with olive oil and nut consumption [5]. The Lyon Diet Heart Study enrolled patients after a first myocardial infarction, and those randomized to receiving a Mediterranean diet had improved survival and fewer myocardial infarctions than those on a usual prudent diet [6]. These studies, however, were conducted in Mediterranean countries and whether this diet can be achieved in non-Mediterranean populations with similar benefits remains to be elucidated [7]. If a Mediterranean diet can be successfully implemented and maintained in a high-risk American population, substantial reductions in cardiovascular morbidity and mortality may follow [8].

Our study tested the hypothesis that experimental asynchronous dietary counseling provided through a custom smartphone app (EXP) would improve implementation and compliance to a Mediterranean diet and secondarily evaluated the effect of this intervention on markers of cardiovascular risk and patient satisfaction compared with standard-of-care (SOC) counseling, that is, face-to-face counseling from a registered dietitian (RD).

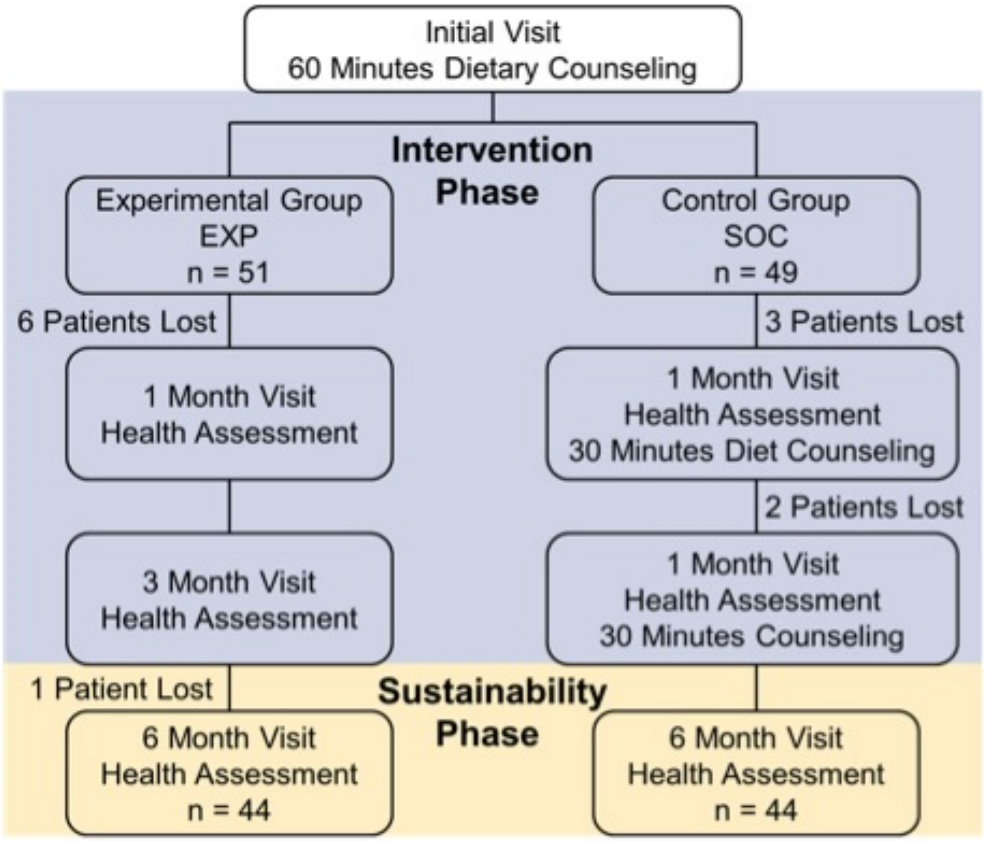
Methods

Recruitment

A total of 100 patients were recruited from the cardiology clinic of an academic medical center in Washington, DC, for this RCT (Figure 1) as cardiology patients may be the ones most likely to benefit from dietary modification to the Mediterranean diet.

All patients scheduled for the clinic were screened via query of the electronic medical record system to see if those potential study participants would meet the eligibility criteria, recruitment posters were posted in exam rooms of the clinic, and those candidates meeting the eligibility criteria were further screened by their cardiologist to see if they would qualify. Potential candidates were approached for participation by a research team member at the time of their clinic visit in coordination with the participant's cardiologist. The inclusion criteria included ongoing cardiology care with a cardiologist that is expected to continue for at least 6 months (to limit the potential for participants in being lost in follow-up), personal ownership of an Android- or iOS (iPhone operating system)-based smartphone with a data plan, English language proficiency, demonstration of the ability to download and install the app, aged at least 18 years, and a minimum fifth grade literacy level per the Rapid Estimate of Adult Literacy in Medicine [9]. The exclusion criteria were clinical instability at the time of enrollment; comorbid medical disease that would preclude the ability to participate in a nutrition intervention study (eg, digestive disease with fat intolerance); life expectancy less than 5 years; severe neurologic, psychiatric, or endocrine abnormalities; immunodeficiency or HIV-positive status; illegal drug use; alcoholism or daily alcohol intake >80 grams/day (ie, 5 12-ounce glasses of beer, 5 5-ounce glasses of wine, or 5 1.5-ounce glasses of spirits); body mass index (BMI) >40 kg/m^2 (as these patients should be on a weight-reducing diet instead); inability or unwillingness to change dietary habits per patient report; inability or unwillingness to adhere to a Mediterranean diet (eg, religious or moral reasons); disorders of chewing or swallowing; allergies to major components of the Mediterranean diet (eg, olive oil and nuts); participation in any drug trial or use of any investigational drug within the past month; institutionalized or nonambulatory status; lack of autonomy; no stable address; acute infection or inflammation (patients could be reconsidered when/if the condition resolves); and patients currently pregnant, breastfeeding, or planning to become pregnant.

Figure 1. Randomized controlled trial design. Of the patients that dropped out of the study, 6 participants (4 standard-of-care, SOC and 2 experimental arm, EXP) did not show up to scheduled appointments, 1 participant (EXP) withdrew from the study because the food was "too expensive," 3 participants (all EXP) withdrew because the app failed to work on their phone, and 2 participants (1 SOC and 1 EXP) were no longer interested in participating in the study and did not give a clarifying reason.



Those patients who met the enrollment criteria were invited to participate by a study coordinator. Patients who wished to enroll signed the informed consent and then received 1 hour of face-to-face counseling on the Mediterranean diet from an RD. During the initial visit, the RD reviewed the rationale of the Mediterranean diet, detailed the expected diet and lifestyle modifications involved, and developed a contract for negotiated change. Participants were provided with shopping lists, weekly plans, and recipes, and the RD administered a quantitative questionnaire of Mediterranean diet compliance, the Mediterranean Diet Score (MDS). Initial blood pressure (BP), weight, total cholesterol, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, triglycerides, hemoglobin A_{1c} (HbA_{1c}), and high-sensitivity C-reactive protein (CRP) were obtained. At the end of the visit, the participant was randomized to either the app-based intervention arm (EXP) or SOC based upon a predetermined random sequence. Participants randomized to the intervention arm were given an instruction manual, directed to a website for additional information on how to use the app, and had the smartphone app set up on their mobile device to establish connectivity with the RD. As reimbursement, patients in both treatment groups received a small stipend to cover the cost of travel for each visit as well as a nominal gift card to a local grocery store that carried foods included in the Mediterranean diet.

Standard-of-Care Arm

Participants could initiate contact and receive brief telephone-based counseling with their assigned RD at any time during the study; however, the RD would not initiate additional face-to-face dietary counseling until the 1-month and 3-month visit. At 1 month, participants returned for a 30-min follow-up counseling session with the RD during which dietary recall was reviewed, education reinforced, and strategies for improvement developed. Repeat BP and weight measurements were obtained. The MDS was readministered, interval medical history events were reviewed, and satisfaction with SOC and the Mediterranean diet was assessed.

At 3 months, the RD administered a final 30 min of dietary counseling; dietary recall was again reviewed, education reinforced, and strategies for improvement further developed. Repeat BP, weight, fasting total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides, HbA_{1c}, and high-sensitivity CRP were also obtained during the 3-month visit. The MDS was readministered, interval medical history events were again reviewed, and satisfaction with SOC and the Mediterranean diet was reassessed.

The participant then entered a sustainability phase during which the RD could not initiate contact. At 6 months, the participant had a final visit with a research coordinator. Repeat BP, weight, fasting total cholesterol, HDL cholesterol, LDL cholesterol,

triglycerides, HbA_{1c}, and high-sensitivity CRP were obtained. The MDS was readministered. Interval medical history events were reviewed. Satisfaction with SOC and the Mediterranean diet was assessed. Participants were also asked if they received any additional dietary counseling not supported through the study.

Intervention Arm

Participants in the intervention arm (EXP) were allotted 60 min with the RD through the customized smartphone app (Figure 2) and encouraged to fully utilize these 60 min in the 3 months following enrollment. The app was created by Vibrent Health (Fairfax). Vibrent Health consulted with the research team on the app's design, and Vibrent Health was free to modify the app throughout the course of the study to load new educational content provided by the research team, refresh challenges, and maintain compatibility with mobile phone operating systems and new smartphone designs. Only time spent by the RD interacting with the participant through the app counted against the allotted time. At the initial visit, the participants received the same educational handouts provided to the SOC arm, but also had this content preloaded onto the app. The app included weekly challenges to encourage dietary modification, and either the RD or the patient could initiate contact through the app. Participants were encouraged to use the app to take pictures of their food, document meals and amounts consumed, ask questions to the RD, document exercise, and monitor their BP and track it in the log if another provider recommended that it be recorded. The challenges were for patients to challenge themselves (ie, they did not compete against other participants). Examples of challenges included increasing daily servings of vegetables or exercising daily. Additional face-to-face counseling was not offered by the RD. The allotted time that was not used in the first 3 months could be used after 3 months, but no counseling was provided after 6 months. If the participants used all of their allotted time, they could continue to initiate contact with their RD through traditional means (eg, by telephone).

Participants in EXP received dietary counseling from their RD via the app in the context of their meal logs (Figure 3). The amount of time the RD spent responding to a participant through the provider portal was tracked to ensure it fell within the 60 min limit. Participants received the RD's feedback directly to their mobile phone to review at their own convenience. After the 6-month visit, the app no longer had RD access and support was not continued.

Similar to the SOC group, participants in EXP had a follow-up visit with a research coordinator at 1 month, 3 months, and 6 months following enrollment. At each visit, interval medical history events were reviewed, the MDS was readministered, and repeat BP and weight obtained. Participant satisfaction with the intervention and Mediterranean diet was also assessed at each visit. Fasting total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides, HbA_{1c}, and high-sensitivity CRP were measured at the 3-month and 6-month visits.

The MDS, the 14-point quantitative score of adherence to the Mediterranean diet, was the primary endpoint of this RCT. Secondary endpoints included BP, weight, fasting lipid parameters (total cholesterol, HDL cholesterol, LDL cholesterol, and triglycerides), HbA_{1c}, high-sensitivity CRP, participant satisfaction with the intervention versus usual care, and percent of participants achieving high compliance with a Mediterranean diet (defined as MDS ≥9).

Figure 2. App-based diet education and tracking to encourage Mediterranean diet compliance. From left to right, the first image illustrates the app-based learning material on the Mediterranean diet, the second image shows a drop-down menu of self-assessment tools offered by the app, and the third displays one of the self-assessment tools—a Mediterranean diet log to record daily compliance.

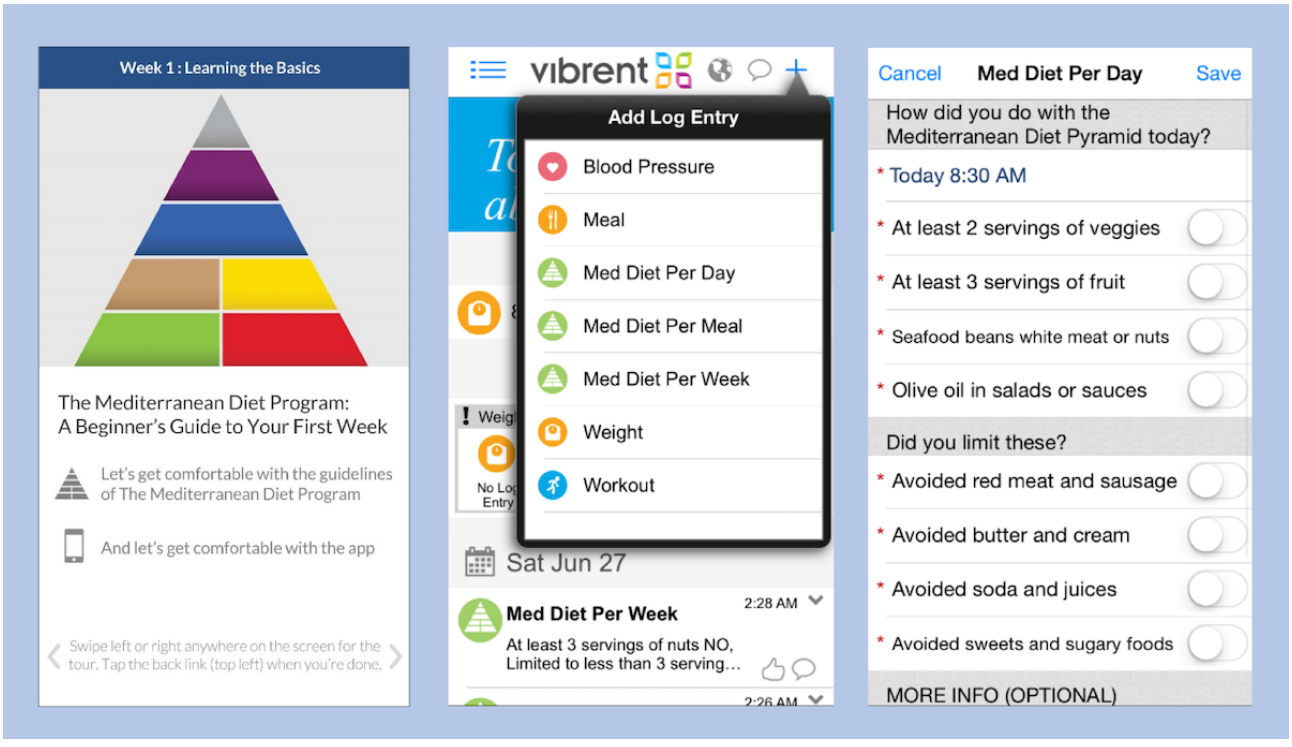
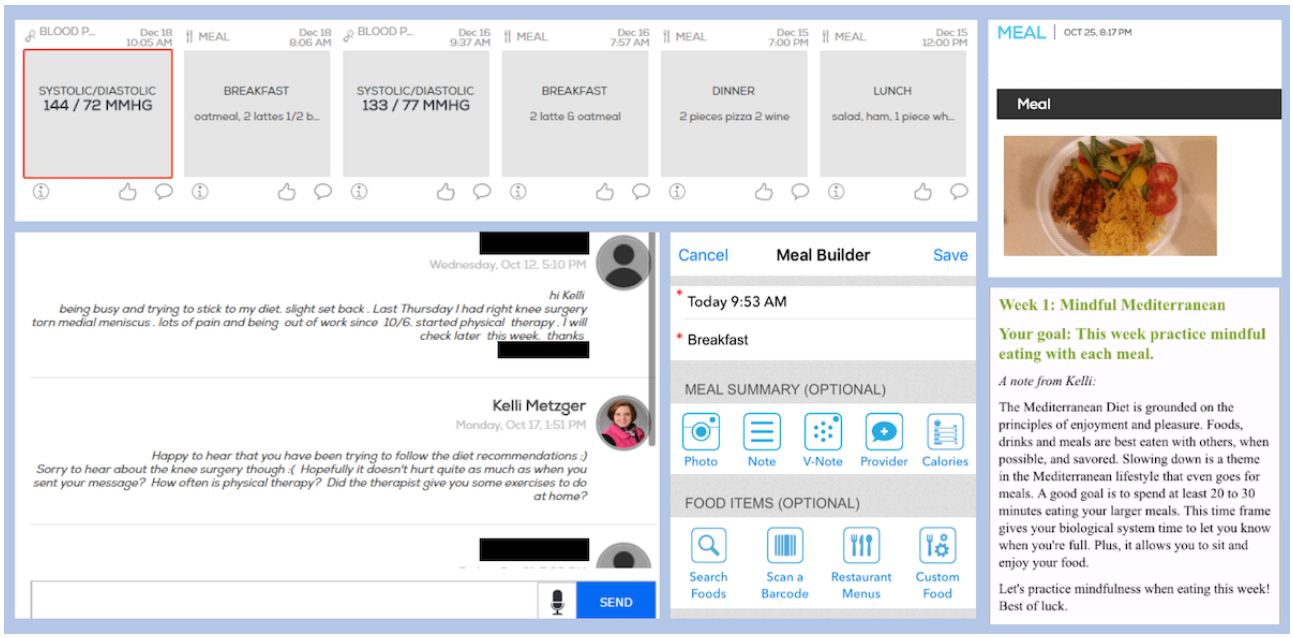


Figure 3. Asynchronous counseling and meal-logging tools of the customized smartphone app. From left to right, in the top row, the first image shows the app's timeline view of various logs (eg, blood pressure and meals) and the second image represents the app's photo log feature via the meal a patient photographed; in the bottom row, the first image illustrates the registered dietitian (RD)-patient interface for app-based nutritional counseling, the second depicts the app's Meal Builder feature with its meal- and food-logging options, and the third exemplifies the app's weekly challenge feature via a sample counseling note sent by the RD through the app.



Statistical Analysis

The sample size was determined using the study's hypotheses and primary endpoint of the MDS. Given an estimated baseline MDS of 4.5 with an SD of 2 units, a total sample size of 100 (n=50 per group) would have adequate power (93%) to detect a difference of 1.4 units in the MDS between the SOC and intervention arm. Even after an assumed attrition rate of 20% overall (total n=80), the study would retain a power of at least 80% to detect a minimum difference of 1.4 units between SOC and EXP.

Descriptive statistics of baseline characteristics plus primary and secondary endpoints at baseline and certain time points were calculated for both groups, and the 2 groups were compared using the chi-square test, Fisher exact test, Wilcoxon signed rank, or 2-sample test, as appropriate. The descriptive statistics for app usage data were also summarized for EXP participants who used each portion of the app.

Repeated measures analysis of variance was used to compare primary and secondary outcomes over time and across groups. Wilcoxon signed rank test was used to assess differences in the participant satisfaction score (PSS) between SOC and EXP at individual time points. As a secondary endpoint, the study also compared the proportion of participants that achieved a high compliance (MDS ≥9) with the Mediterranean diet between the 2 groups; McNemar's chi-square test was used to determine if there were differences between the 2 groups from baseline to last visit.

Differences in MDS from baseline to last visit were calculated and compared between SOC and EXP within subgroups of sex, diabetes, alcohol use, smoking, and atherosclerotic cardiovascular disease (ASCVD) using analysis of variances. App usage data were broken into quartiles based on total app usage time (0 to 8.5 min; 8.5 to 16 min; 16 to 30 min; or more than 30 min); high (>16 min) and low (≤16 min) groups for RD usage; and dichotomous (yes/no) groups for challenge, weight log, workout log, and BP log usage. Repeated measures analysis of variance was used to examine differences in the MDS across app usage groups and over time. Chi-square was employed in examining differences in proportions of weight loss, HDL increase, and lowered systolic BP across app usage groups.

Dropouts were analyzed using intention-to-treat methodology so they would be analyzed in the group they started in up to the period at which they dropped out. Differences in percentages of dropouts were also examined to determine if outcomes differed across groups.

The study was reviewed and approved by the institutional review board. At the time this study was started, trial registration was not required at Clinicaltrials.gov but has since been registered by the study sponsor Vibrent Health (Unique Protocol ID HHSN261201400053C_SBIR 308).

Results

The mean age was 56.6 (SE 1.7) years for SOC and 57.2 (SE 1.8) years for EXP. There was a significantly higher number of participants with diabetes at baseline in SOC (P=.01). For the other baseline characteristics, there was no significant difference between SOC and EXP (Tables 1 and 2). There were 12 dropouts over the course of the study (Figure 1).

Table 1. Baseline characteristics of the study population.

Variable	Standard-of-care arm, n (%)	Experimental arm, n (%)
Gender		
Female	17 (34.7)	22 (43.1)
Male	32 (65.3)	29 (56.9)
Medical history		
Diabetes[a]	10 (20.4)	2 (3.9)
Atherosclerotic CVD[b]	10 (20.4)	18 (35.3)
Previous myocardial infarction	8 (16.3)	10 (19.6)
Previous revascularization	12 (24.5)	18 (35.3)
Previous stroke or transient ischemic attack	2 (4.1)	3 (5.9)
Peripheral vascular disease	1 (2.0)	1 (2.0)
Other CVD	39 (92.9)	38 (90.5)
Alcohol use		
Previous	1 (2.0)	2 (4.1)
Current	32 (68.1)	39 (79.6)
Tobacco use		
Never	29 (59.2)	32 (62.8)
Previous	13 (26.5)	14 (27.5)
Current	4 (8.2)	2 (4.0)

[a]*P*=.01.

[b]CVD: cardiovascular disease.

Table 2. Mean and SE for outcome variables by treatment group and time of visit.

Group	Initial, mean (SE)	1 month, mean (SE)	3 month, mean (SE)	6 month, mean (SE)	$P_{overtime}$	P_{group}
Systolic BP[a] (mmHg)						
SOC[b]	128.1 (2.9)	125.4 (2.4)	130.3 (2.9)	128.7 (2.5)	.43	.34
EXP[c]	129.6 (2.3)	128.5 (2.2)	128.2 (2.2)	129.5 (2.5)	.43	.34
Diastolic BP (mmHg)						
SOC	78.2 (1.1)	77.8 (1.3)	79.5 (1.4)	79.2 (1.5)	.75	.27
EXP	78.3 (1.1)	79.2 (1.5)	77.8 (1.5)	78.7 (1.5)	.75	.27
Body mass index (kg/m^2)						
SOC	30.8 (0.6)	30.4 (0.7)	30.7 (0.7)	30.5 (0.7)	.02	.03
EXP	29.5 (0.6)	29.0 (0.6)	28.9 (0.6)	28.7 (0.7)	.02	.03
Total cholesterol (mg/dL)						
SOC	171.9 (5.4)	—[d]	167.1 (5.1)	172.0 (6.4)	.19	.90
EXP	161.6 (5.4)	—	157.3 (5.8)	159.3 (6.1)	.19	.90
Triglyceride (mg/dL)						
SOC	104.2 (8.0)	—	98.0 (6.5)	97.5 (7.3)	.29	.71
EXP	107.3 (9.2)	—	102.8 (8.4)	99.8 (8.5)	.29	.71
High-density lipoprotein (mg/dL)						
SOC	54.1 (2.4)	—	54.9 (2.7)	54.3 (2.5)	.40	.12
EXP	53.8 (2.4	—	52.9 (2.3)	55.3 (2.8)	.40	.12
Very low-density lipoprotein (mg/dL)						
SOC	22.9 (2.7)	—	19.5 (1.3)	19.5 (1.5)	.31	.71
EXP	21.4 (1.8)	—	20.5 (1.7)	19.9 (1.7)	.31	.71
Low-density lipoprotein (mg/dL)						
SOC	94.5 (5.3)	—	92.7 (4.7)	98.3 (5.4)	.35	.58
EXP	86.4 (4.2)	—	83.9 (4.6)	84.1 (4.9)	.35	.58
Hemoglobin A_{1c} (%)						
SOC	6.2 (0.2)	—	6.0 (0.1)	6.1 (0.1)	.94	.45
EXP	5.8 (0.1)	—	5.9 (0.1)	5.8 (0.1)	.94	.45
C-reactive protein (mg/dL)						
SOC	3.3 (0.9)	—	2.5 (0.4)	2.2 (0.4)	.34	.29
EXP	1.9 (0.3)	—	1.6 (0.2)	1.5 (0.3)	.34	.29
Diet Satisfaction Score (units)						
SOC	—	0.42 (0.04)	0.59 (0.05)	0.53 (0.05)	<.001	.29
EXP	—	0.36 (0.05)	0.46 (0.05)	0.50 (0.05)	<.001	.29

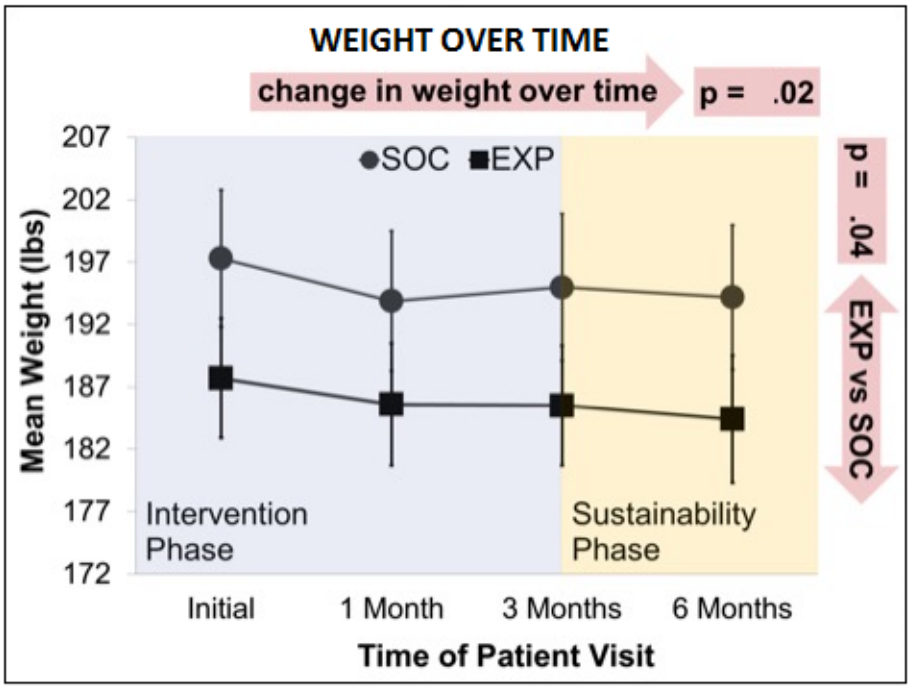
[a]BP: blood pressure.

[b]SOC: standard-of-care arm.

[c]EXP: experimental arm.

[d]Not applicable.

Figure 4. Change in weight over time (mean with SE). SOC: standard-of-care arm; EXP: experimental arm.



From baseline to final visit, there was a significant decrease in BMI (Table 2) and weight (Figure 4) that varied by group—participants in EXP experienced a slightly larger decrease in weight of 3.3 pounds versus 3.1 pounds (*P*=.04) and BMI of 0.8 versus 0.3 (*P*=.03). In addition, the MDS (Figure 5) and Diet Satisfaction Score (Table 2) increased significantly over time for both SOC and EXP, although there was no statistically significant difference in these trends between the 2 groups. Changes in the other outcome variables analyzed were not statistically significant over time or between the 2 treatment groups (Table 2).

There was a significant change in average PSS over time (Figure 6), but the change over time did not vary by group. However, at each time point, SOC had a significantly higher average patient satisfaction score compared with EXP (*P*=.01).

For both SOC and EXP, the proportion of participants that achieved high compliance (defined as MDS ≥9) with the Mediterranean diet increased significantly (McNemar's chi-square *P*<.0001) overtime from the initial visit to the 6-month visit (Figure 7). However, when comparing SOC with EXP at each time point, there were no significant differences in MDS ≥9 at any of the time points (initial: *P*=.33; 1 month: *P*=.50; 3 months: *P*=.44; and 6 months: *P*=.47). There was also no significant difference in any individual component of the MDS with time (at baseline or final visit) or between the 2 groups.

There was a difference in the MDS by ASCVD (*P*=.03). In both the SOC and experimental group, participants without ASCVD showed a larger increase in the MDS over the course of the study than participants with ASCVD (*P*=.02); however, this difference was not statistically significant between the SOC and experimental group. There were no differences in change in MDS from baseline to final visit by sex, diabetes, alcohol use, or smoking.

There were no significant differences in MDS over time by app usage quartiles, high versus low RD use, or challenge participation. There was no significant difference in weight loss between participants who used the weight log versus those who did not. However, of the 14 participants who used the workout log, 12 (86%) experienced weight loss from baseline to final visit. This was significantly higher (*P*=.04) than the 20 of 39 participants (54%) not using the workout log who experienced weight loss.

Of the 15 participants who used the BP log, 9 (60%) lowered their systolic BP from baseline. This was not significantly different from the 13 of 36 participants (38%) not using the BP log who lowered their systolic BP from baseline.

Figure 5. Compliance with the Mediterranean diet overtime (mean with SE). MDS: Mediterranean diet score; SOC: standard-of-care arm; EXP: experimental arm.

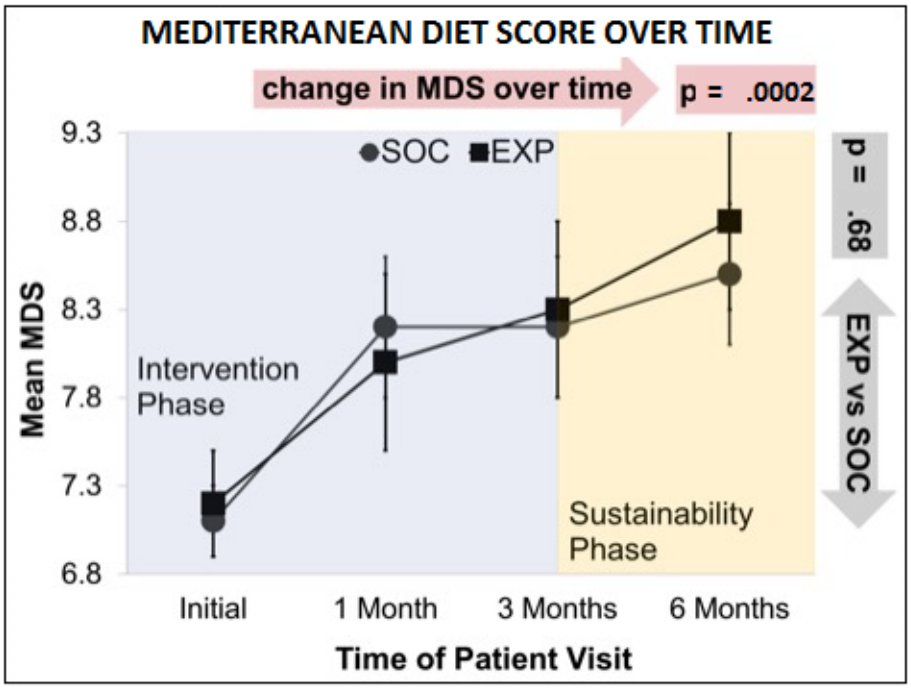


Figure 6. Patient satisfaction over time (mean with SE). PSS: patient satisfaction score; SOC: standard-of-care arm; EXP: experimental arm.

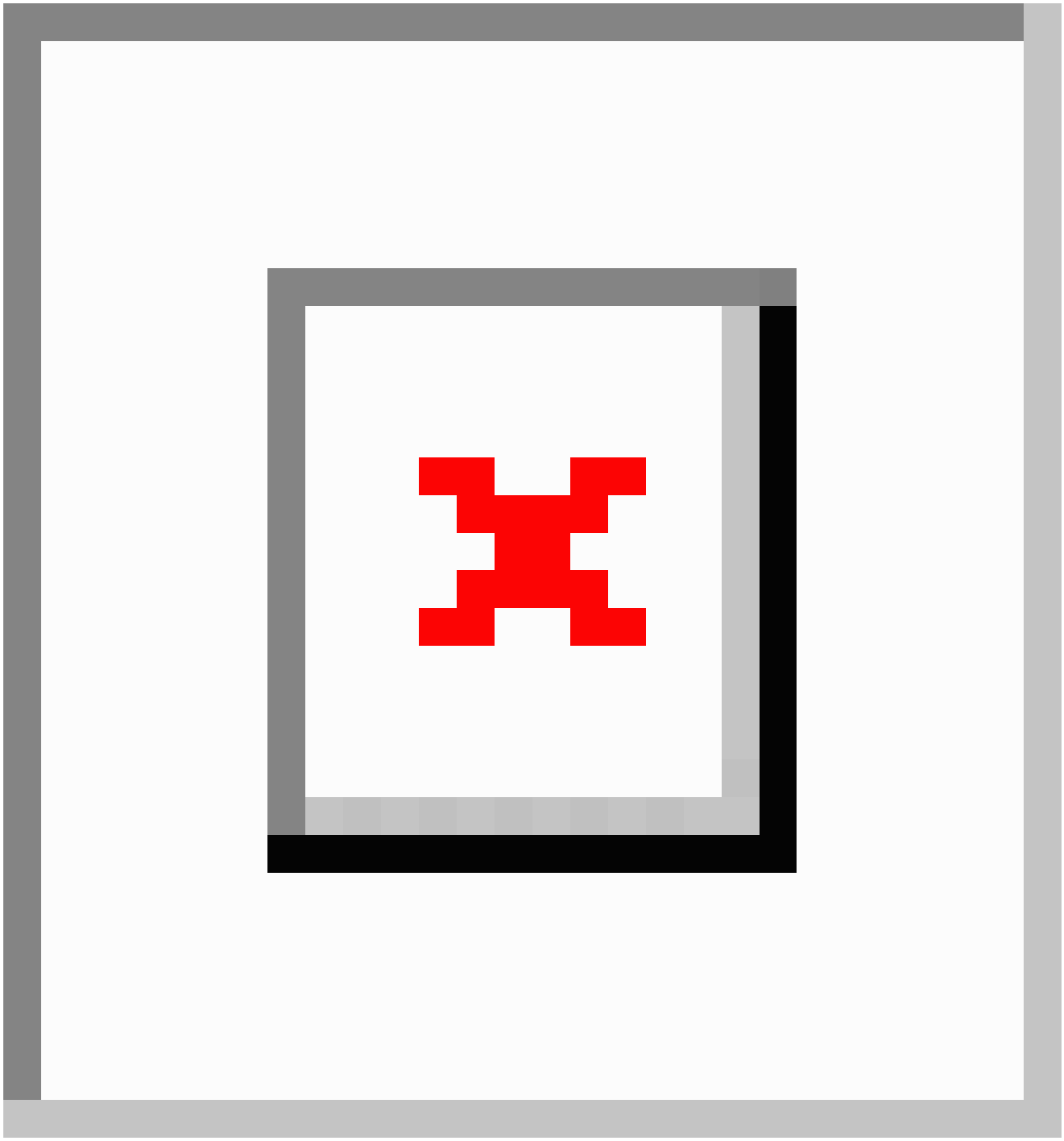
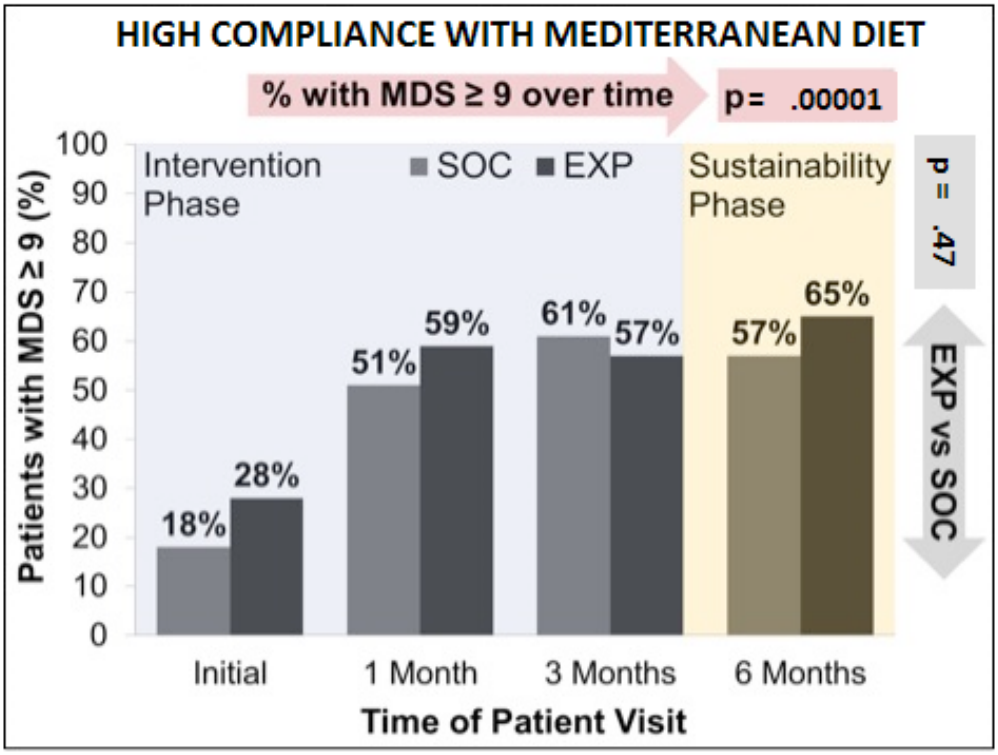


Figure 7. Proportion of patients achieving high compliance with the Mediterranean diet over time. MDS: Mediterranean diet score; SOC: standard-of-care arm; EXP: experimental arm.



Discussion

Principal Findings

Both traditional, SOC face-to-face counseling, and app-based dietary counseling were effective at getting participants to adhere to a Mediterranean diet in this US-based urban study of patients from a cardiology practice, and these dietary changes persisted even after counseling had ended. However, neither method of dietary counseling was more effective than the other at meeting the primary endpoint of Mediterranean diet compliance. These results were consistent whether the app was used heavily or lightly by participants in the experimental arm based on the subgroup analyses.

At each time point, participants were more satisfied with traditional counseling than with the app, but participant satisfaction regressed during the maintenance phase after traditional counseling ended. With the app, satisfaction did not regress. We speculate that this effect may be because participants were permitted to continue to use the app during the maintenance phase. Both groups showed increased satisfaction with their changed diet overall, but there were no differences in the level of satisfaction between the 2 groups.

Although the evidence suggests that participants with ASCVD derive more benefit from a Mediterranean diet based on comparison between the Lyon Diet Heart Study [4] and the PREDIMED investigation [5], participants in this study were more likely to adapt a Mediterranean diet if they did not have ASCVD. This result suggests that dietary patterns in this urban US population may be more modifiable for primary prevention purposes rather than for secondary prevention.

Comparison With Previous Work

Asynchronous mobile health–based dietary counseling has been shown to improve dietary targets. For example, Rossi et al demonstrated that mobile phone software that included an interactive diary for diet management and enabled users to communicate with a dietitian via short text messages was successful in helping users better identify and achieve Mediterranean diet targets such as the increased consumption of fresh fruits and vegetables [10]. However, this was a single-arm longitudinal investigation that did not use a comparison with usual care and was not studied in a non-Mediterranean population. Papadaki and Scott were able to show the efficacy of a Web-based intervention in promoting key components of the Mediterranean diet in a non-Mediterranean population via a quasi-experimental study [11], but their intervention did not use a mobile health–based approach. Our research is unique in that it studies the use of a smartphone-based dietary intervention for the Mediterranean diet by way of the RCT method.

We found that the use of a mobile app for the Mediterranean diet did not result in significant differences between groups or subgroups overall, aside from weight and BMI. No other biomarkers (eg, systolic BP and HbA_{1c}) indicated that the smartphone app was more successful than SOC dietary counseling at differentiating outcomes. Nevertheless, this investigation demonstrates that the Mediterranean diet can be achieved by a non-Mediterranean population. Given that dietary

counseling—which was standardized in both the SOC and intervention arms—resulted in increased adherence to the Mediterranean diet for participants in both treatment groups, a mobile app–based intervention is a practical option for administering dietary counseling when delivered in conjunction with traditional dietitian-based counseling. An app-based dietary intervention may be less resource intensive for maintenance of dietary changes and still deliver the same effect as traditional counseling. In addition, this effect was not altered by app use intensity, suggesting it is the baseline functionality of the app without the RD interaction that drove this benefit. Moreover, although not designed to be a weight-reducing diet, participants who received app-based Mediterranean diet counseling lost more weight than those who received SOC dietary counseling. The subgroup analysis suggests that this weight loss effect may have been driven by the presence of an exercise/activity log within the app.

Our findings are in keeping with previous research; for example, the EVIDENT II trial also demonstrated that dietary counseling is effective in increasing adherence to the Mediterranean diet in a population of primary care patients in Spain and found that the use of a mobile app had no significant impact on this result [12]. As EVIDENT II had greater female participation, these results also support that gender did not influence the outcome, and similar to EVIDENT II, despite finding no difference in the primary endpoint, there was a signal toward increased exercise in the app group (however, in our study, there was no accelerometer data and the signal was simply reflected in exercise log usage).

Limitations

This RCT was conducted as a pilot study. The strength of the investigation was limited by a small sample size of 100 patients. As more patients were screened than qualified for the study, biases based upon unrecorded reasons for study nonparticipation may also have influenced the outcome. In addition, the 6-month duration of the study may have been too short to demonstrate cardiovascular health benefits via parameters such as BP, lipid levels, HbA_{1c}, and CRP, and this pilot study was not adequately powered to detect such an effect.

As this study was performed at a single academic medical center, the ability to extrapolate these results to community-based practices or other locations including the nonurban environment is limited. Furthermore, the typical payer mix of our institution trends toward the underserved, and the results may be different if the participant pool was enriched with those from a higher socioeconomic status.

Emphasis should be made that this study cannot determine whether any sustainable dietary modification could have been achieved in the absence of counseling from a dietitian. Indeed, the benefit appears to have been driven by the initial RD interaction which occurred in both arms of the study, emphasizing the point that apps may best be utilized as an adjunct to traditional counseling as opposed to a replacement [13].

Study design may also have impacted the strength of our results as the nature of the intervention in this RCT prevented blinding of the participants, and several findings of the study (eg, dietary compliance) relied on self-reported information. Moreover, older and elderly adults are often more resistant to behavioral change and may be less comfortable with smartphones and mobile technology than younger adults [14]; this generational and age-related bias may have distorted the measured and reported efficacy of and satisfaction with the mobile app and Mediterranean diet. As the survey instrument was administered by the dietitian in SOC at the 1- and 3-month time points, these survey responses may have been biased; however, the results at the interim time points did not differ from the final time point (when all surveys were administered by the same coordinator). Despite its limitations, this RCT is hypothesis generating and can guide strategies for improving preventative cardiology. As the effects were not different between those who used the app heavily and those who did not use the app much, it is possible that the results would have been different with a different, more engaging app given the low PSS in the experimental group.

Conclusions

In summary, the dietary pattern of the urban US cardiology patient is modifiable to greater adherence to a Mediterranean diet, whether that dietary intervention is face-to-face or delivered via a mobile smartphone app. Supplemental features that promote weight loss within a smartphone app may, in addition, help weight loss. Although participants were less satisfied with the app, it may have presented a more time-effective means of delivering dietary counseling and resulted in less loss of effectiveness as compared with the period after completion of face-to-face counseling.

Acknowledgments

The authors thank the National Institutes of Health Small Business Innovation Research program for providing a grant to Vibrent Health to fund this study and the cardiologists of the GW Medical Faculty Associates for graciously screening their patients for participation. The authors also thank Vibrent Health for allowing them to have full editorial control of this manuscript, and no employee of Vibrent Health participated in its writing (and none of the authors have a financial relationship to Vibrent Health).

Conflicts of Interest

None declared.

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2017 Update: a report from the American Heart Association. Circulation 2017 Mar 7;135(10):e146-e603. [doi: 10.1161/CIR.0000000000000485] [Medline: 28122885]
2. Wilson MM, Reedy J, Krebs-Smith SM. American diet quality: where it is, where it is heading, and what it could be. J Acad Nutr Diet 2016 Feb;116(2):302-10.e1 [FREE Full text] [doi: 10.1016/j.jand.2015.09.020] [Medline: 26612769]
3. Murray CJ, Atkinson C, Bhalla K, Birbeck G, Burstein R, Chou D, Foreman, Lopez, Murray, Dahodwala, Jarlais, Fahami, Murray, Jarlais, Foreman, Lopez, Murray, US Burden of Disease Collaborators. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. J Am Med Assoc 2013 Aug 14;310(6):591-608 [FREE Full text] [doi: 10.1001/jama.2013.13805] [Medline: 23842577]
4. de Lorgeril M, Salen P, Martin JL, Monjaud I, Delaye J, Mamelle N. Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction: final report of the Lyon Diet Heart Study. Circulation 1999 Feb 16;99(6):779-785. [doi: 10.1161/01.cir.99.6.779] [Medline: 9989963]
5. Estruch R, Ros E, Salas-Salvadó J, Covas M, Corella D, Arós F, PREDIMED Study Investigators. Primary prevention of cardiovascular disease with a Mediterranean diet. N Engl J Med 2013 Dec 4;368(14):1279-1290. [doi: 10.1056/NEJMoa1200303] [Medline: 23432189]
6. de Lorgeril M, Renaud S, Mamelle N, Salen P, Martin JL, Monjaud I, et al. Mediterranean alpha-linolenic acid-rich diet in secondary prevention of coronary heart disease. Lancet 1994 Jun 11;343(8911):1454-1459. [doi: 10.1016/S0140-6736(94)92580-1] [Medline: 7911176]
7. Hoffman R, Gerber M. Evaluating and adapting the Mediterranean diet for non-Mediterranean populations: a critical appraisal. Nutr Rev 2013 Sep;71(9):573-584. [doi: 10.1111/nure.12040] [Medline: 24032362]
8. McManus DD, Gore J, Yarzebski J, Spencer F, Lessard D, Goldberg RJ. Recent trends in the incidence, treatment, and outcomes of patients with STEMI and NSTEMI. Am J Med 2011 Jan;124(1):40-47 [FREE Full text] [doi: 10.1016/j.amjmed.2010.07.023] [Medline: 21187184]
9. Davis T, Long S, Jackson R, Mayeaux E, George R, Murphy P, et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med 1993 Jun;25(6):391-395. [Medline: 8349060]
10. Rossi MC, Perozzi C, Consorti C, Almonti T, Foglini P, Giostra N, et al. An interactive diary for diet management (DAI): a new telemedicine system able to promote body weight reduction, nutritional education, and consumption of fresh local produce. Diabetes Technol Ther 2010 Aug;12(8):641-647. [doi: 10.1089/dia.2010.0025] [Medline: 20615106]
11. Papadaki A, Scott JA. Follow-up of a web-based tailored intervention promoting the Mediterranean diet in Scotland. Patient Educ Couns 2008;73(2):256-263. [doi: 10.1016/j.pec.2008.05.030] [Medline: 18640000]
12. Recio-Rodriguez J, Agudo-Conde C, Martin-Cantera C, González-Viejo MA, Fernandez-Alonso M, Arietaleanizbeaskoa M, EVIDENT Investigators. Short-term effectiveness of a mobile phone app for increasing physical activity and adherence to the Mediterranean diet in primary care: a randomized controlled trial (EVIDENT II Study). J Med Internet Res 2016 Dec 19;18(12):e331 [FREE Full text] [doi: 10.2196/jmir.6814] [Medline: 27993759]
13. Chen J, Gemming L, Hanning R, Allman-Farinelli M. Smartphone apps and the nutrition care process: current perspectives and future considerations. Patient Educ Couns 2018 Dec;101(4):750-757. [doi: 10.1016/j.pec.2017.11.011] [Medline: 29169863]
14. Pew Research Center. 2018. Mobile fact sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2018-10-24] [WebCite Cache ID 73PSfIKay]

Abbreviations

ASCVD: atherosclerotic cardiovascular disease
BMI: body mass index
BP: blood pressure
CRP: C-reactive protein
CVD: cardiovascular disease
EXP: experimental arm
HbA $_{1c}$: hemoglobin A_{1c}
HDL: high-density lipoprotein
LDL: low-density lipoprotein
MDS: Mediterranean diet score
PREDIMED: Prevención con Dieta Mediterránea
PSS: participant satisfaction score
RCT: randomized controlled trial
RD: registered dietitian
SOC: standard-of-care

Edited by I Weber; submitted 24.04.18; peer-reviewed by R Jansen, C Pellegrini, J Porter, J Chen; comments to author 04.06.18; revised version received 14.11.18; accepted 13.02.19; published 23.04.19.

Please cite as:
Choi BG, Dhawan T, Metzger K, Marshall L, Akbar A, Jain T, Young HA, Katz RJ
Image-Based Mobile System for Dietary Management in an American Cardiology Population: Pilot Randomized Controlled Trial to Assess the Efficacy of Dietary Coaching Delivered via a Smartphone App Versus Traditional Counseling
JMIR Mhealth Uhealth 2019;7(4):e10755
URL: https://mhealth.jmir.org/2019/4/e10755/
doi:10.2196/10755
PMID:31012860

Original Paper

A Mobile Smoking Cessation Intervention for Mexico (Vive sin Tabaco... ¡Decídete!): Single-Arm Pilot Study

Ana Paula Cupertino[1], PhD; Francisco Cartujano-Barrera[1], MD; Mariana Ramírez[2], LSMW; Rosibel Rodríguez-Bolaños[3], PhD; James F Thrasher[4], PhD; Gloria Pérez-Rubio[5], PhD; Ramcés Falfán-Valencia[5], PhD; Edward F Ellerbeck[2], MD, MPH; Luz Myriam Reynales-Shigematsu[3], PhD

[1]Department of Cancer Prevention and Control, Hackensack University Medical Center, Hackensack, NJ, United States

[2]Department of Preventive Medicine and Public Health, University of Kansas Medical Center, Kansas City, KS, United States

[3]Department of Tobacco Research, National Institute of Public Health, Cuernavaca, Mexico

[4]Department of Health Promotion, Education and Behavior, University of South Carolina, Columbia, SC, United States

[5]Instituto Nacional de Enfermedades Respiratorias Ismael Cosio Villegas, Laboratorio HLA, Mexico City, Mexico

Corresponding Author:
Luz Myriam Reynales-Shigematsu, PhD
Department of Tobacco Research
National Institute of Public Health
Avenida Universidad 655, Colonia Santa María Ahuacatitlán
Cuernavaca, 62100
Mexico
Phone: 52 777 101 2977
Email: lreynales@insp.mx

Abstract

Background: Of the 14.3 million Mexicans who smoke, only a minority take advantage of evidence-based approaches to smoking cessation. Mobile health interventions have the potential to increase the reach of effective cessation interventions in Mexico.

Objective: This study aimed to assess the feasibility and acceptability of an innovative, personalized, and interactive smoking cessation mobile intervention developed for Mexican smokers.

Methods: We recruited 40 Mexican smokers to participate in *Vive sin Tabaco... ¡Decídete!*, a smoking cessation program that uses a tablet-based decision support software to drive a 12-week text messaging smoking cessation program and pharmacotherapy support. Outcome measures included participant text messaging interactivity with the program, participant satisfaction, and 12-week verified abstinence using urinary cotinine testing or exhaled carbon monoxide.

Results: Average age of the participants was 36 years (SD 10.7), and they were primarily male (65%, 26/40) with at least an undergraduate degree (62%, 25/40). Most participants (95%, 38/40) smoked daily and were interested in quitting in the next 7 days. As an indicator of participant interactivity, participants sent an average of 21 text messages during the 12-week intervention (SD 17.62). Of the 843 messages that participants sent to the program, only 96 messages (11.3%, 96/843) used keywords. At 12 weeks, 40% (16/40) of participants were biochemically verified (87%, 35/40, follow-up rate). The majority of participants (85%, 30/35) reported being very satisfied or extremely satisfied with the program.

Conclusions: The *Vive sin Tabaco... ¡Decídete!* smoking cessation mobile intervention was accepted by participants, generated high satisfaction and high text messaging interactivity, and resulted in a noteworthy cessation rate at the end of treatment. This intervention is a promising strategy for smoking cessation in Mexico. Additional testing as a formal randomized clinical trial appears warranted.

(JMIR Mhealth Uhealth 2019;7(4):e12482) doi:10.2196/12482

KEYWORDS

smoking; smoking cessation; mHealth; text messages; global health

Introduction

Background

Currently, 14.3 million Mexican adults (16.4%) smoke [1], and it is expected that more than 4 million will die of tobacco-related diseases in the next decade if smoking prevalence remains unchecked [2]. The prevalence of smoking has remained relatively consistent over the last decade [1,3], despite implementation of taxes [4,5], smoke-free policies [6,7], health warnings on cigarette packages [8-10], and advertising restrictions [11], as recommended by the World Health Organization's Framework Convention on Tobacco Control [12]. The potential for reducing the projected morbidity and mortality associated with smoking depends greatly on reaching smokers and delivering cost-effective cessation interventions.

Currently, 8 in 10 Mexican smokers are interested in quitting smoking [1]. However, less than 10% of Mexican smokers take advantage of the evidence-based approaches to smoking cessation (pharmacotherapy and counseling) [1] that are offered by the public health care system [13-15]. Although health care providers in Mexico are encouraged to address smoking with their patients in every visit [16,17], most health care providers fail to initiate cessation treatment [18]. Overcoming the burden of tobacco use in Mexico demands affordable, accessible, and effective solutions.

Mexicans are more likely to be nondaily and light smokers (<10 cigarettes per day [CPD]) [1,19]. The 2015 Global Adult Tobacco Survey in Mexico found that 53.7% of current smokers are nondaily smokers and, among daily smokers, the average number of CPD was 7.7 [1]. Light smokers believe their lower level of smoking reduces or eliminates their health risk despite evidence to the contrary [20]. Light smoking significantly increases the risk for cancer, all-cause mortality, and adverse cardiovascular outcomes [21]. It is important to identify innovative smoking cessation strategies tailored to the needs of Mexican smokers.

Developments in the sophistication of mobile technologies allow for flexible delivery of text messages, with algorithms used to tailor content to individual motivational and behavioral needs for smoking cessation [22-26]. A number of studies have examined the effectiveness of text messaging interventions to promote and support smoking cessation [27-29]. A Cochrane meta-analysis of these studies indicates that text messaging interventions increase the likelihood of staying quit by approximately 1.7 times (9.3% quit rate with text messages vs 5.6% quit rate with no program) [27]. All the studies included in the meta-analysis were conducted in high-income countries with limited generalizability to low- and middle-income countries. Reflecting the global trend in the uptake of cell phones, Mexico is the eleventh largest mobile market in the world, with 107.8 million active cell phones [30] for its 123 million inhabitants [31]. The very high rate of cell phone ownership, the low use of smoking cessation treatments, and the evidence that text messaging can enhance a smoking cessation intervention provide a unique opportunity to assess a smoking cessation mobile intervention in Mexico. Considering this premise, the US National Cancer Institutes' text messaging program [32] was adapted for Mexican smokers interested in quitting using focus groups and interviews [33]; however, no formal evaluation of that effort has been done.

Objective

This pilot study aimed to assess the feasibility and acceptability of *Vive sin Tabaco... ¡Decídete!*, an innovative smoking cessation mobile intervention for Mexican smokers. This mobile innovation is achieved by connecting a Web-based decision-making tool used to develop a personalized quit plan with the delivery of tailored text messages over 12 weeks. Results from this study will inform future implementation and dissemination studies to achieve significant reductions in tobacco-related morbidity and mortality in Mexico and provide a model for population-based smoking cessation mobile interventions.

Methods

Setting

This study was conducted between March and August 2017 at the Medical Center of the Autonomous University of the State of Morelos, located in Cuernavaca, Morelos, Mexico. This urban primary health care clinic serves an average of 100 individuals on a daily basis. None of the services provided by the clinic address smoking cessation.

Participants

Participants were recruited through printed posters and multimedia venues including ads through the National Institute of Public Health's website and Facebook and local radio announcements. Potential participants emailed or called the study personnel to learn more about the study. Eligibility assessment was conducted over the phone. Eligible participants were of Mexican origin, aged 18 years or older, had smoked for at least 6 months, smoked at least 3 days per week, were interested in quitting within the next 30 days, had a cell phone with text messaging capacity, and were willing to complete baseline and 12-week follow-up surveys. Participants were excluded from the study if they were planning to move within the next 6 months, consumed other forms of tobacco (including electronic cigarettes) or illicit drugs in Mexico (eg, cannabis and cocaine), or had another household member enrolled in the study. All subjects gave informed consent before participation in the study. Participants received 300 Mexican pesos (approximately US $17) at baseline and follow-up as an incentive for their time and transportation. The Human Subjects Committee of the National Institute of Public Health approved the study procedures.

Intervention

Vive sin Tabaco... ¡Decídete! is a smoking cessation mobile intervention that encompasses 3 integrated components: (1) a tablet-based software that collects personal smoking-related information to support the development of an individualized quit plan and guides the ensuing text messages program, (2) a 12-week individually tailored text messages program with interactive capabilities, and (3) pharmacotherapy support when applicable.

Vive sin Tabaco... ¡Decídete! Tablet-Based Software

The tablet-based, decision support tool was designed to help smokers create a personalized smoking cessation plan and to collect data that tailored the text messages delivered over the ensuing 12 weeks [34]. This tool was adapted from 2 smoking cessation Web-based, informed decision-making tools for Latinos in the United States [35] and Mexico [36,37]. This tablet-based tool consisted of interactive features that lead smokers through various steps in the quitting process. The program included testimonies from ex-smokers and features short video clips (0:21-2:03 min) and narrated graphics on the benefits of quitting while also describing how cessation pharmacotherapy (nicotine replacement therapy [NRT]) can support abstinence. The program also collected basic information about the participants' smoking history, including the number of days they smoke each week and the number of cigarettes smoked per day. At the end of the 10- to 15-min session, participants were prompted to request pharmacotherapy if interested and clinically recommended and to select a quit date within a 30-day timeframe. Upon completion of the tablet-based software, participants were provided with a 1-page summary of their personalized cessation plan (eg, the selected quit date and pharmacotherapy with the recommended dose and regimen). Next, participants automatically began receiving the text message portion of the intervention.

Vive sin Tabaco... ¡Decídete! Text Messaging

We developed a library of 304 text messages in Spanish to support a 12-week cessation program [38] based on the social cognitive theory [39]. Text messages were informed by literature reviews on educational facts and strategies for smoking cessation, feedback from national tobacco control experts, and results from focus groups with Mexican smokers and ex-smokers. The text message intervention allowed 3 levels of interactivity: (1) prescheduled standard messages, (2) keyword-triggered standard messages, and (3) counselor-personalized responses.

Prescheduled Standard Messages

The main goal of the prescheduled standard messages was to provide counseling through *educational* (eg, health risks of smoking, immediate and long-term benefits of quitting smoking, and how to correctly use the pharmacotherapy), *motivational* (eg, intrinsic and extrinsic motivation), and *behavioral* (eg, reminders to use strategies to cope with triggers, self-control through goals, self-monitoring, and prompts to order NRT) messages to facilitate quitting and supporting abstinence. Messages were automatically tailored to the participant's name(s), gender, pharmacotherapy indication, and the selected quit date. Text messages were organized along a 12-week timeline designed to support a personalized quit plan: (1) prequit (29 days), (2) quit day (1 day), (3) maintenance (28 days), and (4) relapse prevention (8 weeks). The content and number of messages varied as the intervention progressed (see Table 1).

Keyword-Triggered Standard Messages

These messages consisted of automated immediate responses sent to participants who texted 1 of the following keywords: *Antojo* (Spanish for "Crave"), *Estrés* (Spanish for "Stress"), *Recaída* (Spanish for "Relapse"), *Familia* (Spanish for "Family"), *Tristeza* (Spanish for "Sadness"), and *Consejo* (Spanish for "Advice"). In addition, throughout the 12-week program, participants received 7 response-triggered (YES or NO) text messages to assess their smoking status (eg, *¿Sigues sin fumar? Responde SÍ o NO y te ayudaremos. ¡Seguimos contigo!* [Spanish for "Are you still smoke-free? Reply YES or NO and we will help you. We are here with you!"]). If participants indicated that they were smoking, these automated messages encouraged them to set a new quit date. Following the Mexican Federal Telecommunications Institute regulations, participants could withdraw from the text message program at any moment by sending the keyword *Alto* (Spanish for "Stop") [40].

Counselor-Personalized Responses

Taking advantage of the text message platform's capability to recognize free texting (nonkeyword) from participants, *Vive sin Tabaco... ¡Decídete!* encouraged participants to text any feelings, concerns, and/or questions to the program (eg, *Puedes escribirnos en todo momento. Te apreciamos y nuestro compromiso es ayudarte. ¡Recuérdalo!* [Spanish for "You can text us at any time. We appreciate you and we are committed to help you. Remember it!"]). A trained smoking cessation counselor answered these messages following standardized protocols (eg, motivational interview and pharmacotherapy delivery, use, adherence, and side effects). The counselor was trained on the Basic Skills for Working with Smokers course by The University of Massachusetts Medical School [41]. The counselor monitored and triaged queries daily, responding within 24 hours of receipt of text messages sent by participants.

Table 1. Types of messages, stages, duration, number, and examples of text messages.

Type of message	Stage	Duration	Number	Examples
Prescheduled standard messages	Prequit	29 days	2 to 3 messages a day	*[Nombre], crea un grupo que te apoye en este proceso, pueden ser tus familiares y amistades más cercanas. Diles que a partir de [Fecha para dejar de fumar] dejarás de fumar. ¡Cuenta con ellos(as), cuenta con nosotros!*; [Name], create a group that will support you through this process, it can be your family and close friends. Tell them that from [Quit date] you will quit smoking. Count on them, count on us!;
				¡BUENAS NOTICIAS! Dejando de fumar cuidas el medio ambiente. ¿Cómo? Cada año, la industria tabacalera es responsable de la tala de 600 millones de árboles; GOOD NEWS! By quitting smoking, you are taking care of the environment. How? Each year, the tobacco industry is responsible for cutting down 600 million trees.;
				[Nombre], al dejar el cigarro te convertirás en un ejemplo para otros, ¡muchos querrán lograrlo como tú!; [Name], when you quit smoking, you become a role model to others, many will want to stop smoking like you!;
				La vida está llena de retos, dejar de fumar es uno más que vas a superar. Piensa en lo bien que te sentirás cuando lo hayas logrado; Life is full of challenges, quitting smoking is one more challenge you will overcome. Think about how good you will feel once you achieve it!
Prescheduled standard messages	Quit day	1 day	4 messages	*[Nombre], ¡FELICIDADES! Ha llegado el día. Hoy comienzas una nueva vida. Recuerda por qué estás dejando de fumar*; [Name], CONGRATULATIONS! The day has arrived. Today you start a new life. Remember why you are quitting smoking;
				[Nombre], ¿cómo te sientes? Sabemos que las primeras 24 horas son las más difíciles, ¡recuerda que cuentas con nosotros! Escríbenos; [Name], how are you feeling? We know the first 24 hours are the most difficult, remember that you can count on us! Text us
Prescheduled standard messages	Maintenance	28 days	3 to 4 messages a day	*Dejar de fumar es una decisión que se toma todos los días. Reafírmala cada mañana. ¡Lo estás logrando!*; Quitting smoking is a decision you make every day. Reaffirm it each morning. You are doing it!;
				[Nombre], ¿estás disfrutando esta nueva etapa de tu vida? Cuéntanos lo que más te gusta de haber dejado de fumar; [Name], are you enjoying this new stage in your life? Tell us what you like the most about no longer smoking.;
				¿Ya publicaste en Facebook que dejaste de fumar? ¡Te sorprenderá ver la respuesta positiva de todos!; Have you posted on Facebook that you quit smoking? You will be surprised to see the positive responses from everyone!;
				¿Ya probaste tu platillo favorito ahora que dejaste el cigarro? Redescubre su sabor, ¡disfrútalo!; Have you tried your favorite dish now that you quit smoking? Rediscover its flavor. Enjoy it!
Prescheduled standard messages	Relapse prevention	56 days	1 to 2 messages a day	*Hoy puedes decir: "¡Soy [Nombre] y llevo 35 días sin fumar! ¡Sigo adelante!*; Today you can say, "I am [Name] and I have gone 35 days without smoking! I can keep going!";
				¡Felicidades [Nombre]! Ahora que dejaste de fumar, es menos probable que tus hijos(as) lo hagan. ¡Sigue cuidando tu futuro y el de los tuyos!; Congratulations [Name]! Now that you stopped smoking, it is less likely that your children will do it. Keep taking care of your future and the future of your loved ones!;
				[Nombre], planea un bonito fin de semana con las personas que más quieres. ¡6 semanas sin fumar es un gran motivo para celebrar!; [Name], plan a nice weekend with the people you love the most. 6 weeks without smoking is a great reason to celebrate!

Type of message	Stage	Duration	Number	Examples
Prescheduled standard messages	Relapse	3 days	4 messages a day	*Recaer no significa que has perdido la batalla, recaer nos da la oportunidad de aprender y salir victoriosos. ¡Inténtalo una vez más!*; To relapse does not mean you have lost the battle; to relapse gives you the opportunity to learn and become victorious. Try it one more time!; *[Nombre], ¿qué te hace fumar? Cuéntanos. Juntos encontraremos nuevas soluciones para dejar el tabaco;* [Name], what makes you smoke? Tell us. We can find new solutions together to quit tobacco
Keyword-triggered standard messages	Family	Per participant request	14 messages	*Dejar de fumar vale la pena por ti y por tus seres queridos. [Nombre], ¡te felicitamos por tu esfuerzo! ¡Sigue adelante!;* Quitting smoking is worth it for the sake of you and your loved ones. [Name], we congratulate for your effort! Keep going!; *[Nombre], dejar de fumar también demuestra amor a tu familia. Sigue así, ¡vas muy bien!;* [Name], quitting smoking also shows love for your family. Keep going; you are doing great!
Keyword-triggered standard messages	Crave	Per participant request	10 messages	*Deja lo que estás haciendo y camina por 5 minutos. Te distraerás y el antojo desaparecerá;* Drop what you are doing and walk for 5 minutes. It will distract you and the cravings will go away.; *[Nombre], piensa en lo lejos que has llegado. No vale la pena volver a fumar, ¡sigue adelante!;* [Name], think of how far you have come. It is not worth it to smoke again, keep going!
Keyword-triggered standard messages	Stress	Per participant request	11 messages	*Inhala profundo por 3 segundos, aguanta el aire por otros 3, y exhala por 6 segundos. Mientras respiras hondo, imagina que estás en un lugar bello y tranquilo. Repítelo, ¡esto te relajará!;* Inhale deeply for 3 seconds, hold your breath for another 3 seconds, and exhale for 6 seconds. While you breathe deeply, imagine that you are in a beautiful and peaceful place. Repeat it; this will relax you!; *[Nombre], escucha tu música favorita, te servirá para controlar el estrés. ¡Será de mucha ayuda!;* [Name], listen to your favorite music, it will help you manage stress. It will be a great help!
Keyword-triggered standard messages	Advice	Per participant request	9 messages	*¡Mantente ocupado! Hay tantas cosas que puedes hacer sin fumar: ir al cine, ir por un café, leer, bailar, ejercitarse;* Keep yourself busy! There are so many thing that you can do without smoking: go to the movies, go for a coffee, read, dance, and exercise; *[Nombre], caminar al menos 15 minutos al día es excelente durante el proceso para dejar de fumar, ¡muévete!;* [Name], walking at least 15 minutes a day is excellent during the quitting process, get moving!
Keyword-triggered standard messages	Sadness	Per participant request	24 messages	*[Nombre], ¿sabías que el ejercicio mejora tu estado de ánimo? El ejercicio te hará sentir bien física y mentalmente. Camina, baila o ve al gimnasio. Te ayudará a conseguir tu objetivo;* [Name], did you know exercising improves your mood? Exercising will make you feel better, both physically and mentally. Walk, dance, or go to the gym. It will help you reach your goals; *[Nombre], la gente que nos rodea influye en nuestro estado de ánimo. Trata de rodearte de personas positivas, que te hagan sentir bien;* [Name], people around us affect our mood. Try to surround yourself with positive people that make you feel good

Nicotine Replacement Therapy

The choice of pharmacotherapy followed the practice guidelines for treating smokers in Mexico [16,17]. Only daily smokers who smoked 6 or more CPD were eligible to use nicotine patches. Nicotine patches were contraindicated in participants who (1) had a heart attack in the last 2 months, (2) had a stroke in the last 6 months, (3) have been diagnosed with an arrhythmia or tachycardia, (4) have uncontrolled hypertension, and (5) were using warfarin. Nicotine patches were not offered to these participants. Participants who smoked 10 or more CPD and had no contraindications were offered to use 10 weeks of nicotine patches: 21-mg nicotine patches to be used during the first 6 weeks, followed by 14-mg nicotine patches for 2 weeks, and 7-mg patches for the last 2 weeks. Participants who smoked between 6 and 9 CPD and had no contraindications were offered

to use 8 weeks of nicotine patches: 14-mg nicotine patches to be used during the first 6 weeks, followed by 7-mg patches for the last 2 weeks. The NRT was provided in 2 phases. At baseline, each participant received a 4-week supply if they were eligible and interested in using it. Beginning at the second week of the intervention, participants received text message queries to see if they were interested in receiving more NRT. If a participant indicated such an interest, a 4- or 6-week supply was shipped to their home. Participants were prompted to start using their nicotine patches on their selected quit date.

Gateway Infrastructure

To implement the text messages system, we worked with *Agile Health Inc* [42], a text messaging health company, which hosted the software to interface with the *Vive sin Tabaco... ¡Decídete!* tablet-based software [34]. We utilized *Agile Health* 's platform and application program interface to create a customized system to support the smoking cessation program. The platform allowed the counselor to monitor the text messages sent by the participants and categorize the messages using keywords (which triggered an automatic response) and counselor-required messages. The platform allowed the counselor to interact with participants while being able to see the text message history and other participant information such as age, gender, and the date they started the intervention. Text messages were delivered through *Auronix* [43], a text message gateway in Mexico identified through market research and technical testing. *Auronix* 's service provider had the capacity to (1) engage bidirectional real-time communication through a short code, (2) connect to all carriers in Mexico, (3) receive participants' text messages containing special characters used in Spanish (ñ á é í ó ú ¡ ¿), and (4) send a large number of text messages from *Agile Health* 's platform.

Measures

The in-person baseline survey assessed sociodemographic variables such as age, gender, education level, marital status, and type of health insurance. Other variables collected were physical nicotine dependence (the Fagerström Test for Nicotine Dependence [44]—a 6-item test that evaluates the quantity of cigarette consumption, the compulsion to use, and the dependence), the number of cigarettes smoked per day, and the number of previous quit attempts. We also assessed the frequency of messages that participants sent to the program, including the use of keywords, across each stage of the intervention. At 12 weeks after enrollment, an in-person follow-up survey was conducted by trained research staff and biological samples were collected to verify cessation status. Acceptability measures included satisfaction questions such as "How satisfied are you with the smoking cessation text message program?"

Outcomes

The primary outcome was cotinine-verified 7-day point prevalence abstinence (no cigarettes in the past 7 days) at 12 weeks. This was biochemically verified using urinary cotinine testing, with a cutoff of 200 ng/ml cotinine [45,46]. If the participant was still using NRT, exhaled carbon monoxide, with a cutoff of 6 ppm [45], was used to verify smoking abstinence. The secondary outcomes were acceptability of the program and text messaging interactivity.

Data Analysis

We calculated simple frequencies for categorical variables and means and SDs for continuous variables. The primary analysis on cessation was conducted using an intention-to-treat analysis, in which participants lost to follow-up are considered smokers.

Results

Participant Recruitment and Characteristics

During a single week of recruitment, 122 smokers contacted the study personnel via phone or email for information; among them, 106 were contacted and assessed for eligibility by telephone and 72 were identified as eligible for the study. Overall, 41 smokers consented to participate and completed the baseline assessment in person; 1 smoker was removed from the study because of a carrier blockage that could not be solved, resulting in 40 smokers enrolled in the study (Figure 1).

Participants' age at baseline ranged from 20 to 59 years (mean 36.0, SD 10.7); 65% (26/40) of the participants were men, 50% (20/40) were single, 62% (25/40) had college or postgraduate education, and 80% (32/40) had health insurance coverage. Most participants smoked daily (95%, 38/40) and were interested in quitting in the next 7 days (95%, 38/40). Half of the participants were light smokers (smoked 10 or less CPD) and, according to the Fagerström test, 70% (28/40) of the participants reported low levels of nicotine dependence (Table 2).

Figure 1. Vive Sin Tabaco… ¡Decídete! intervention flow.

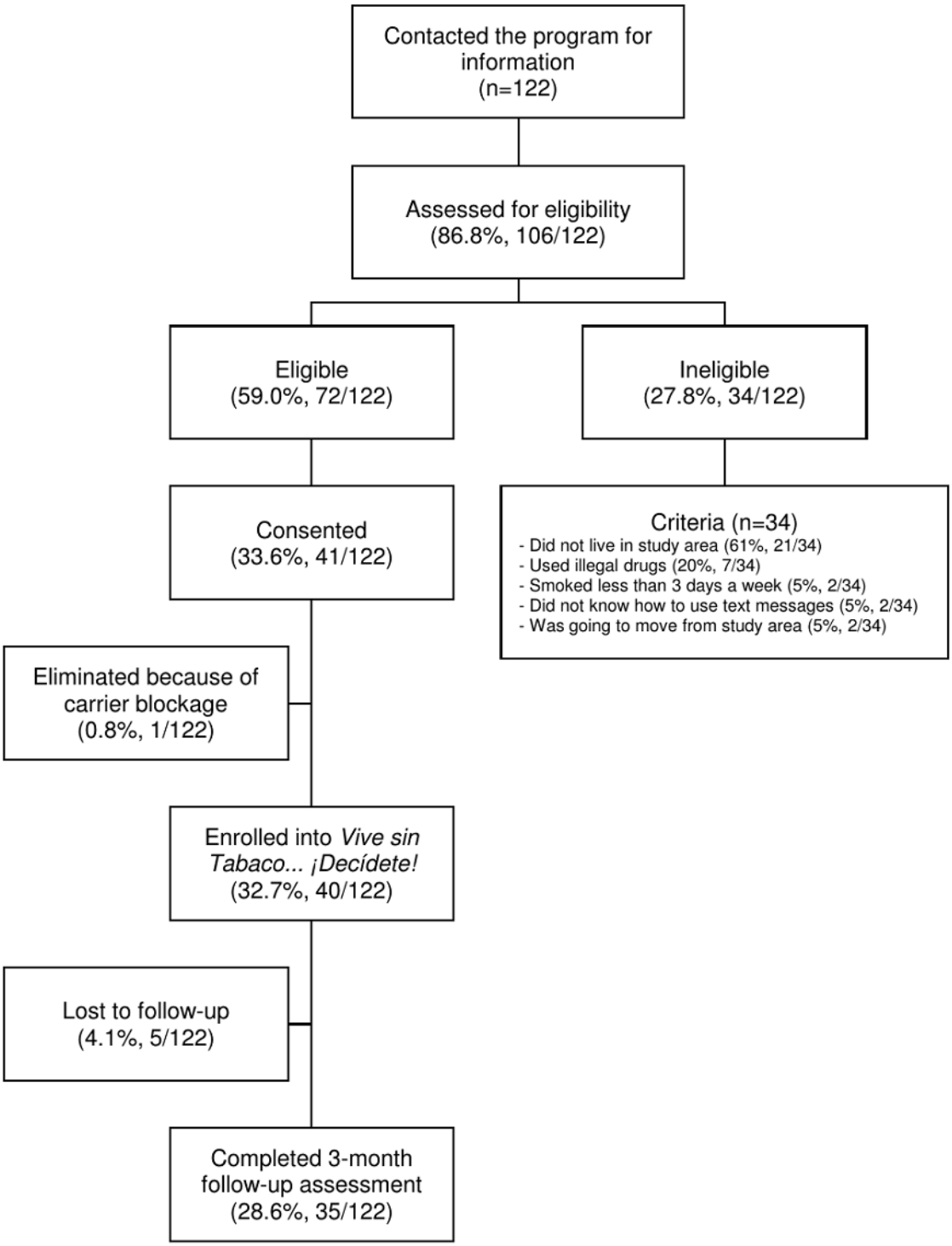


Table 2. Baseline characteristics of participants (n=40).

Profile characteristics	Statistics
Age, mean (SD)	36.0 (10.7)
Sex, n (%)	
Men	26 (65)
Education level, n (%)	
Less than high school graduate	3 (7)
High school graduate	7 (17)
Technical school	5 (12)
College graduate	18 (45)
Postgraduate	7 (17)
Marital status, n (%)	
Married or cohabitating	16 (40)
Single	20 (50)
Divorced, separated, or widowed	4 (10)
Health coverage, n (%)	
Instituto Mexicano del Seguro Social (English: Mexican Social Security Institute)	25 (62)
Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (English: Institute for Social Security and Services for State Workers)	7 (17)
None	8 (20)
Smoking pattern, n (%)	
Nondaily	2 (5)
Daily, 1-9 CPD[a]	20 (50)
Daily, 10-19 CPD	12 (30)
Daily, 20 or more CPD	6 (15)
Fagerström test for nicotine dependence, n (%)	
Low dependence	28 (70)
Moderate dependence	11 (27)
High dependence	1 (2)
Quit attempt in the previous year , n (%)	
Yes	26 (65)
No	14 (35)

[a]CPD: cigarettes per day.

Text Messaging Utilization

Participants received approximately 180 automated messages during the 12 weeks; none of the participants texted the word STOP to disenroll from the program. During the 12-week intervention period, participants sent 843 text messages, an average of 21 text messages per participant (SD 17.62). Of the 843 messages that participants sent to the program, only 96 (11.3%) used keywords. Participants varied in the frequency of sending text messages: 3 (7%, 3/40) never interacted with the program, 16 (40%, 16/40) had low interaction (1-9 messages), 17 (37%, 17/40) had medium interaction (10-49 messages), and 4 (10%, 4/40) had high interaction (>50 messages). Interaction varied across the different stages of the program (Figure 2). Interaction was very high at the beginning of the intervention and on the quit date, decreasing progressively as the program continued except for spikes on days 7, 14, 28, 42, 56, and 77 after the quit date. These spikes were because of smoking status being assessed on those days via text message. Overall, 15 (37%, 15/40) participants notified the program that they had relapsed and 7 (17%, 7/4-) set up a new quit date.

Pharmacotherapy Utilization

Three-quarters (75%, 30/40) of the participants were eligible to use NRT, all of whom requested an initial supply of NRT. Of these 30 participants who requested NRT at baseline, 18 (60%) requested a refill at 4 weeks (Figure 3).

Figure 2. Text messages interaction by participants during the intervention. QD: quit day.

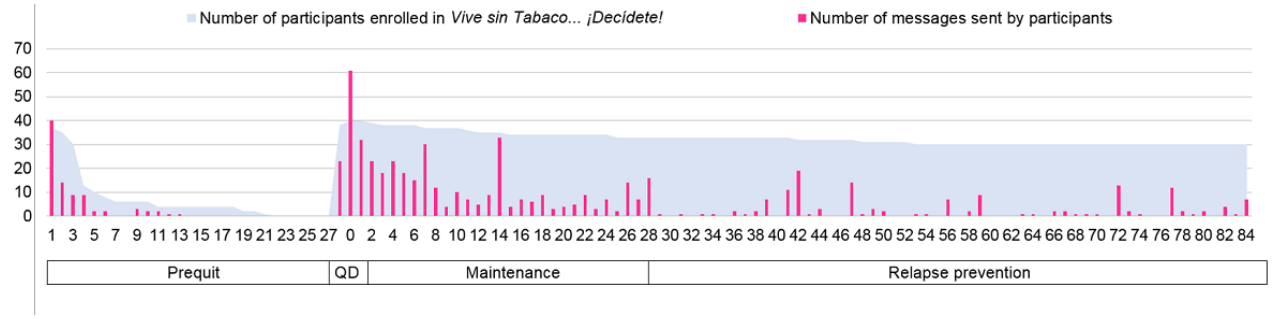
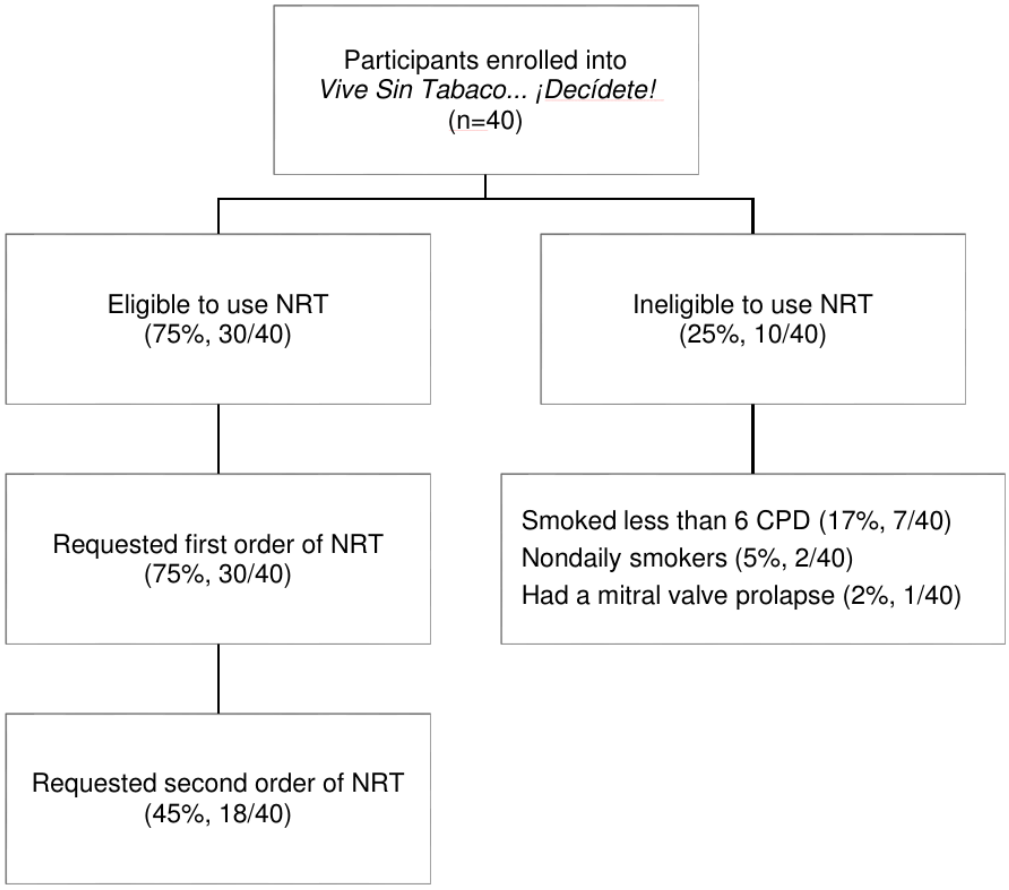


Figure 3. Utilization of nicotine replacement therapy (NRT) during the intervention. CPD: cigarettes per day.



Cessation, Retention, and Satisfaction

At 12 weeks, 16 participants (40%, 16/40) were biochemically verified abstinent using intent-to-treat analysis (Table 3). The follow up-rate at 12 weeks was 87.5%. Of the participants who completed the follow-up assessment, most (85.7%) reported being very satisfied or extremely satisfied with the program. In addition, 17 participants (48.5%) reported not being able to send text messages at some point of the intervention because of not having enough credit on their cell phones.

Table 3. The 3-month follow-up outcomes.

Outcome	n (%)
Smoking Abstinence (n=40)[a]	
Self-reported 7-days smoking abstinence	18 (45)
Biochemically verified abstinence urine cotinine (≤200 ng/ml)	15 (37)
Biochemically verified abstinence exhaled carbon monoxide (≤6 ppm)	1 (2)
Satisfaction (n=35)	
Extremely satisfied	13 (37)
Very satisfied	17 (48)
Satisfied	5 (14)

[a]Intention-to-treat analysis was used.

Discussion

Principal Findings

To the best of our knowledge, *Vive sin Tabaco... ¡Decídete!* is the first smoking cessation mobile intervention specifically tailored to the needs of Mexican smokers. The program has a unique 2-component platform that includes a tablet-based program at baseline to collect information used to develop a personalized quit plan and the delivery of the 12-week messaging program. This system generated substantial interest among smokers, as indicated by the rapid recruitment of our sample within a single week. The program was well received by the participants, most of whom engaged in high levels of interactivity with the program (eg, bidirectional messaging) and indicated high levels of participant satisfaction. The participants also expressed a high level of interest in using NRT in conjunction with the text messaging program. Although we did not assess medication adherence, the vast majority of participants used the text messaging program to request additional NRT, suggesting that most participants completed at least a 4-week course of therapy. The 40% rate of smoking cessation seen at week 12 (end of treatment) appears promising and is in line with end-of-treatment cessation rates seen in trials of NRT that used substantial in-person counseling.

Implications for Future Research

A smoking cessation mobile intervention can only be effective and sustainable if it is properly deployed in an environment that reaches a large number of smokers in need of evidence-based services. In Mexico, the most logical setting for the deployment of *Vive sin Tabaco... ¡Decídete!* is within the universal health care system, which is founded on a network of comprehensive primary care clinics [47,48]. Primary health within Mexico's universal health care system follows established guidelines for the identification and treatment of smokers [16,17]. However, implementation of these guidelines has been limited because of the lack of time during routine care, inadequate training of personnel, and competing patient demands [18]. *Vive sin Tabaco... ¡Decídete!* has the potential to overcome these barriers as it was designed for easy integration into primary health centers without disrupting clinical workflows.

The text messaging program appears to be a promising, low-cost alternative to in-person or telephone counseling to prompt smoking cessation, although additional strategies to eliminate the costs incurred by participants generating text messages to interact with the program may be needed. In this study, participants preferred to send their own, self-composed text messages rather than relying on keywords from the program for a response. This suggests that reliance on keywords may be insufficient for smoking cessation counseling via text messaging in Mexico. Hence, there may be additional costs involved in having trained personnel responding to participants' text messages, as occurred in this study. Participants' text messages content should be analyzed using qualitative methods to identify common themes. These methods can guide the creation of a categorized codebook that would be able to retrieve and send responses automatically, thus reducing the need for trained personnel responding to self-composed participants' text messages.

Limitations

This study had a number of limitations. This was a pilot study and did not have a control group. Due to the small sample size, the results are not generalizable to all Mexican smokers. Follow-up was limited to a single assessment at week 12, when the program ended. Analyses were limited to quantitative assessments of participant interactions. Furthermore, the sample was more highly educated and smoked more heavily than the general population of smokers in Mexico; future research is warranted to determine whether the effectiveness of this type of intervention is generalizable to those who are from lower socioeconomic status groups. Contrary to the US clinical guidelines [49], NRT is contraindicated in Mexico for those who smoke less than 6 cigarettes a day [17], which is a group that represents about 75% of smokers in Mexico [1]. It is possible that the cessation rate in this study could have been higher if light smokers had access to NRT as determined in the US clinical guidelines [49]. Despite these limitations, the study suggests that *Vive sin Tabaco... ¡Decídete!* is highly acceptable and holds promise for further testing, including a cost-effectiveness analysis.

Conclusions

The *Vive sin Tabaco... ¡Decídete!* smoking cessation mobile intervention was well accepted by participants, generated high

satisfaction and frequent 2-way interactivity, and resulted in noteworthy cessation rates at the end of treatment. The program appears to offer a promising strategy for smoking cessation in Mexico, particularly in the context of primary care clinics that could deploy the *Vive sin Tabaco... ¡Decídete!* tablet to assist participant enrollment. Additional testing in a formal randomized clinical trial is needed before widespread dissemination of *Vive sin Tabaco... ¡Decídete!*

Acknowledgments

This study was funded by the National Institutes of Health 3P30CA168524-04S1. *Vive sin Tabaco... ¡Decídete!* would not have been possible without the help from the multidisciplinary and multinational teams involved. The authors acknowledge the research teams from (1) The University of Kansas Medical Center, Kansas, US (Jaime Perales, PhD; Kendra Cruz, MPH; and Mitzi Ramírez, BS); (2) The National Institute of Public Health, Cuernavaca, Mexico (Anabel Rojas-Carmona, MPH; Moisés Ortíz-Vargas; María Guadalupe Flores-Escartín; Aurelio Cruz, MD, PhD; Janet Pacheco, MD, MPH; Steyce Aguilar, RN; and Rosalba García, RN); and (3) The National Autonomous University of Mexico, Mexico City, Mexico (Guadalupe Ponciano-Rodríguez, PhD). The authors also acknowledge the software development teams from (1) *Viva la Gráfica*, Bogota, Colombia (Catalina Lineros and David Gómez-Bonilla); (2) *In Touch Productions*, Mexico City, Mexico (Marco Polo Ramírez-Bonilla and René Navarro); (3) *Assertia Solutions*, Bahia Blanca, Argentina (Ariel Trellini and Fabricio Ziliotto); (4) *Agile Health Inc*, Chicago, US (Gary Slagle, Scott Werntz, and Vova Kagan); and (5) *Auronix*, Mexico City, Mexico (Martha Cepeda, Isadora Hernández, Blanca Ortega, Alejandra Rivas, and Carlos Piña). Finally, the authors also acknowledge Ernesto Suarez and Juan Núñez-Guadarrama (communication experts) and the administrative team from The National Institute of Public Health (Pilar Cuellar-Rodríguez, MBA; Edith Reyes-Martínez, MBA; and Uvaldo Marcelo-Landa).

Conflicts of Interest

None declared.

References

1. Organización Panamericana de la Salud, Instituto Nacional de Salud Pública. [Global Survey of Smoking in Adults Mexico 2015]. Cuernavaca, México: Organización Panamericana de la Salud; 2017.
2. World Health Organization. Country Statistics and Global Health Estimates: Mexico. Mexico: World Health Organization; 2015.
3. Organización Panamericana de la Salud, Instituto Nacional de Salud Pública. Global Survey of Smoking in Adults Mexico 2009. Cuernavaca, México: Organización Panamericana de la Salud; 2009.
4. Saenz-de-Miera B, Thrasher JF, Chaloupka FJ, Waters HR, Hernandez-Avila M, Fong GT. Self-reported price of cigarettes, consumption and compensatory behaviours in a cohort of Mexican smokers before and after a cigarette tax increase. Tob Control 2010 Dec;19(6):481-487 [FREE Full text] [doi: 10.1136/tc.2009.032177] [Medline: 20870740]
5. Sáenz de Miera Juárez B, Thrasher JF, Reynales Shigematsu LM, Hernández Ávila M, Chaloupka FJ. Tax, price and cigarette brand preferences: a longitudinal study of adult smokers from the ITC Mexico Survey. Tob Control 2014 Mar;23(Suppl 1):i80-i85 [FREE Full text] [doi: 10.1136/tobaccocontrol-2012-050939] [Medline: 24114563]
6. Thrasher JF, Pérez-Hernández R, Swayampakala K, Arillo-Santillán E, Bottai M. Policy support, norms, and secondhand smoke exposure before and after implementation of a comprehensive smoke-free law in Mexico city. Am J Public Health 2010 Sep;100(9):1789-1798. [doi: 10.2105/AJPH.2009.180950] [Medline: 20466952]
7. Thrasher JF, Nayeli Abad-Vivero E, Sebrié EM, Barrientos-Gutierrez T, Boado M, Yong HH, et al. Tobacco smoke exposure in public places and workplaces after smoke-free policy implementation: a longitudinal analysis of smoker cohorts in Mexico and Uruguay. Health Policy Plan 2013 Dec;28(8):789-798 [FREE Full text] [doi: 10.1093/heapol/czs118] [Medline: 23172895]
8. Thrasher JF, Pérez-Hernández R, Arillo-Santillán E, Barrientos-Gutiérrez I. [Towards informed tobacco consumption in Mexico: effect of pictorial warning labels in smokers]. Salud Publica Mex 2012 Jun;54(3):242-253 [FREE Full text] [Medline: 22689162]
9. Thrasher JF, Abad-Vivero EN, Huang L, O'Connor RJ, Hammond D, Bansal-Travers M, et al. Interpersonal communication about pictorial health warnings on cigarette packages: policy-related influences and relationships with smoking cessation attempts. Soc Sci Med 2016 Dec;164:141-149 [FREE Full text] [doi: 10.1016/j.socscimed.2015.05.042] [Medline: 26092600]
10. Swayampakala K, Thrasher JF, Hammond D, Yong H, Bansal-Travers M, Krugman D, et al. Pictorial health warning label content and smokers' understanding of smoking-related risks-a cross-country comparison. Health Educ Res 2015 Feb;30(1):35-45 [FREE Full text] [doi: 10.1093/her/cyu022] [Medline: 24848554]
11. Pérez-Hernández R, Thrasher JF, Rodríguez-Bolaños R, Barrientos-Gutiérrez I, Ibáñez-Hernández NA. [Tobacco advertising and promotions: changes in reported exposure in a cohort of Mexican smokers]. Salud Publica Mex 2012 Jun;54(3):204-212 [FREE Full text] [Medline: 22689158]

12. Blanco A, Sandoval R, Martínez-López L, Caixeta R. [Ten years of the WHO Framework Convention on Tobacco Control: progress in the Americas]. Salud Publica Mex 2017;59(Suppl 1):117-125 [FREE Full text] [doi: 10.21149/8682] [Medline: 28658460]
13. [Federal Government of Mexico, National Commission of Social Protection in Health, Seguro Popular]. 2012. [Universal catalog of health services 2012] URL: http://www.censida.salud.gob.mx/descargas/biblioteca/documentos/CAUSES2012.pdf [accessed 2019-04-04] [WebCite Cache ID 77NzfZFkO]
14. [Mexican Social Security Institute]. 2018. [Basic Medicine Chart of the Mexican Institute of Social Security] URL: http://www.imss.gob.mx/sites/all/statics/pdf/cuadros-basicos/CBM.pdf [accessed 2019-04-04] [WebCite Cache ID 77NzkFrxd]
15. [Institute of Security and Social Services of State Workers]. 2018. [Clinics to Stop Smoking] URL: http://sgm.issste.gob.mx/medica/clinicaspdejarcigarro_principal.php [accessed 2019-04-04] [WebCite Cache ID 77NztdeBz]
16. [Health Secretary]. [Prevention, Diagnosis and Treatment of Tobacco Consumption and Extraneous Smoke, in the First Level of Care]. Mexico: Secretaría de Salud; 2009.
17. [Mexican Guide for the Treatment of Smoking. A National Council of Experts. Pneumology and Thoracic Surgery]. Mexico: [Mexican Society of Pulmonology and Thoracic Surgery]; 2005.
18. Ponciano-Rodríguez G. The urgent need to change the current medical approach on tobacco cessation in Latin America. Salud Publica Mex 2010;52(Suppl 2):S366-S372 [FREE Full text] [Medline: 21243210]
19. Swayampakala K, Thrasher J, Carpenter MJ, Shigematsu LM, Cupertino A, Berg CJ. Level of cigarette consumption and quit behavior in a population of low-intensity smokers--longitudinal results from the International Tobacco Control (ITC) survey in Mexico. Addict Behav 2013 Apr;38(4):1958-1965 [FREE Full text] [doi: 10.1016/j.addbeh.2012.12.007] [Medline: 23380491]
20. Etter JF. The psychological determinants of low-rate daily smoking. Addiction 2004 Oct;99(10):1342-1350. [doi: 10.1111/j.1360-0443.2004.00839.x] [Medline: 15369573]
21. Schane RE, Ling PM, Glantz SA. Health effects of light and intermittent smoking: a review. Circulation 2010 Apr 06;121(13):1518-1522 [FREE Full text] [Medline: 20368531]
22. Op den Akker H, Cabrita M, Op den Akker R, Jones VM, Hermens HJ. Tailored motivational message generation: a model and practical framework for real-time physical activity coaching. J Biomed Inform 2015 Jun;55:104-115 [FREE Full text] [doi: 10.1016/j.jbi.2015.03.005] [Medline: 25843359]
23. Bock BC, Rosen RK, Barnett NP, Thind H, Walaska K, Foster R, et al. Translating behavioral interventions onto mHealth platforms: developing text message interventions for smoking and alcohol. JMIR Mhealth Uhealth 2015;3(1):e22 [FREE Full text] [doi: 10.2196/mhealth.3779] [Medline: 25714907]
24. Haug S, Schaub MP, Venzin V, Meyer C, John U. Efficacy of a text message-based smoking cessation intervention for young people: a cluster randomized controlled trial. J Med Internet Res 2013;15(8):e171 [FREE Full text] [doi: 10.2196/jmir.2636] [Medline: 23956024]
25. Ybarra ML, Holtrop JS, Bağci BA, Emri S. Design considerations in developing a text messaging program aimed at smoking cessation. J Med Internet Res 2012;14(4):e103 [FREE Full text] [doi: 10.2196/jmir.2061] [Medline: 22832182]
26. Spears CA, Bell SA, Scarlett CA, Anderson NK, Cottrell-Daniels C, Lotfalian S, et al. Text messaging to enhance mindfulness-based smoking cessation treatment: program development through qualitative research. JMIR Mhealth Uhealth 2019 Jan 07;7(1):e11246 [FREE Full text] [doi: 10.2196/11246] [Medline: 30617043]
27. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
28. Kong G, Ells DM, Camenga DR, Krishnan-Sarin S. Text messaging-based smoking cessation intervention: a narrative review. Addict Behav 2014 May;39(5):907-917 [FREE Full text] [doi: 10.1016/j.addbeh.2013.11.024] [Medline: 24462528]
29. Scott-Sheldon LA, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text messaging-based interventions for smoking cessation: a systematic review and meta-analysis. JMIR Mhealth Uhealth 2016;4(2):e49 [FREE Full text] [doi: 10.2196/mhealth.5436] [Medline: 27207211]
30. The Competitive Intelligence Unit. 2015. Market research, business opportunity analysis and impact of regulation in telecommunications URL: http://www.the-ciu.net/ [accessed 2019-04-04] [WebCite Cache ID 77O4it6D9]
31. [National Institute of Statistic and Geography]. 2017. URL: https://www.inegi.org.mx/ [accessed 2019-04-04] [WebCite Cache ID 77O6XesA5]
32. [National Cancer Institute]. SmokefreeTXT URL: https://smokefree.gov/smokefreetxt/ [accessed 2019-04-04] [WebCite Cache ID 77O6fxTSS]
33. Ortiz-León M, Thrasher J, Morales-Romero J, De San Jorge Cárdenas X. [Transcultural adaptation of messages via mobile phones as support to quit smoking]. Rev Salud Publica 2017;21(2). [doi: 10.31052/853.1180.v21.n2.16868]
34. Cartujano-Barrera F, Cruz K, Cupertino A, Flores-Escartín M, Ortíz-Vargas M, Perales J, et al. e-Decídete: Herramienta Interactiva para Cesación de Consumo de Tabaco. Mexico: Instituto Nacional del Derecho de Autor. Public Registry of Copyright; 2017.
35. Cupertino AP, Richter K, Cox LS, Garrett S, Ramirez R, Mujica F, et al. Feasibility of a Spanish/English computerized decision aid to facilitate smoking cessation efforts in underserved communities. J Health Care Poor Underserved 2010 May;21(2):504-517. [doi: 10.1353/hpu.0.0307] [Medline: 20453353]

36. Cupertino AP, Cartujano-Barrera F, Perales J, Formagini T, Rodríguez-Bolaños R, Ellerbeck EF, et al. "Vive Sin Tabaco… ¡Decídete!" feasibility and acceptability of an e-Health smoking cessation informed decision-making tool integrated in primary healthcare in Mexico. Telemed J E Health 2018 Jul 26 Epub ahead of print. [doi: 10.1089/tmj.2017.0299] [Medline: 30048208]
37. Ponciano-Rodríguez G, Reynales-Shigematsu L, Rodríguez-Bolaños R, Pruñonosa-Santana J, Cartujano-Barrera F, Cupertino A. Enhancing smoking cessation in Mexico using an e-Health tool in primary healthcare. Salud Publica Mex 2018;60(5):549-558. [doi: 10.21149/9348] [Medline: 30550116]
38. Cartujano-Barrera F, Cruz K, Cupertino A, Perales J, Ramírez-Mantilla M, Ramírez-Mantilla M, et al. e-Decídete: Biblioteca de Mensajes de Texto para Cesación de Consumo de Tabaco. Mexico: Instituto Nacional del Derecho de Autor. Public Registry of Copyright; 2017.
39. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
40. Organisation for Economic Co-operation and Development. 2012. Review of Telecommunication Policy and Regulation in Mexico URL: https://www.oecd.org/sti/broadband/50550219.pdf [accessed 2019-04-04] [WebCite Cache ID 77O7GxhbT]
41. University of Massachusetts Medical School. Center for Tobacco Treatment Research and Training URL: https://www.umassmed.edu/tobacco/ [accessed 2019-04-04] [WebCite Cache ID 77O7NXirJ]
42. Agile Health. URL: https://agilehealth.com/ [accessed 2019-04-04] [WebCite Cache ID 77O7YSH8w]
43. Auronix. URL: https://www.auronix.mx/ [accessed 2019-04-04] [WebCite Cache ID 77O7iPopj]
44. Heatherton T, Kozlowski L, Frecker R, Fagerström KO. The Fagerström test for nicotine dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
45. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
46. Kim S. Overview of cotinine cutoff values for smoking status classification. Int J Environ Res Public Health 2016 Dec 14;13(12) [FREE Full text] [doi: 10.3390/ijerph13121236] [Medline: 27983665]
47. Atun R, de Andrade LO, Almeida G, Cotlear D, Dmytraczenko T, Frenz P, et al. Health-system reform and universal health coverage in Latin America. Lancet 2015 Mar 28;385(9974):1230-1247. [doi: 10.1016/S0140-6736(14)61646-9] [Medline: 25458725]
48. Urquieta-Salomón JE, Villarreal HJ. Evolution of health coverage in Mexico: evidence of progress and challenges in the Mexican health system. Health Policy Plan 2016 Feb 29;31(1):28-36. [doi: 10.1093/heapol/czv015] [Medline: 25823751]
49. Fiore M, Jaén C, Baker T. Treating Tobacco Use and Dependence: 2008 Update. In: Clinical Practice Guideline. Rockville, MD: US Department of Health and Human Services. Public Health Service; 2008.

Abbreviations

CPD: cigarettes per day
NRT: nicotine replacement therapy

Edited by G Eysenbach; submitted 11.10.18; peer-reviewed by B Pineiro, A Galil; comments to author 26.12.18; revised version received 14.02.19; accepted 17.02.19; published 25.04.19.

Please cite as:
Cupertino AP, Cartujano-Barrera F, Ramírez M, Rodríguez-Bolaños R, Thrasher JF, Pérez-Rubio G, Falfán-Valencia R, Ellerbeck EF, Reynales-Shigematsu LM
A Mobile Smoking Cessation Intervention for Mexico (Vive sin Tabaco... ¡Decídete!): Single-Arm Pilot Study
JMIR Mhealth Uhealth 2019;7(4):e12482
URL: http://mhealth.jmir.org/2019/4/e12482/
doi:10.2196/12482
PMID:31021326

Original Paper

A Mobile Phone–Based Approach for Hearing Screening of School-Age Children: Cross-Sectional Validation Study

Yuan-Chia Chu[1,2,3*], MS; Yen-Fu Cheng[4,5,6,7*], MD, PhD; Ying-Hui Lai[8], PhD; Yu Tsao[9], PhD; Tzong-Yang Tu[4,6], MD, PhD; Shuenn Tsong Young[10], PhD; Tzer-Shyong Chen[11], PhD; Yu-Fang Chung[12], PhD; Feipei Lai[1,13,14*], PhD; Wen-Huei Liao[4,6], MD, PhD

[1]Graduate Institute of Biomedical Electronics & Bioinformatics, National Taiwan University, Taipei, Taiwan
[2]Information Management Office, Taipei Veterans General Hospital, Taipei City, Taiwan
[3]Big Data Center, Taipei Veterans General Hospital, Taipei City, Taiwan
[4]Department of Otolaryngology-Head and Neck Surgery, Taipei Veterans General Hospital, Taipei, Taiwan
[5]Department of Medical Research, Taipei Veterans General Hospital, Taipei, Taiwan
[6]Department of Otolaryngology-Head and Neck Surgery, School of Medicine, National Yang-Ming University, Taipei, Taiwan
[7]Department of Speech Language Pathology and Audiology, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan
[8]Department of Biomedical Engineering, National Yang-Ming University, Taipei, Taiwan
[9]Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan
[10]Holistic Education Center, Mackay Medical College, Taipei, Taiwan
[11]Department of Information Management, Tunghai University, Taipei, Taiwan
[12]Department of Electrical Engineering, Tunghai University, Taipei, Taiwan
[13]Department of Computer Science & Information Engineering, National Taiwan University, Taipei, Taiwan
[14]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan
[*]these authors contributed equally

Corresponding Author:
Wen-Huei Liao, MD, PhD
Department of Otolaryngology-Head and Neck Surgery
Taipei Veterans General Hospital
No 201, Sec 2, Shipai Rd, Beitou District
Taipei, 11217
Taiwan
Phone: 886 938102333
Email: whliaovictor@gmail.com

Abstract

Background: Pure-tone screening (PTS) is considered as the gold standard for hearing screening programs in school-age children. Mobile devices, such as mobile phones, have the potential for audiometric testing.

Objective: This study aimed to demonstrate a new approach to rapidly screen hearing status and provide stratified test values, using a smartphone-based hearing screening app, for each screened ear of school-age children.

Method: This was a prospective cohort study design. The proposed smartphone-based screening method and a standard sound-treated booth with PTS were used to assess 85 school-age children (170 ears). Sound-treated PTS involved applying 4 test tones to each tested ear: 500 Hz at 25 dB and 1000 Hz, 2000 Hz, and 4000 Hz at 20 dB. The results were classified as *pass* (normal hearing in the ear) or *fail* (possible hearing impairment). The proposed smartphone-based screening employs 20 stratified hearing scales. Thresholds were compared with those of pure-tone average (PTA).

Results: A total of 85 subjects (170 ears), including 38 males and 47 females, aged between 11 and 12 years with a mean (SD) of 11 (0.5) years, participated in the trial. Both screening methods produced comparable *pass* and *fail* results (pass in 168 ears and fail in 2 ears). The smartphone-based screening detected moderate or worse hearing loss (average PTA>25 dB) accurately. Both the sensitivity and specificity of the smartphone-based screening method were calculated at 100%.

Conclusions: The results of the proposed smartphone-based self-hearing test demonstrated high concordance with conventional PTS in a sound-treated booth. Our results suggested the potential use of the proposed smartphone-based hearing screening in a school-age population.

(JMIR Mhealth Uhealth 2019;7(4):e12033) doi:10.2196/12033

KEYWORDS

hearing tests; telemedicine; mobile apps; audiometry, pure-tone

Introduction

Background

Worldwide, more than 466 million (over 5%) people, including 34 million children, are estimated to have a hearing impairment. Hearing impairment is difficult to monitor because of the limited availability of testing equipment and trained specialists in many developing countries [1,2]. Unidentified hearing impairment has been one of the most common disorders in school-age children [1,3,4]. Several studies have shown that children with hearing impairments remain unidentified, and if they do not receive treatment, these children may experience a delay in the acquisition of speech and language skills [5-7]. The burden of hearing loss is the greatest in developing countries and more than 80% of people with hearing loss live in these areas [3,8]. However, hearing care services in these areas are either very limited or absent altogether [8,9]. Early detection and early intervention are key factors in reducing the impact of hearing impairment on the development and future achievement in school-age children [10].

Pure-tone screening (PTS) is considered as the gold standard for hearing screening programs for school-age children [11,12]. PTS is usually administered by a hearing professional or a nurse, using a portable instrument that produces a limited set of test stimuli often at a predetermined level between 20 and 40 dB hearing level (HL), depending on the age of the group being tested [2]. Current school-based hearing screening protocols have not been standardized, and numerous screening criteria vary according to the guidelines of the agency, state, or country. For example, the American Speech-Language-Hearing Association (ASHA) and the American Academy of Audiology published professional recommendations that specify screening at 20 dB at frequencies of 1000 Hz, 2000 Hz, and 4000 Hz [2,4]. In 2003, the American Academy of Pediatrics (AAP) also suggested screening at 20 dB at frequencies of 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz [3]. One major drawback of the current hearing screening methods is the lack of sensitivity and specificity in determining hearing ability and indicating hearing loss candidacy. As a result, conventional PTS provides only a pass or fail result for each screened ear and lacks hearing status assessment and further stratified test values as provided by tools such as the Landolt C eye chart for follow-ups [6,7].

The Hearing Scale Test (HST) is a novel hearing screening method derived from the consecutive hearing screening procedures for approaching the current hearing status of each screened ear of children [5,8]. The HST employs stratified hearing scales containing 4 test tones (500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz), where adjacent scales differ from each other by 5 dB (Table 1). In addition to the pass/fail results that most PTS-based screening programs offer, the HST also offers current hearing status and provides stratified test values that can be recorded for follow-ups. Our previous studies have shown that the automated audiometry devices based on personal computers built with the hearing protocol of the HST, which offer a user-friendly interface and measure hearing threshold values, are useful for monitoring progressive hearing changes in school-age children [5,8].

Automated audiometry devices have demonstrated that comparable hearing threshold values, compared with those obtained by automated audiometry, such as computer-assisted audiometry [9,10,12] or smartphone-based audiometry [11,13-20], and results obtained by audiologists using conventional manual audiometry can be achieved. Automated audiometry devices using mobile phone require the use of earphones, and given the huge variety of combinations of earphones and mobile phone, standardized and calibrated software and devices continue to be the key for performing reliable hearing tests [15,16,21-27]. Apple, iOS-based devices provide standardized hardware and software components; therefore, most apps can potentially be universally shared with all iOS-based device models [19]. Numerous audiometric apps have been developed for hearing assessments on Apple mobile devices [19,21,28], most of which calibrate mobile devices using a biological method to determine a reference sound level in relation to the hearing threshold of normal people [11,15,22]. To avoid possible variability and inconsistency caused by biological calibration, our previous study has shown that reference equivalent threshold sound pressure levels (RETSPLs) represent a reliable calibration method for output levels across different Apple mobile devices with bundled earphones [23].

Objectives

In this study, we developed an iOS-based smartphone hearing test app *Ear Scale* and evaluated its performance and feasibility as a hearing screening program for school-age children. We investigated the accuracy of the hearing tests conducted on mobile devices calibrated by RETSPLs for Apple EarPod [23]. We compared the performance of the smartphone-based automated hearing screening with that of audiologist-assisted pure-tone audiometry (PTA) performed in a sound-treated booth. Different screening protocols, including those suggested by the AAP and ASHA, were also compared with the built-in HST protocol of the Ear Scale app [15,16,19,21-27,29].

Table 1. Stimulus levels in dB hearing level for tested frequencies in the proposed Hearing Scale Test.

Stimulation level	Hearing Scale Test									
	Normal (pure-tone audiometry ≤25 dB)					Possible hearing impairment (pure-tone audiometry >25 dB)				
	S^a_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Frequency (Hz)										
1000 Hz, 2000 Hz, and 4000 Hz	0	5	10	15	20	25	30	35	40	45
500 Hz	5	10	15	20	25	30	35	40	45	50

[a]S: stratified hearing scale.

Methods

Study Setting and Participants

This prospective cohort study was conducted at an elementary school in Taipei, Taiwan. We recruited children from grades 5 and 6, aged between 11 and 12 years. A total of 85 children (38 boys and 47 girls) were enrolled, with 170 ears tested. The trial was approved by the Institutional Review Board of Taipei Veterans General Hospital (2017-10-003CC). Written informed consent was collected by the teachers from the parents, before the scheduled date of the hearing screening tests. After instruction by the researchers, each child, in a random order, underwent smartphone-based and booth-based hearing screening consecutively. The smartphone-based hearing screening procedures were performed in a quiet room in the school. Before the hearing screening, the students were taught how to wear the headphones and push a button when hearing the tone. The air conditioner was turned off during the measurements to reduce ambient noise, the level of which was monitored every 30 min by a sound level meter to ensure an ambient noise level of less than 50 dB at test frequencies of 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz.

Measurements

Pure-Tone Screening Procedures in a Sound-Treated Booth

The audiologist manually controlled a GrasonStadler GSI 18 screening audiometer that was used with a Telephonics TDH-39 supraaural earphones previously calibrated according to International Organization for Standardization (ISO) 389-1. A *pass* result for an ear indicated that the child responded correctly to all 4 test tones. If the child did not respond to all 4 test tones after 2 consecutive testing procedures, then the ear was assigned a *fail* result. PTA hearing thresholds of more than 25 dB at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz in the sound-treated booth were designated as *hearing impairment*.

iOS Automated Audiometry App

The iOS-based automated *Ear Scale* app (version 2.0) was developed to perform pure-tone air conduction hearing testing and was made freely accessible as a download through the Apple iTunes store in 2018. The HST, a new modified hearing screening method derived from consecutive hearing screening procedures to assess the current hearing status of each screened ear of children, was used to determine the hearing threshold [5] of each screened ear in children (Table 1). The test tones were 1.5 seconds in duration, whereas the silent interval between successive tones randomly varied between 2 and 3 seconds, and depending on the user response, the sound intensity was changed in steps of 5 dB semiautomatically [19]. The test tone's amplitude was modulated with a depth of 100% [11]. At the end of the test, an audiogram was displayed, which could be saved on the device (Figure 1). The Ear Scale app involved computerized self-determination of the lowest audible sound generated by the mobile device. The computerized smartphone-based audiometer presented the 4 test tones of the HST at the appropriate stimulus levels semiautomatically, as shown in Figure 2. The Ear Scale app started with a hearing scale of 25 dB (S_5; Figure 2). The 4 test tones were automatically presented in a fixed order: 1000 Hz, 2000 Hz, 4000 Hz, and 500 Hz. If the child responded correctly to all test tones of a particular hearing scale, then the test stimulus level was decreased (corresponding to hearing scales decreasing from S_4 to S_1) until the child did not respond to any of the 4 test tones; otherwise, the test stimulus level was increased (corresponding to hearing scales increasing from S_6 to S_{10}; Figure 2). The minimum audible hearing scale on the HST indicated the stimulus level at which the child responded correctly to all 4 test tones. If the child did not respond correctly to hearing scale S_{10}, then the result was designated as *no response* (NR). Scales S_1 to S_5 of the HST are equivalent to a PTS *pass* result, whereas scales S_6 to S_{10} and NR are equivalent to a PTS *fail* result (Figure 2). The tests on mobile devices were conducted twice, test and retest.

iOS Automated Audiometry Calibration

Calibration of iOS-based devices with Apple EarPod RETSPLs was described in detail in a previous paper [23]. Briefly, the RETSPL method of the hearing self-test carried out on mobile devices with calibrated bundled headphones is used when calibrating audiometric equipment to a hearing threshold of 0 dB at various frequencies. Pure-tone stimuli at 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz were generated on the iOS mobile device and delivered by the Apple EarPods. The KEMAR manikin was developed to meet the needs of hearing aid designers and other manikin users. The EarPods were placed in the left and right pinna of the KEMAR manikin for eardrum-pressure recording. Hearing thresholds were determined by the ascending method described in ISO 8253-1 [24], where the step size was set to 1 dB. The initial level was set at 10 dB below the lowest subject response level, which was

predetermined using a conventional audiometer. Subjects were instructed to respond when they heard the stimulus. Final thresholds were determined using a 2-down, 1-up adaptive staircase procedure [25] after 3 reversals. All devices were standardized by setting the user-controllable volume to 100% of its maximum limit. The maximum difference between right and left EarPods was less than 1 dB and the maximum difference among devices (iPhone 5s, iPhone 6, iPhone 6 Plus, iPhone 7, iPhone 7 Plus, and iPad mini) was less than 1.5 dB with output levels across 5 EarPods between 250 and 8000 Hz on a single device (iPad mini 4). The maximum difference was less than 1.0 dB. The microphone of the ear simulators and the electrical and acoustical measurement systems were calibrated using a GRAS model 42AA pistonphone. The output levels of the EarPods at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz were calibrated in units of dB sound pressure level (SPL) when the volume of the Apple mobile device was set to maximum. The output level (dB) of the pure-tone sound corresponding to each hearing test frequency is similar to that of the apparatus previously described for sound output calibration [19,23]. Apple EarPod RETSPLs have stable output levels between right and left EarPods, which can be applied to calibrate output levels of various Apple mobile devices with EarPods [23].

Statistical Analysis

For hearing screening, the presence or absence of hearing loss (PTA>25 dB) in each ear was determined by sound-treated booth audiometry. The results from the Ear Scale app were compared with the threshold obtained from sound-treated booth PTA measurement. These data were entered into 2×2 tables to calculate the sensitivity, specificity, positive predictive value, and negative predictive value. The hearing scale obtained from the Ear Scale app and the corresponding mean pure-tone threshold obtained from the sound-treated booth are shown by a box plot (Figure 2). The corresponding pure-tone threshold of each grade of the HST is shown by a box plot (Figure 3). The correlation coefficient was calculated to estimate the average correlation coefficient across both methods. The Kruskal-Wallis test was performed to determine significance. Analyses were performed using the SPSS version 23.0 (SPSS Inc) and Microsoft Excel version 2016 (Microsoft Inc) for personal computers. *P* values less than .05 were considered statistically significant. The PTA thresholds at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz were summarized as the mean (SD) values (Table 1).

Figure 1. Screenshot of the Ear Scale app includes instructions for the testers and the hearing test process.

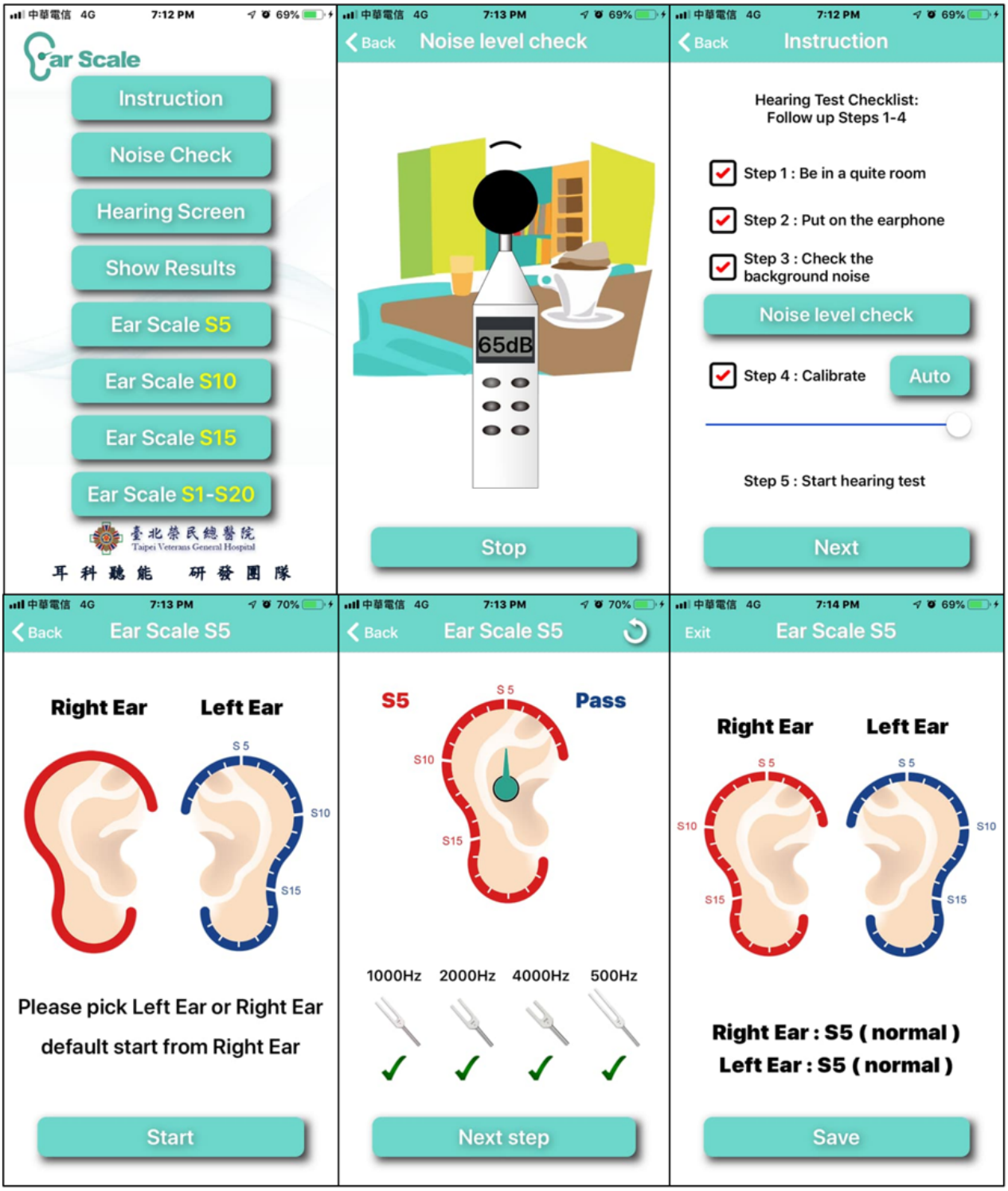


Figure 2. The computerized smartphone-based hearing screening flow diagram. S: stratified hearing scale.

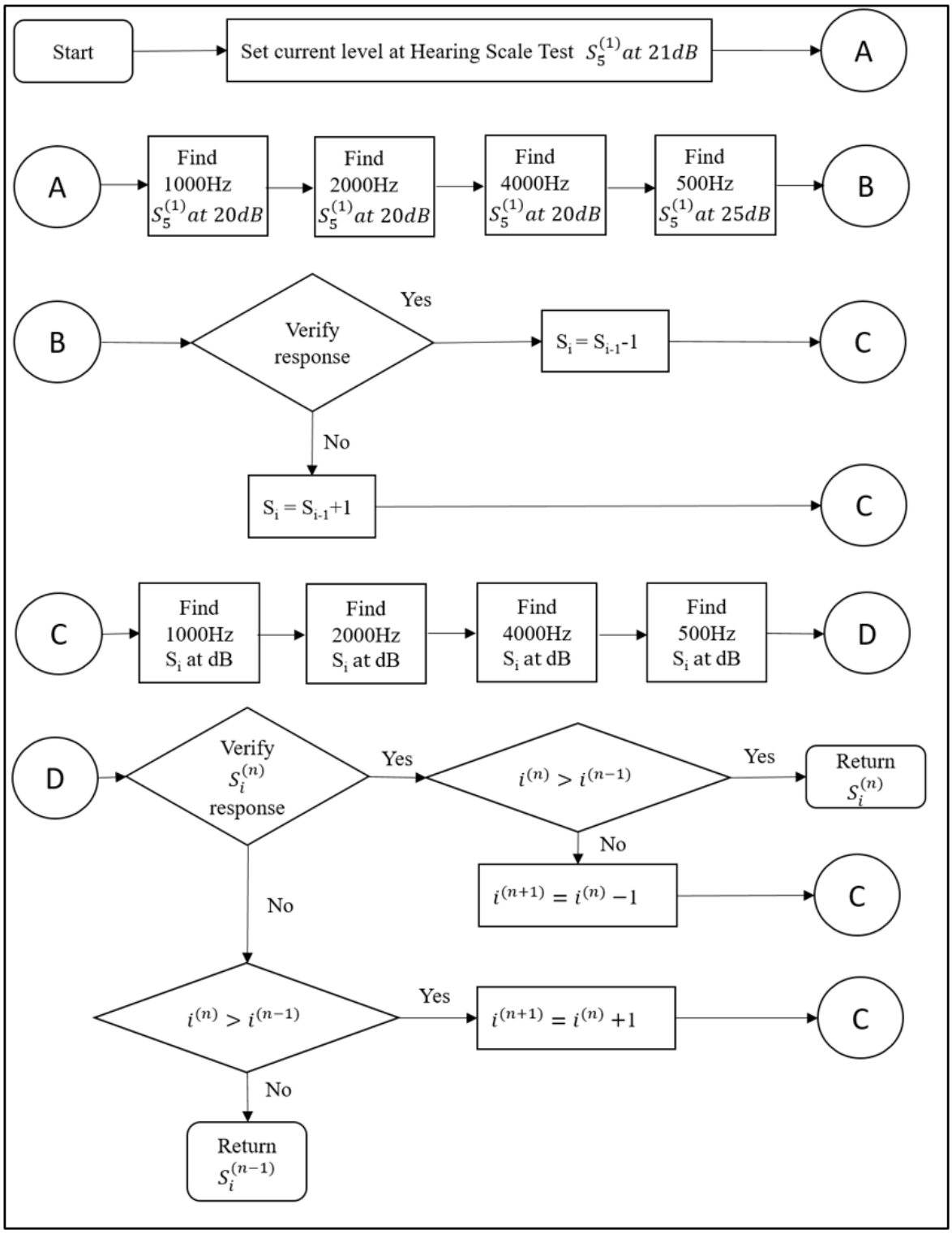
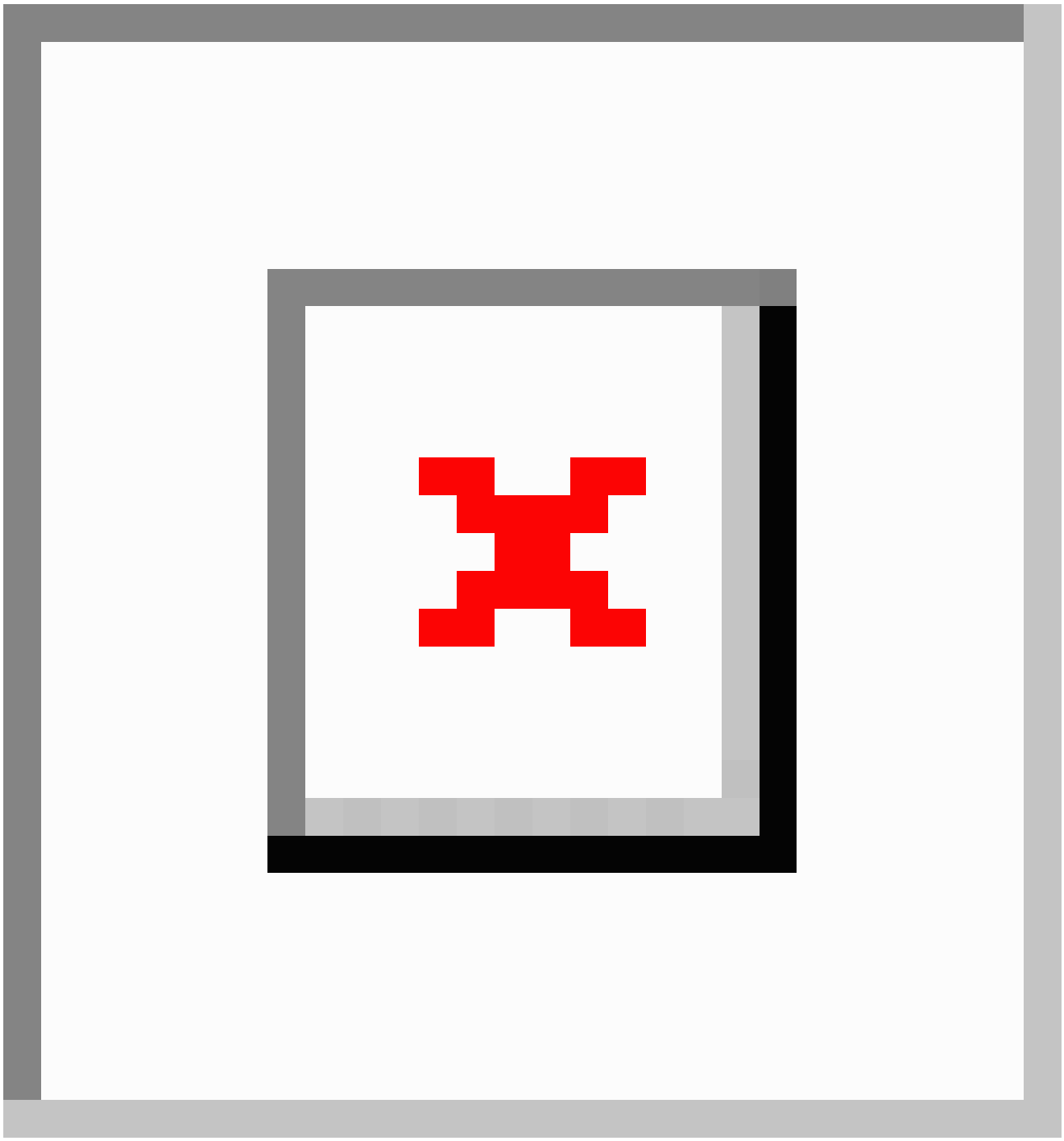


Figure 3. Box plots of the hearing results of right ears and left ears obtained from the Ear Scale app in relation to those obtained from pure-tone screening. The box includes the median (heavy line) and represents the first and third quartiles, whereas the vertical bar indicates the SD. Blue lines represent best-fit linear regressions of the means of the boxes, whereas the gray areas around the line represent the 95% CI of the model (P<.05, differences were found between groups). S: stratified hearing scale.



Results

Comparing 2 Hearing Screening Methods: Conventional Pure-Tone Screening Versus the Ear Scale App

Of the 170 ears tested by sound-treated booth PTA, 98.8% (168/170) and 1.2% (2/170) were assigned *pass* and *fail* results, respectively. Similarly, of the 170 ears tested by the Ear Scale app, 98.8% (168/170) and 1.2% (2/170) of the tests were assigned *pass* and *fail* results, respectively (Table 2). The results using these 2 methods of hearing screening were calculated in a 2×2 table to determine the sensitivity, specificity, positive predictive value, and negative predictive value (Figure 3). In addition to the dichotomous *pass* or *fail* results, the Ear Scale app provided stratified hearing scales for each screened ear. The results of 84 left ears with a *pass* result were stratified as 0 dB (S_1) of 13% (11/84), 5 dB (S_2) of 38% (32/85), 15 dB (S_3) of 33% (28/85), 20 dB (S_4) of 11% (9/85), and 25 dB (S_5) of 4% (4/85), whereas *fail* results were stratified as 35 dB (S_7) of 1% (1/85). Similarly, 84 *pass* results and 1 *fail* result for right ears were also further stratified. The results of 168 *pass* ears and 2 *fail* ears are pooled and shown in Table 3.

Table 2. Participants' demographics and hearing impairment candidacy (as graded by the pure-tone screening and Hearing Scale Test).

Variables	Statistics
Participants, n	85
Age (years), mean (SD)	11 (0.5)
Gender, n	
Male	38
Female	47
Pure-tone screening, n	
≤25 dB (normal)	168
26-40 dB (mild loss)	2
41-55 dB (moderate loss)	0
56-70 dB (moderate to severe loss)	0
71-90 dB (severe loss)	0
≥91 dB (profound loss)	0
Ear Scale app with the Hearing Scale Test, n	
≤ 25 dB (S^a_1-S_5, normal)	168
>25 dB (S_6-S_{10}, hearing loss candidate)	2

[a]S: stratified hearing scale.

Validation of the Built-In Hearing Scale Test Hearing Screening Protocol for the Ear Scale App

As the HST was used in our Ear Scale app for the default screening protocol, we also compared the HST with other popular protocols, including those suggested by the AAP and ASHA. The Ear Scale app was highly accurate at the tested frequencies (500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz) for all 3 screening protocols. The specificity was 100% and the sensitivity was 100% for HST (1000 Hz, 2000 Hz, and 4000 Hz at 20 dB and 500 Hz at 25 dB), 95.2% for AAP (500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz at 20 dB), and 95.2% for ASHA (500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz at 15 dB). The false-positive rate was 0% in all 3 screening protocols, whereas the false-negative rates were 0% of HST, 4.8% of AAP, and 4.8% of ASHA, respectively. A summary of the results from all 3 tested screening protocols is provided in Table 4.

Accuracy of Ear Scale App Calibration at All Hearing Scale Test Grades

The correlation between the 2 measurements by utilizing the Ear Scale app in a quiet conference room and the clinical audiometer in a sound-treated room was significant at the .01 level (Figure 3). Statistically significant differences were found in all tested HST scales (S_1, S_2, S_3, S_4, and S_5) in right ears and left ears (Kruskal-Wallis test with 5 degrees of $P<.01$; Figure 3). Similarly, the pooled data from both ears also showed a significant difference, indicating the usefulness of the proposed Ear Scale app in not only distinguishing ears with *pass* or *fail* results but also providing an accurate measurement of the HL of school children.

Table 3. The Hearing Scale Test and the mean difference between thresholds (dB) for the Ear Scale app and sound-treated booth (N=170 ears).

Ear Scale app with Hearing Scale Test	Sound-treated booth in pure-tone screening	
	Mean (SD)	n
Left ear (mean thresholds)		
≤5 dB (S^a_1)	4 (3.14)	11
6-10 dB (S_2)	7 (2.7)	32
11-15 dB (S_3)	8 (2.9)	28
16-20 dB (S_4)	11 (4.2)	9
21-25 dB (S_5)	14 (4.3)	4
26-30 dB (S_6)	0	0
31-35 dB (S_7)	31 (NaN[b])	1
36-40 dB (S_8)	0	0
41-45 dB (S_9)	0	0
46-50 dB (S_{10})	0	0
Right ear (mean thresholds)		
≤5 dB (S_1)	6 (2.1)	5
6-10 dB (S_2)	7 (3.3)	26
11-15 dB (S_3)	10 (2.6)	31
16-20 dB (S_4)	11 (3.8)	18
21-25 dB (S_5)	11 (2.6)	4
26-30 dB (S_6)	0	0
31-35 dB (S_7)	0	0
36-40 dB (S_8)	36 (NaN)	1
41-45 dB (S_9)	0	0
46-50 dB (S_{10})	0	0
Both ears (mean thresholds)		
≤5 dB (S_1)	5 (2.9)	16
6-10 dB (S_2)	7 (3.0)	58
11-15 dB (S_3)	9 (2.8)	59
16-20 dB (S_4)	11 (3.8)	27
21-25 dB (S_5)	12 3.6)	8
26-30 dB (S_6)	0	0
31-35 dB (S_7)	31 (NaN)	1
36-40 dB (S_8)	36 (NaN)	1
41-45 dB (S_9)	0	0
46-50 dB (S_{10})	0	0

[a]S: stratified hearing scale.

[b]NaN: not a number.

Table 4. Comparison of the hearing screening protocols for both ears of all subjects participating in the study.

Results	Hearing screening protocols		
	Hearing Scale Test, %	American Academy of Pediatrics, %	American Speech-Language-Hearing Association, %
Sensitivity	100	95.2	95.2
Specificity	100	100	100
False-positive	0	0	0
False-negative	0	4.8	4.8

Discussion

Principal Findings

The findings from this study support the use of the Ear Scale app in smartphone-based hearing screening of school children. To the best of our knowledge, this is the first report proposing a method for stratifying hearing test results on a smartphone and then using it for hearing screening in school children. As hearing screening is useful for detecting hearing impairment in the school system [26], we developed the Ear Scale app to evaluate school children's HL ranges on the basis of 20 stratified hearing scales, that is, 5 dB (S_1) to 100 dB (S_{20}), plus an NR result. Our Ear Scale 25 dB (S_5) menu item fit a normal hearing range, the Ear Scale 50 dB (S_{10}) menu item fit a mild hearing loss range, the Ear Scale 75 dB (S_{15}) menu item fit a moderate hearing loss range, and the Ear Scale app with the HST from 5 dB (S_1) to 100 dB (S_{20}) menu item can be customized for a wide range of hearing loss for school-age children. Conventional PTS provides a *pass* / *fail* result, and it therefore provides little information regarding a child's hearing ability. The Ear Scale app with the HST proposed in this study has 10 stratified hearing scales from 0 dB (S_1) to 45 dB (S_{10}) plus an NR result. The Ear Scale app with the HST is derived from the hearing screening concept of dichotomized test results (*pass* or *fail*), but the use of computerized hearing screening procedures and hearing scales with different test stimulus levels allows the minimum audible hearing scale to be determined. The scale determined by the Ear Scale app can present the current hearing status of each tested ear. The Ear Scale app with the HST can rapidly evaluate the hearing status of the tested ear, typically within 3 to 5 min.

Many different ear screening protocols have been established in the past [7,30], but the methods suitable for children and school-age groups have not been standardized [27,30]. The Ear Scale app described in this study has several implications for hearing screening programs. First, the built-in HST protocol stratifies the hearing scales of each screened ear, whereas PTS provides only *pass* or *fail* results (Table 2). These stratified hearing scales from 0 dB (S_1) to 45 dB (S_{10}) recorded in an initial hearing assessment can be used for further follow-up surveillance in hearing screening programs [5,8]. Second, the results of the HST show the distribution of different stratified hearing scales (representing different degrees of hearing status) of all screened ears with the same median reference standard (S_5), thus facilitating comparisons of hearing screening results among classes or schools (Table 3). The Ear Scale app with a computerized audiometer typically requires only 3 to 5 min per child, whereas PTS conducted manually requires 1 to 2 min per child. The longer testing time of the Ear Scale app is because of the stratification performed by consecutive tests to determine the minimum audible hearing scale. However, this small increase in the time spent in the test is worthwhile to achieve the goal of determining a more informative hearing status associated with the use of stratified hearing scales in the Ear Scale app.

It is projected that the smartphone subscription will increase from 5 billion in 2018 to 7.2 billion in 2024 [29], and there has been a surge of health-related smartphone apps in recent years [31-36]. Smartphone hearing screening audiometry has been widely implemented as mobile phone gained popularity, and several studies have compared hearing thresholds with standardized automated hearing thresholds obtained in a sound-treated booth [11,13,14,18,28,37-39]. However, none of these studies integrated a computerized hearing screening flow diagram with a graphical interface for school-age children. Our Ear Scale app is based on a series of distinct steps and is implemented in the form of an automated process, which improves standardization of the test procedures and therefore avoids inconsistency [40,41].

Our results indicate that the iOS-based Ear Scale app is reasonably accurate for hearing screening. The sensitivity and specificity were high (100%), whereas the false-positive (0%) and false-negative rates (0%) were low when the hearing tests were performed in a quiet room in the school library, ensuring an ideal test for hearing screening. The Ear Scale app was also found to be highly accurate in testing several hearing screening protocols in addition to the built-in HST [5], including those recommended by the AAP [3] and ASHA [2]. The Ear Scale app can be used to screen school-age children and individuals at a high risk of developing hearing loss and facilitate early detection of abnormal or worsening thresholds. The Ear Scale app is therefore an appropriate tool to screen for disabling hearing loss and detect hearing loss in a nonsoundproof environment. Children who have limited access to audiologists may benefit from a smartphone-based, freely available self-assessment hearing screening test such as this. With increasing rates of age- and noise-related hearing loss globally, further studies are required to examine the suitability of the Ear Scale app for early detection or prevention of hearing loss in the future.

Limitations

The environmental noise level is one of the most common concerns in hearing screening [7,11,27,30,42,43]. This study was conducted at a school, where ambient noise levels were increased but not excessive at various times, which may have influenced the findings. Therefore, recalibration is required to

reset RETSPLs and maximum output levels with bundled earphones (Apple EarPods) for each new device model. At the same time, we must recalibrate the mobile devices with the KEMAR manikin, following the same procedures to obtain the mean values [44].

Conclusion

This paper proposes an innovative approach to hearing screening of school-age children. We developed an Ear Scale app that is comparable with clinical-grade PTS in a sound-treated booth in terms of hearing test results. With favorable high sensitivity and specificity rates and low false-positive and false-negative rates, this study demonstrated that using the proposed Ear Scale app can rapidly screen hearing status and provide stratified test values for each screened ear, and it is therefore an ideal tool for hearing screening in schools.

Acknowledgments

The authors would like to thank Shang-Liang Wu, Hsin-Yi Huang, Wen-Tsung Kuo, the Big Data Center of Taipei Veterans General Hospital, and Biostatistics Task Force of Taipei Veterans General Hospital for their assistance during the study. This study was supported by grants from Taipei Veterans General Hospital (V107E-004-2(108), 108VACS-003, V108C-178) and the Ministry of Health and Welfare (MOHW108-TDU-B-211-124019, MOHW108-TDU-B-211-133001). The funders had no role in the study design, data collection and analysis, the decision to publish, or preparation of the manuscript.

Authors' Contributions

YCC and WHL built the Ear Scale app and participated in writing the final draft. YCC and YFC designed the study, interpreted the results, and wrote the draft. YT and YHL and FL conceptualized and designed the study, interpreted the data, and critically revised the manuscript. All authors helped critically review and revise the manuscript and approved the final version.

Conflicts of Interest

None declared.

References

1. Berg AL, Papri H, Ferdous S, Khan NZ, Durkin MS. Screening methods for childhood hearing impairment in rural Bangladesh. Int J Pediatr Otorhinolaryngol 2006 Jan;70(1):107-114. [doi: 10.1016/j.ijporl.2005.05.029] [Medline: 16023223]
2. Guidelines for Audiologic Screening. USA: American Speech-Language-Hearing Association; 1997.
3. Cunningham M, Cox EO, Committee on Practice and Ambulatory Medicine and the Section on Otolaryngology and Bronchoesophagology. Hearing assessment in infants and children: recommendations beyond neonatal screening. Pediatrics 2003 Feb;111(2):436-440. [Medline: 12563074]
4. Roush J. Identification of hearing loss and middle ear dysfunction in preschool and school age children (American Academy of Audiology, Report and Position Statement). Semin Hear 2008 Aug 26;11(4):357-371. [doi: 10.1055/s-0028-1085515]
5. Liao WH, Lien CF, Young ST. The Hearing Scale Test for hearing screening of school-age children. Int J Pediatr Otorhinolaryngol 2010 Jul;74(7):760-764. [doi: 10.1016/j.ijporl.2010.03.033] [Medline: 20394995]
6. Snijders T, Bosker R. The Worship and Principles of the Church of England: A Sermon, Preached at the Opening of Christ's Church, in Great-Barrington, on Christ-Mass Day, M.DCC.LXIV. Multilevel modeling: Printed and sold by William Goddard, at the Post-Office; 1999.
7. Meinke DK, Dice N. Comparison of audiometric screening criteria for the identification of noise-induced hearing loss in adolescents. Am J Audiol 2007 Dec 1;16(2):S190. [doi: 10.1044/1059-0889(2007/023)]
8. Liao W, Young S, Tang S, Shiao A, Wang S, Lien C. A novel method for quick hearing assessment of children. 2010 Presented at: 2010 International Conference on Electronics and Information Engineering; August 1-3, 2010; Kyoto, Japan. [doi: 10.1109/ICEIE.2010.5559833]
9. Masalski M, Kręcicki T. Self-test web-based pure-tone audiometry: validity evaluation and measurement error analysis. J Med Internet Res 2013 Apr 12;15(4):e71 [FREE Full text] [doi: 10.2196/jmir.2222] [Medline: 23583917]
10. Liao WH, Young ST, Lien CF, Wang SJ. An audiometer to monitor progressive hearing change in school-aged children. J Med Screen 2011;18(1):8-11. [doi: 10.1258/jms.2011.010120] [Medline: 21536810]
11. Masalski M, Grysiński T, Kręcicki T. Hearing tests based on biologically calibrated mobile devices: comparison with pure-tone audiometry. JMIR Mhealth Uhealth 2018 Jan 10;6(1):e10 [FREE Full text] [doi: 10.2196/mhealth.7800] [Medline: 29321124]
12. Honeth L, Bexelius C, Eriksson M, Sandin S, Litton J, Rosenhall U, et al. An internet-based hearing test for simple audiometry in nonclinical settings: preliminary validation and proof of principle. Otol Neurotol 2010 Jul;31(5):708-714. [doi: 10.1097/MAO.0b013e3181de467a] [Medline: 20458255]
13. Sandström J, Swanepoel DW, Carel Myburgh H, Laurent C. Smartphone threshold audiometry in underserved primary health-care contexts. Int J Audiol 2016;55(4):232-238. [doi: 10.3109/14992027.2015.1124294] [Medline: 26795898]

14. Renda L, Selçuk OT, Eyigör H, Osma U, Yılmaz MD. Smartphone based audiometric test for confirming the level of hearing; is it useable in underserved areas. J Int Adv Otol 2016 Apr;12(1):61-66 [FREE Full text] [doi: 10.5152/iao.2016.1421] [Medline: 27340985]
15. Masalski M, Kipiński L, Grysiński T, Kręcicki T. Hearing tests on mobile devices: evaluation of the reference sound level by means of biological calibration. J Med Internet Res 2016 May 30;18(5):e130 [FREE Full text] [doi: 10.2196/jmir.4987] [Medline: 27241793]
16. Bright T, Pallawela D. Validated smartphone-based apps for ear and hearing assessments: a review. JMIR Rehabil Assist Technol 2016 Dec 23;3(2):e13 [FREE Full text] [doi: 10.2196/rehab.6074] [Medline: 28582261]
17. Abu-Ghanem S, Handzel O, Ness L, Ben-Artzi-Blima M, Fait-Ghelbendorf K, Himmelfarb M. Smartphone-based audiometric test for screening hearing loss in the elderly. Eur Arch Otorhinolaryngol 2016 Feb;273(2):333-339. [doi: 10.1007/s00405-015-3533-9] [Medline: 25655259]
18. Swanepoel DW, Myburgh HC, Howe DM, Mahomed F, Eikelboom RH. Smartphone hearing screening with integrated quality control and data management. Int J Audiol 2014 Dec;53(12):841-849. [doi: 10.3109/14992027.2014.920965] [Medline: 24998412]
19. Foulad A, Bui P, Djalilian H. Automated audiometry using apple iOS-based application technology. Otolaryngol Head Neck Surg 2013 Nov;149(5):700-706. [doi: 10.1177/0194599813501461] [Medline: 23963611]
20. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. J Med Internet Res 2013 Jun 14;15(6):e120 [FREE Full text] [doi: 10.2196/jmir.2600] [Medline: 23770578]
21. Xing Y, Fu Z, Wu X, Chen J. Evaluation of Apple iOS-based automated audiometry. 2016 Sep 1 Presented at: 22nd International Congress on Acoustics; September 5-9, 2016; Buenos Aires URL: https://www.icacommission.org/Proceedings/ICA2016BuenosAires/papers/ICA2016-0114.pdf
22. Masalski M, Grysiński T, Kręcicki T. Biological calibration for web-based hearing tests: evaluation of the methods. J Med Internet Res 2014 Jan 15;16(1):e11 [FREE Full text] [doi: 10.2196/jmir.2798] [Medline: 24429353]
23. Ho CY, Li PC, Young ST. Reference equivalent threshold sound pressure levels for Apple EarPods. J Acoust Soc Am 2017 Dec;141(2):EL115. [doi: 10.1121/1.4976110] [Medline: 28253642]
24. International Organization for Standardization. Reference equivalent threshold sound pressure levels for Apple EarPods URL: https://www.iso.org/standard/43601.html [accessed 2019-03-01] [WebCite Cache ID 76YADA31k]
25. Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am 1971 Feb;49(2B):467-477. [doi: 10.1121/1.1912375]
26. Flanary VA, Flanary CJ, Colombo J, Kloss D. Mass hearing screening in kindergarten students. Int J Pediatr Otorhinolaryngol 1999 Oct 25;50(2):93-98. [doi: 10.1016/S0165-5876(99)00210-4] [Medline: 10576608]
27. Mason CA, Gaffney M, Green DR, Grosse SD. Measures of follow-up in early hearing detection and intervention programs: a need for standardization. Am J Audiol 2008 Jun;17(1):60-67. [doi: 10.1044/1059-0889(2008/007)] [Medline: 18519580]
28. Khoza-Shangase K, Kassner L. Automated screening audiometry in the digital age: exploring uHear™ and its use in a resource-stricken developing country. Int J Technol Assess Health Care 2013 Jan;29(1):42-47. [doi: 10.1017/S0266462312000761] [Medline: 23298579]
29. Cerwall P, Lundvall A, Jonsson P, Carson S, Möller R, Jonssonricsson MR. GSMA. 2018. Ericsson Mobility Report: On the pulse of the Networked Society URL: https://www.gsma.com/latinamerica/ericsson-mobility-report-networked-society/ [accessed 2019-03-01] [WebCite Cache ID 76YAdHGt7]
30. Bamford J, Fortnum H, Bristow K, Smith J, Vamvakas G, Davies L, et al. Current practice, accuracy, effectiveness and cost-effectiveness of the school entry hearing screen. Health Technol Assess 2007 Aug;11(32):1-168, iii [FREE Full text] [doi: 10.3310/hta11320] [Medline: 17683682]
31. Jimoh F, Lund EK, Harvey LJ, Frost C, Lay WJ, Roe MA, et al. Comparing diet and exercise monitoring using smartphone app and paper diary: a two-phase intervention study. JMIR Mhealth Uhealth 2018 Jan 15;6(1):e17 [FREE Full text] [doi: 10.2196/mhealth.7702] [Medline: 29335239]
32. Swendeman D, Comulada WS, Koussa M, Worthman CM, Estrin D, Rotheram-Borus MJ, et al. Longitudinal validity and reliability of brief smartphone self-monitoring of diet, stress, and physical activity in a diverse sample of mothers. JMIR Mhealth Uhealth 2018 Sep 21;6(9):e176 [FREE Full text] [doi: 10.2196/mhealth.9378] [Medline: 30249576]
33. El Shafie RA, Weber D, Bougatf N, Sprave T, Oetzel D, Huber PE, et al. Supportive care in radiotherapy based on a mobile app: prospective multicenter survey. JMIR Mhealth Uhealth 2018 Aug 30;6(8):e10916 [FREE Full text] [doi: 10.2196/10916] [Medline: 30166275]
34. Webb MJ, Wadley G, Sanci LA. Experiences of general practitioners and practice support staff using a health and lifestyle screening app in primary health care: implementation case study. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e105 [FREE Full text] [doi: 10.2196/mhealth.8778] [Medline: 29691209]
35. Robbins RN, Gouse H, Brown HG, Ehlers A, Scott TM, Leu C, et al. A mobile app to screen for neurocognitive impairment: preliminary validation of neuroscreen among HIV-infected South African adults. JMIR Mhealth Uhealth 2018 Jan 5;6(1):e5 [FREE Full text] [doi: 10.2196/mhealth.9148] [Medline: 29305338]

36. March S, Day J, Zieschank K, Ireland M. The Interactive Child Distress Screener: development and preliminary feasibility testing. JMIR Mhealth Uhealth 2018 Apr 19;6(4):e90 [FREE Full text] [doi: 10.2196/mhealth.9456] [Medline: 29674310]
37. Paglialonga A, Tognola G, Pinciroli F. Apps for hearing science and care. Am J Audiol 2015 Sep;24(3):293-298. [doi: 10.1044/2015_AJA-14-0093] [Medline: 26649533]
38. Yeh CH, Wei ST, Chen TW, Wang CY, Tsai MH, Lin CD. A web-based audiometry database system. J Formos Med Assoc 2014 Jul;113(7):477-480 [FREE Full text] [doi: 10.1016/j.jfma.2013.10.006] [Medline: 24262921]
39. Yimtae K, Israsena P, Thanawirattananit P, Seesutas S, Saibua S, Kasemsiri P, et al. A tablet-based mobile hearing screening system for preschoolers: design and validation study. JMIR Mhealth Uhealth 2018 Oct 23;6(10):e186 [FREE Full text] [doi: 10.2196/mhealth.9560] [Medline: 30355558]
40. Mahomed F, Swanepoel DW, Eikelboom RH, Soer M. Validity of automated threshold audiometry: a systematic review and meta-analysis. Ear Hear 2013;34(6):745-752. [doi: 10.1097/01.aud.0000436255.53747.a4] [Medline: 24165302]
41. Margolis RH, Morgan DE. Automated pure-tone audiometry: an analysis of capacity, need, and benefit. Am J Audiol 2008 Dec;17(2):109-113. [doi: 10.1044/1059-0889(2008/07-0047)] [Medline: 18840703]
42. FitzZaland RE, Zink GD. A comparative study of hearing screening procedures. Ear Hear 1984;5(4):205-210. [Medline: 6468778]
43. Sabo MP, Winston R, Macias JD. Comparison of pure tone and transient otoacoustic emissions screening in a grade school population. Am J Otol 2000 Jan;21(1):88-91. [doi: 10.1016/S0196-0709(00)80080-0] [Medline: 10651440]
44. Gardner WG, Martin KD. HRTF measurements of a KEMAR. J Acoust Soc Am 1995 Jun;97(6):3907-3908. [doi: 10.1121/1.412407]

Abbreviations

AAP: American Academy of Pediatrics
ASHA: American Speech-Language-Hearing Association
HL: hearing level
HST: Hearing Scale Test
ISO: International Organization for Standardization
NR: no response
PTA: pure-tone average
PTS: pure-tone screening
RETSPL: reference equivalent threshold sound pressure level
S: stratified hearing scale
SPL: sound pressure level

Edited by G Eysenbach; submitted 25.08.18; peer-reviewed by R Dewey, C He; comments to author 27.10.18; revised version received 20.12.18; accepted 21.01.19; published 01.04.19.

Please cite as:
Chu YC, Cheng YF, Lai YH, Tsao Y, Tu TY, Young ST, Chen TS, Chung YF, Lai F, Liao WH
A Mobile Phone–Based Approach for Hearing Screening of School-Age Children: Cross-Sectional Validation Study
JMIR Mhealth Uhealth 2019;7(4):e12033
URL: https://mhealth.jmir.org/2019/4/e12033/
doi:10.2196/12033
PMID:30932870

Original Paper

A Mobile Web App to Improve Health Screening Uptake in Men (ScreenMen): Utility and Usability Evaluation Study

Chin Hai Teo[1*], BMedSc, PhD; Chirk Jenn Ng[1*], MBBS, MMed (FamMed), PhD; Sin Kuang Lo[1*], BCompSc; Chip Dong Lim[1*], BDes; Alan White[2*], RN, PhD

[1]Department of Primary Care Medicine, University of Malaya eHealth Initiative, Faculty of Medicine, Kuala Lumpur, Malaysia

[2]Institute for Health & Wellbeing, Leeds Beckett University, Leeds, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Chirk Jenn Ng, MBBS, MMed (FamMed), PhD
Department of Primary Care Medicine
University of Malaya eHealth Initiative
Faculty of Medicine
University of Malaya
Kuala Lumpur, 50603
Malaysia
Phone: 60 37949 ext 2306
Email: ngcj@um.edu.my

Abstract

Background: Globally, the uptake of health screening is suboptimal, especially in men and those of younger age. In view of the increasing internet access and mobile phone ownership, ScreenMen, a mobile Web app, was developed to improve health screening uptake in men.

Objective: This study aimed to evaluate the utility and usability of ScreenMen.

Methods: This study used both qualitative and quantitative methods. Healthy men working in a banking institution were recruited to participate in this study. They were purposively sampled according to job position, age, education level, and screening status. Men were asked to use ScreenMen independently while the screen activities were being recorded. Once completed, retrospective think aloud with playback was conducted with men to obtain their feedback. They were asked to answer the System Usability Scale (SUS). Intention to undergo screening pre- and postintervention was also measured. Qualitative data were analyzed using a framework approach followed by thematic analysis. For quantitative data, the mean SUS score was calculated and change in intention to screening was analyzed using McNemar test.

Results: In total, 24 men participated in this study. On the basis of the qualitative data, men found ScreenMen useful as they could learn more about their health risks and screening. They found ScreenMen convenient to use, which might trigger men to undergo screening. In terms of usability, men thought that ScreenMen was user-friendly and easy to understand. The key revision done on utility was the addition of a reminder function, whereas for usability, the revisions done were in terms of attracting and gaining users' trust, improving learnability, and making ScreenMen usable to all types of users. To attract men to use it, ScreenMen was introduced to users in terms of *improving health* instead of *going for screening*. Another important revision made was emphasizing the screening tests the users do not need, instead of just informing them about the screening tests they need. A *Quick Assessment Mode* was also added for users with limited attention span. The quantitative data showed that 8 out of 23 men (35%) planned to attend screening earlier than intended after using the ScreenMen. Furthermore, 4 out of 12 (33%) men who were in the precontemplation stage changed to either contemplation or preparation stage after using ScreenMen with P=.13. In terms of usability, the mean SUS score of 76.4 (SD 7.72) indicated that ScreenMen had good usability.

Conclusions: This study showed that ScreenMen was acceptable to men in terms of its utility and usability. The preliminary data suggested that ScreenMen might increase men's intention to undergo screening. This paper also presented key lessons learned from the beta testing, which is useful for public health experts and researchers when developing a user-centered mobile Web app.

(JMIR Mhealth Uhealth 2019;7(4):e10216) doi:10.2196/10216

KEYWORDS

internet; mHealth; eHealth; mass screening; health behavior; men's health

Introduction

Background

In the past decade, many Web-based interventions have been developed to improve health outcomes of the public. Web-based interventions not only have a wider reach but also are less labor intensive and less resource exhaustive as compared with conventional health interventions; in addition, they can be interactive, personalized, and fun, which makes learning more effective and ultimately leads to improved users' health behavior. In addition, the impact of Web-based health interventions is further amplified with the flux of mobile technology into health care. The mobile phone, which is a good platform to deliver health care to people anytime and anywhere, is widely available and affordable these days, including in developing countries [1]. Many studies have shown that Web-based interventions, including mobile Web apps, are effective in improving health outcomes such as improving physical activity level, asthma treatment knowledge, psoriasis knowledge, and weight loss, as well as reducing depression symptoms and preventing low back pain [2-10].

Web-based interventions could be deployed to improve health screening uptake as well. Health screening plays an important role to detect and treat diseases at an early stage, which leads to reduced mortality rate [11]. Despite its benefits, the uptake of health screening remains suboptimal, especially in men and those of younger age [12-14]. This pattern has also been observed in Malaysia as reported in the National Health and Morbidity Survey (NHMS), where only 34.9% of Malaysian men attended health screening in 2011 [15]. The NHMS also reported that the prevalence of undiagnosed hypertension was higher than known hypertension in men younger than 55 years [16]. In Malaysia, health screening can be conducted at a public hospital, public health clinic, private hospital, private clinic, as well as blood test lab. There is a public health care facility within every 5-km radius, including in the rural areas. Although the fee for utilizing a public outpatient clinic, including health screening is as low as Malaysian Ringgit (MYR) 1 (approximately US$ 0.25), health screening uptake remains low. There are many contributing factors for this trend; however, lack of awareness and knowledge on health screening could be one of the key factors [17,18].

Many types of interventions to improve the uptake of health screening in men, including those using partners' involvement, educational workshops, reminder phone calls, and letters have been evaluated [19]. However, only educational interventions were found to be effective in improving men's intention to undergo screening and increasing the actual screening uptake; others were inconclusive because of poor study design in terms of blinding of participant and allocation concealment [19]. There are also many Web-based interventions on health screening in men that have been evaluated, such as Web-based patient decision aid on prostate cancer screening, as well as educational Web and social media to encourage HIV screening in men [20-22]. However, these interventions are disease-specific, and there is a lack of intervention that promotes comprehensive (all in one) health screening in men, which is crucial in ensuring holistic care for men [19]. Among the recommended health screening for men by the United States Preventive Services Task Force (USPSTF) are hypertension, diabetes, dyslipidemia, colorectal cancer, lung cancer, HIV, hepatitis, sexually transmitted infections, depression, as well as lifestyle risk factors including smoking status, alcohol usage, obesity, diet, and physical activity [23]. These recommendations should be applied on the basis of men's health profile such as age, ethnicity, and family history.

In view of the increasing internet access and mobile phone ownership in Malaysia, as well as in the world [1], ScreenMen, a mobile Web app, which aims to promote comprehensive health screening in men was developed. ScreenMen is mobile-responsive and aimed to be disseminated via mobile phone to all Malaysian men. It was developed on the basis of theories, evidence, as well as the needs of users [17-19,24]. ScreenMen was developed focusing on men in view of the lower level of health outcomes and behavior in men, as well as answering the call to use a gender sensitive approach in health intervention [25,26]. Before the development of ScreenMen, a need assessment was conducted with working men from a banking institution in Kuala Lumpur to identify their needs on health screening and to find out what men want in a health screening mobile app [18,24]. At the time of development, the prototype of ScreenMen was tested with experts from various backgrounds (alpha testing) and was revised iteratively to improve it.

Before ScreenMen was finalized, a beta testing was conducted. Beta testing aims to test a software with end users in a real-world setting to identify and rectify any potential issue before being released. This is particularly important for a mobile Web app as Web-based technologies are growing and changing rapidly [27]. Poor usability is often reported as one of the main reasons why users stopped using a mobile Web app, as a consequence of inadequate user testing [28]. To ensure that a mobile Web app is useful, experts recommend that it should be evaluated in terms of its utility (whether a website provides the features the users need), as well as usability (how easy and pleasant the features are) with users [29,30].

Objectives

Thus, this study aimed to evaluate ScreenMen with men from the community in terms of its utility and usability, as well as to present the key revisions made to improve the utility and usability in ScreenMen, on the basis of the feedback obtained.

Methods

Study Design Overview

This study used a mixed-methods design to evaluate the utility and usability of ScreenMen with end users. The mixed-methods approach helps to triangulate the findings using qualitative and

quantitative methods. The qualitative assessment was done using the retrospective think-aloud method with the aid of a topic guide, which contained questions on utility and usability [31,32]. Instead of the prospective think-aloud method, the retrospective think-aloud method was chosen to simulate the actual usage and to resolve any unforeseen navigability issues at users' real-life settings [30]. A questionnaire was also used to score and evaluate ScreenMen in terms of utility and usability quantitatively. This study was approved by the University of Malaya Medical Centre Medical Research Ethics Committee (MRECID No. 201610144372).

Study Setting, Sampling, and Recruitment

This study was conducted with healthy men from a banking institution in Kuala Lumpur, the capital of Malaysia. Unlike alpha testing with experts, which was done at the developer's site, beta testing is conducted at the users' settings, which was men's working place in this study. Men from a banking institution were chosen as they were reported to have increased work-related stress, which affected their psychological as well as physical health, including smoking and alcohol overuse, depression, body posture issue, and visual problem [33,34]. Their busy schedule and excessive work demand also contributed to nonattendance of health screening [35,36]. These men represent a group of *hard-to-reach* men in the community, who often do not seek health care services despite having easy access to them.

The same banking institution where the needs assessment was conducted in the earlier phase was selected as the recruitment site for this beta testing. This study was approved by the banking institution. Men who have a mobile phone and were from the banking institution were recruited to participate in the beta testing. They were purposively sampled according to their job position, age, education level, and screening status to achieve maximal variation of the feedback on ScreenMen. A Microsoft Excel spreadsheet that contained the participants' demography was used to plan the recruitment to ensure equal representation from each sampling criterion. Men who participated in the needs assessment phase were first contacted and arranged for interviews. Then, the snowballing method was used to recruit new participants, where the recruited participants were asked to recommend their colleagues to participate in the study. New participants were also included in addition to those who had participated in the needs assessment phase to gather more feedback on ScreenMen. All participants were reimbursed with MYR 50 (approximately US $ 12) for their time participating in this study, which took about an hour.

The sample size of a usability study is often small. Studies have shown that the optimum sample size to detect sufficient usability problem is 10 users [37]. As this study involved quantitative evaluation, at least 20 participants were aimed to be recruited to obtain statistically significant number [38]. The recruitment was stopped once data saturation was reached.

The ScreenMen Web App (Beta Testing Version)

ScreenMen is a mobile-responsive Web app aimed to be disseminated via smartphone. It aims to educate men, empower men, and improve men's behavior on health screening. ScreenMen was developed to contain male-sensitive attributes (such as using car maintenance analogy), as well as evidence-based recommendation for health screening gathered from the USPSTF [23]. Apart from that, 4 key sections of ScreenMen were developed following a framework modified from the health literacy principle, to guide the learning process in ScreenMen [39]. The 4 sections are as follows:

- Learn: This section contains a short educational video to demystify the misconceptions on health screening, which were identified in the needs assessment.
- Assess: This is an interactive section where users can interact with ScreenMen to assess their health risks and obtain personalized health advice, as well as the evidence-based health screening they need, on the basis of their health profile. There were 15 health conditions that were being assessed in ScreenMen, and they include obesity, unhealthy diet, physical activity, tobacco use, alcohol misuse, high blood pressure, diabetes, dyslipidemia, colorectal cancer, lung cancer, HIV, syphilis, hepatitis B, hepatitis C, and depression. This section is algorithm-driven and attempts to mimic a real-life clinical consultation with a doctor. A health report can be generated at the end of this section.
- Ask: In this section, there is a list of frequently asked questions about screening, which men can read if they would like to have further clarifications about screening. The questions were developed on the basis of the comprehensive framework on barriers and facilitators to health screening in men [17].
- Prepare: This section aims to prepare the users to undergo health screening by providing basic logistic information such as where to screen, when to screen, and cost of screening.

Data Collection

In-depth interviews (IDIs) and focus group discussions (FGDs) were conducted for data collection. At the time of the appointment, the researchers first briefed the participants about the study using a participant information sheet. The participants were encouraged to ask questions and were informed that they could stop the study at any time. Once agreed to participate, the participants were asked to sign a consent form and fill up the demography form, including intention to undergo screening. Then, the participants were given a smartphone with ScreenMen activated on the screen. They were asked to use it themselves and notify the researchers once they had finished using it. All on-screen activities were being recorded using a free screen recording app (AZ Screen Recorder by Hecorat). The researchers were present in the same room to observe the participants' behavior when using ScreenMen and take field notes, as well as assist the users, only when necessary.

Table 1. Postintervention questionnaire.

No.	Item	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
1	I think that I would like to use this website frequently	__[a]	—	—	—	—
2	I found the website unnecessarily complex	—	—	—	—	—
3	I thought the website was easy to use	—	—	—	—	—
4	I think that I would need the support of a technical person to be able to use this website	—	—	—	—	—
5	I found the various functions in this website well integrated	—	—	—	—	—
6	I thought there was too much inconsistency in this website	—	—	—	—	—
7	I would imagine that most people would learn to use this website very quickly	—	—	—	—	—
8	I found the website very cumbersome to use	—	—	—	—	—
9	I felt very confident using the website	—	—	—	—	—
10	I needed to learn a lot of things before I could get going with this website	—	—	—	—	—
11	Does the website help you to understand more about your health risks?	—	Yes			
		—	No			
12	Does the website help you to understand more about health screening?	—	Yes			
		—	No			
13	Do you intend to go for health screening in the future?	—	Yes, in the next 1 month			
		—	Yes, in the next 6 months			
		—	Yes, in the next 1 year			
		—	Yes, in the next 2 years			
		—	Yes, in the next 5 years			
		—	No, I do not intend to go for health screening			
14	Would you recommend this website to your family or friends?	—	Yes			
		—	No			

[a]Indicate boxes for participant's response.

Once completed, the participants were asked to answer the postintervention questionnaire, which contains the validated 10-question System Usability Scale (SUS) [40] and 4 utility questions, including intention to undergo screening (Table 1). The scale of question 13, *Do you intend to go for health screening in the future?* was developed on the basis of the transtheoretical model of health behavior change [41]. This model explained the stages of behavior change in a person, from precontemplation (no intention of change), contemplation (intent to change in the next 6 months), preparation (intent to change in the next 1 month), action (acted and maintained behavior within 6 months), to maintenance (maintained behavior more than 6 months). Only precontemplation, contemplation, and preparation were adopted for the scale of question 13 as these are related to intention to change.

Then, retrospective think aloud with playback was conducted. Using a topic guide, the researchers started the interview by asking the participants to provide their overall opinion on the Web app; to comment on its contents and layout (usability); to explain how, if the Web app helped them to understand more about health and screening; and to suggest any other part of the Web app that can be improved. The on-screen recording was played to assist the participant in the retrospective think-aloud process. They were probed to comment on the content and layout when going through each section of ScreenMen. All conversations during the retrospective think aloud were audio-recorded. For FGDs, all procedures were similar to those of IDIs, except that only 1 participant was given the project's mobile phone, whereas others used their own mobile phones and during the feedback session, and the ScreenMen Web app was projected on the screen and navigated by the researcher, page by page, to assist the retrospective think aloud, instead of playing the on-screen recording of each participant.

Data Analysis

The analysis was performed after completing all data collection. Subsequently, the researchers met to discuss the issues and proposed the revisions to be done on ScreenMen. The qualitative data obtained were analyzed using a framework approach to systematically guide the revision of the Web app. After each interview or discussion session, the researchers discussed and compiled a list of comments and issues on ScreenMen, with the aid of the field notes taken. The researchers then listened to the audio recording to triangulate and check for additional

comments. Unlike the usual approach in a qualitative study, the audio-recordings were not transcribed verbatim in this study as the purpose of this beta testing was to capture the users' feedback on the Web app rather than to provide an in-depth understanding of the users' experience. The list of comments compiled was then coded under utility or usability and by section (Multimedia Appendix 1), which were then used to revise ScreenMen. In addition, to present the data in a more meaningful way in this paper, the comments and issues identified were grouped and categorized according to common themes. This was done by the first author and discussed and agreed by all authors.

For the quantitative data, all data were managed and analyzed using the IBM SPSS Statistics version 21. First, the SUS score from each participant and a mean SUS score for all participants were calculated. The SUS score was interpreted using the adjective rating scale developed by Bangor [42]. The utility questions were analyzed using descriptive statistics (percentage of yes). For intention to undergo screening, the percentages of participants who plan to screen earlier than intended, later than intended, and no change in intention after using ScreenMen, were calculated by comparing the intention to screen pre- and postintervention. Intention to screen was also analyzed according to stage of behavior change, specifically by comparing the number of participants in the precontemplation stage (more than 6 months) with the number of participants in either contemplation or preparation stage (6 months or less) after using ScreenMen, using the McNemar test.

Results

Participant Demography

In total, 24 men participated in the beta testing with 14 IDIs and 2 FGDs (5 participants in each FGD). They were conducted from February to March 2017. The details of the participants are shown in Table 2.

Qualitative Evaluation (Observation and Retrospective Think Aloud)

Utility

The participants found the ScreenMen useful as they could learn more about their health risks, what screening to go for, and what they could do to improve their health. The same was found for older men who had undergone screening, as ScreenMen contains information they never knew before, such as colorectal cancer and the unnecessary screening tests:

> *I like all of it, it tells you your health, everything about where you are (in terms of health), and what you should do to improve it.* [40-59 years, Senior manager]
>
> *Using this web, people know what diseases they should check* [40-59 years, Clerk]
>
> *Now I understand about the importance of health screening. We don't know that our lifestyle could actually affect our health. Using this website, you know what to be improved upon.* [20-29 years, Sales advisor]

Some mentioned that they were glad to learn about the unnecessary screening tests. One participant suggested to highlight this more to ensure all users get it:

> *My key take home message from this website is some of the tests are unnecessary, for example, the liver or kidney test, as it may over or under detect the disease. Nowadays, there are a lot of external blood test centers, they normally package ECG, heart stress test, and everything together and sell you thousands of Ringgit (Malaysian currency). I didn't know that those are actually unnecessary. So, this is something I got to know now. It's good to know that those are actually not useful for screening. This information is a little secluded and need to be highlighted better.* [20-29 years, Officer]

A participant was glad as:

> *...it contains localized contents for us (Malaysian), unlike the UK or US websites.* [20-29 years, Officer]

One participant mentioned that using ScreenMen may trigger men to take care of their health:

> *ScreenMen is easy to use, can add more knowledge and act as a trigger to take care of health when going through the website, unlike those who do not receive anything and do not do anything about health.* [30-39 years, Clerk]

Men also felt that ScreenMen was convenient to use.

> *We have limited time for screening. With this, we can check at anywhere, we can have the information and what can we do (to improve health). It is just like talking to a doctor or consultant.* [30-39 years, Clerk]
>
> *It is good for people. People are always with handphone. With application like this, one doesn't need to go anywhere, at home also can do, at office also can do.* [30-39 years, Clerk]
>
> *will share this website with friends via Facebook Group.* [40-59 years, Clerk]

On the other hand, 1 participant raised the issue that users may not use ScreenMen again after using it once:

> *It's good. Will I use the website? Yes. But subsequently will I continue to use it repeatedly again, it remains a question mark.* [30-39 years, Senior manager]

The participants also suggested that a reminder function may be useful as users may not act instantly after using ScreenMen. Thus, the research team added a function where users can input their email, and an event entitled *My Check-up Day* would appear in their email calendar, calculated on the basis of their past screening date. It serves as a reminder for users as they check their calendar daily and come across that added event.

Some participants suggested that it would be good to have a list of screening centers with phone numbers to make appointments on the website, as that may facilitate users to take action to screen. However, the research teams decided not to include the list to avoid being perceived as using it for commercial reason.

Table 2. Characteristics of participants in the beta testing (N=24).

Characteristics	Statistics
Age range (years), n (%)	
20-29	7 (29)
30-39	10 (42)
40-59	7 (29)
Age (years), mean (range)	37 (23-56)
Ethnicity, n (%)	
Malay	10 (42)
Chinese	10 (42)
Indian	3 (13)
Others	1 (4)
Position, n (%)	
Senior manager	7 (29)
Officer	5 (21)
Sales advisor	5 (21)
Clerk	7 (29)
Education, n (%)	
Secondary school	5 (21)
Certificate/diploma	4 (17)
Degree	13 (54)
Postgraduate	2 (8)
Marital status, n (%)	
Unmarried	10 (42)
Married	14 (58)
Screened in the past 1 year, n (%)	
Yes	9 (38)
Mobile phone operating system, n (%)	
iOS	11 (46)
Android	12 (50)
Windows	1 (4)
Participated in needs assessment, n (%)	
Yes	13 (54)

Usability

Overall, participants mentioned:

> *[ScreenMen] is quite user-friendly, comfortable to look at, and not too cluttered. The interface is easy to understand, not too complicated.* [20-29 years, Officer]

The participants also felt assured to use the Web app as it was stated upfront that the Web app does not capture any identifiable information from them.

There were several key issues with revisions to improve ScreenMen, and they were grouped under 3 themes: (1) attracting and engaging users, (2) ensuring effective learning, and (3) catering for the widest range of users' characteristics.

Theme 1: Attracting and Engaging Users

Designing a Simple and Focused Home Page

The home page of ScreenMen outlined the 4 key sections of the Web app. Some participants found that the home page contained too much information to read and felt that it may put off users. Thus, the home page was simplified to include only the main objective of what users may gain from this Web app (Figure 1).

Promoting the Concept of Health Instead of Just Screening

Men were less interested in screening as they did not understand about screening and its importance. Describing the Web app as a platform to learn about screening did not interest the users. However, men wanted to be healthy, and they more readily received information that can keep them healthy. Thus, the objective on the home page was framed in terms of learning about users' health risks and ways to stay healthy, instead of learning about health screening. Additional health information, such as erectile dysfunction and urinary symptoms, were also added as requested by men, to provide more information than just screening.

Highlighting the Credibility of the Web App

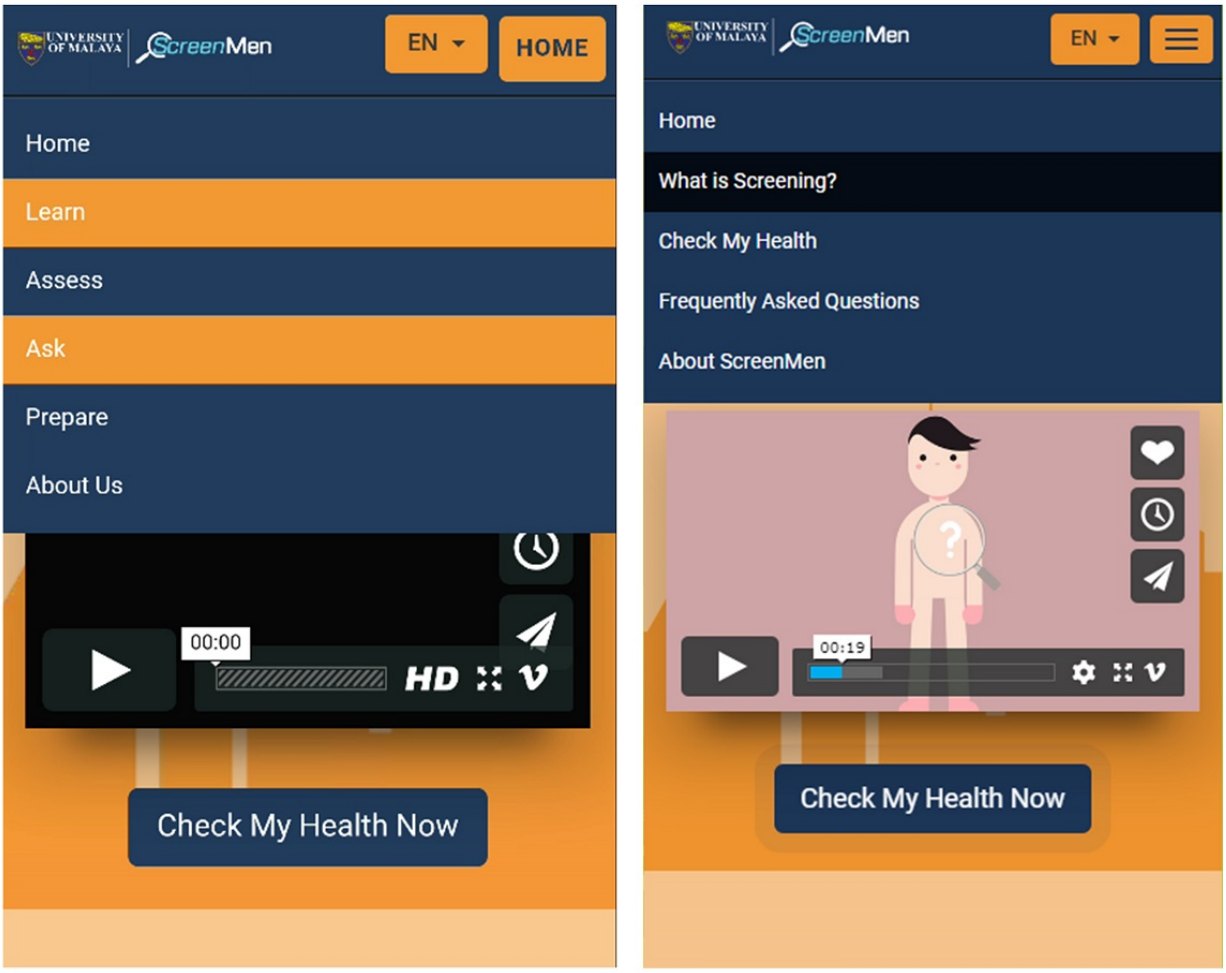
The participants felt that there was a lack of credibility on the home page. They mentioned that the credibility of Web app is crucial to gain users' trust so that they continued to use the Web app. To address this issue, they suggested to enlarge the university's logo on the home page.

Incorporating a Male-Favored Avatar

The ScreenMen Web app attempted to attract men using Dr ScreenMen, a Superman-resembling doctor avatar at the home page. Although some liked Dr ScreenMen figure as it encourages them to be strong, especially those from lower educational level, others had no comment on the Dr ScreenMen figure. One participant suggested to make Dr ScreenMen provide various types of reaction, but this was not done because of technical complexity and the potential impact on Web loading time.

Figure 1. The home page of ScreenMen before and after revision.

Figure 2. The menus of ScreenMen before and after revision.



Theme 2: Ensuring Effective Learning

Using Practical Terms Instead of Theoretical Concepts

From the researchers' observation, the Learn, Assess, Ask, and Prepare menus were unclear, and participants were confused about these concepts. For example, some thought that they could ask questions to a doctor at the *Ask* section and were lost looking for that function. Theoretical concepts were difficult to be understood by users, and thus, the menus *Learn*, *Assess*, and *Ask* were revised to *What is Screening?*, *Check My Health*, and *Frequently Asked Questions*, to more accurately represent the content of each Web app section (Figure 2). The *Prepare* section was removed and merged into *Check My Health*.

Using Linear Learning Design for a More Structured Learning

Users are allowed to navigate freely to any section of ScreenMen by using the icons on the home page. This was done to cater to users who had already understood the basics of health screening and the repeat users. However, some users were confused as they went to the third section directly from the home page and did not go through the first and second section. Thus, the navigation links on the home page were removed. Users who liked to skip any section could use the *hamburger* button.

Incorporating Concepts That Are Familiar to Men

Most participants agreed that the car maintenance analogy was very useful in helping them learn about health screening. The only comment on this was to use the word *car service* instead of *car maintenance*, as the term is more commonly used among men. However, this change was not made as *maintenance* is closer to health screening concept, where maintenance is about the routine schedule for service, whereas service is about the task performed on a vehicle.

Showing Important Information First Instead of Optional Information

At the time of the usability testing, it was found that some users lost their attention at the third section (Ask). The *Ask* section contains a long list of frequently asked questions, and most users only skimmed through them. The fourth section contains a small amount of information to prepare users for screening, which is crucial for them to learn. To ensure that users learn this crucial information before losing attention, they were brought forward and merged into the last part of the second *Check My Health* section. The *Frequently Asked Questions* section was made optional as most information in this section was presented in the earlier sections.

Emphasizing the "Negatives" When Addressing Misconceptions

ScreenMen was developed on the basis of the USPSTF guidelines. It advocates evidence-based screening and encourages users to avoid unnecessary screening. To fulfill the *personalized content* factor as suggested by the users during needs assessment, ScreenMen only states the screening tests users need to undergo on the basis of their health profile. However, after using the Web app, it was found from the retrospective think aloud that the users still had the mindset of *undergoing more screening tests or full body screening is better*. It is insufficient to inform men only on the screening tests they need, but it is also necessary to emphasize the screening tests they do not need, especially when addressing misconceptions. To more effectively educate men to avoid unnecessary health screening, ScreenMen was revised to emphasize the tests that they do not need to go for (Figure 3). Some of the unnecessary screening tests, which were commonly done in the community, were highlighted with reasons why they should be avoided (Figure 4). ScreenMen also empowers men to avoid unnecessary screening by encouraging them to ask the doctors 3 questions when choosing screening tests.

Figure 3. User's list of screening recommendation without and with emphasis of not recommended screening.

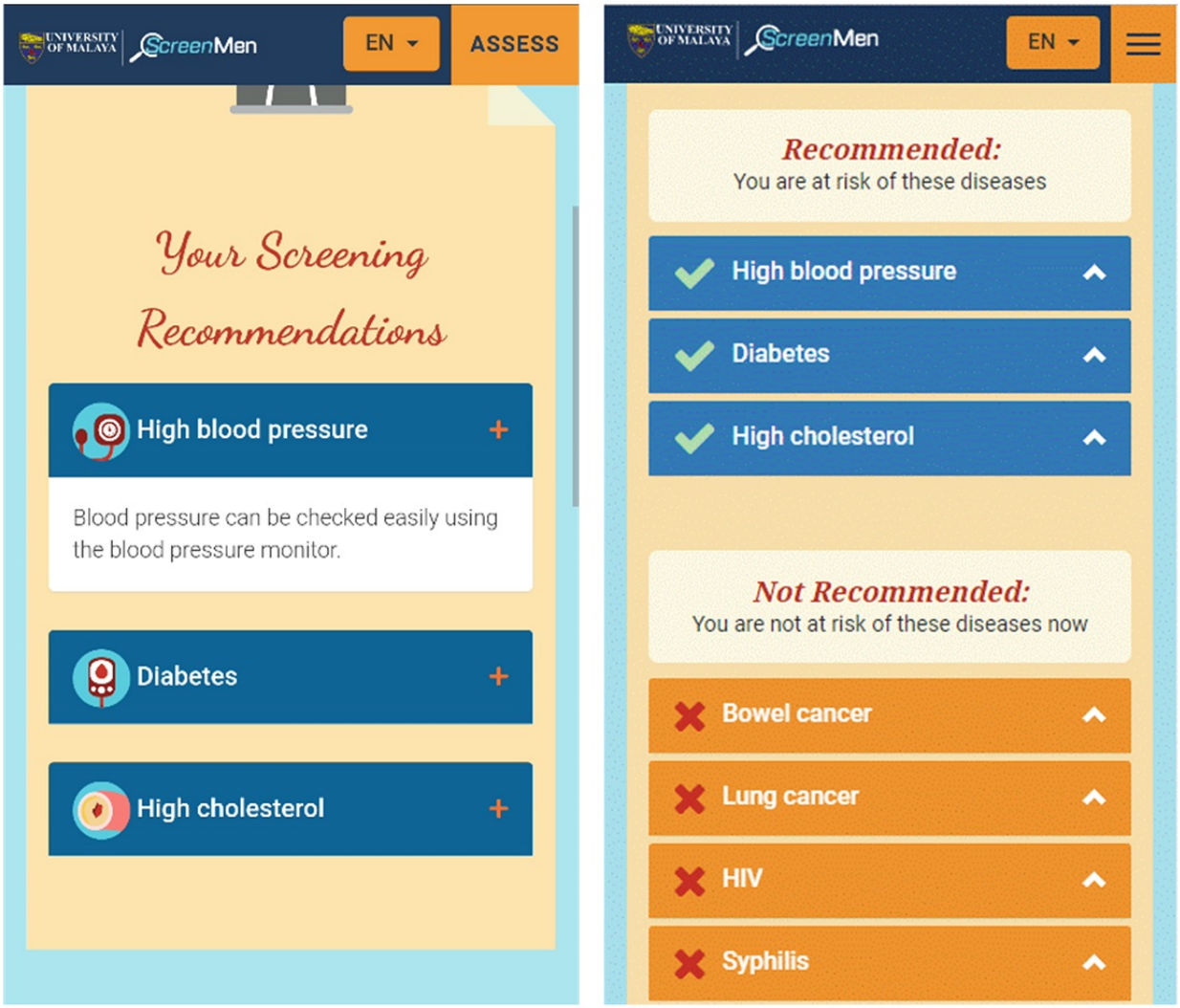
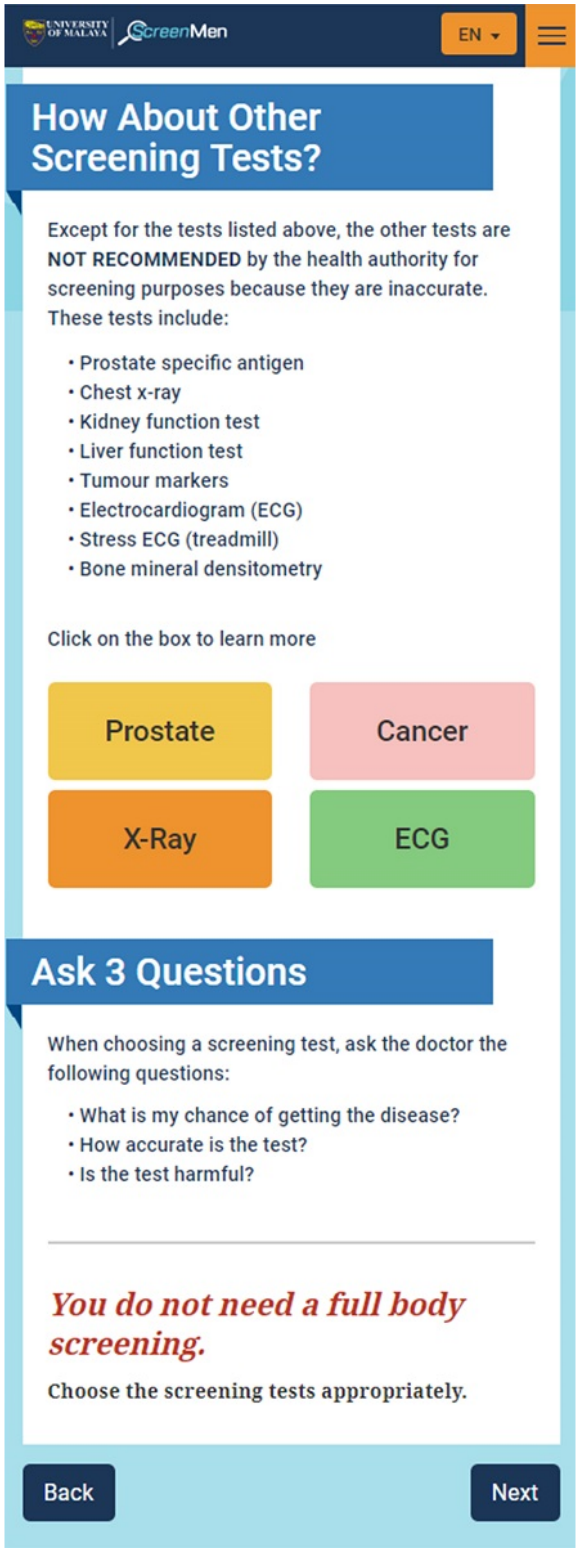


Figure 4. The newly added section to encourage users to avoid unnecessary screening.



Theme 3: Catering for the Widest Range of Users' Characteristics

Suiting Lower Literacy Users

Some participants commented that there was too much information to be read in ScreenMen, especially for people with a lower literacy level. The information in ScreenMen was thus simplified to present only relevant and brief information. Links for additional information were incorporated throughout the Web app for users who may want more information about certain topics.

Anticipating the Lowest Level of Users' Health Behavior

Developed from the medical perspective, users had problems answering some of the questions in ScreenMen. For example, some participants had difficulties answering *I take fruit __ time(s) a day*, as they do not eat fruit every day. This question was developed partly for the recommendation of *5 servings of*

fruits and vegetables per day. As a result of the problem, this question was revised to *weekly* instead of *daily*, as weekly fruit intake was more prevalent among users. The algorithm was also revised to recalculate the users' input to compare against the recommended level.

Another example was the blood pressure reading. Most users could not remember their blood pressure reading but remembered that their blood pressure was normal from the previous health screening. To provide a more accurate advice, an option of *I don't know but I know my BP is normal* was added instead of letting these users select *I don't know*.

Providing a Quick Mode Option for Busy Men

The health assessment section was developed to mimic a real-life clinical consultation with a doctor. This section starts with Dr ScreenMen greeting the users, obtaining users' age, and followed by health assessment, topic by topic. Dr ScreenMen asks question and provides advice on each health condition on the basis of users' answers. At the end of this section, users can view the summary of their health status with screening recommendation.

However, some participants commented that men who are busy may not like to go through this process and would prefer a shorter mode. Though the consultation mode is more ideal for learning as it breaks the session into chunks, a *Quick Health Check* mode was added as an alternative to cater to *busy users* (Figure 5).

Accommodating Female Users

Though ScreenMen was developed for men, some participants suggested that it could also be used by women as a woman might be the person taking care of her husband or father in a family. Some of the sentences were thus rephrased to accommodate female users; for example, *only men 18 years old or above should use this website* to *this website is meant for men 18 years old and above*.

Taking Into Account the Difference in Culture

Malaysia consists of 3 main ethnic groups: Malay, Chinese, and Indian. There were only 2 languages available in ScreenMen for beta testing (English and Malay). Some Chinese participants mentioned that their parents may need the Mandarin version as they were not literate in English and Malay languages. The Mandarin language was thus added to ScreenMen. No issue was raised regarding having Tamil language on ScreenMen as Indians are usually literate in English and Malay.

Apart from language, some sections of ScreenMen might be sensitive to certain ethnic groups. For example, all users were assessed in terms of alcohol intake, which may not be relevant to Muslim users as alcohol intake is prohibited in the religion. However, the Muslim participants reassured the research team that it was not an issue as the option *I never drink alcohol* was already in place.

Another concern was the sexually transmitted disease assessment. Personal information such as having multiple sexual partners, having sex with men, and injecting drugs was being asked of the users. However, the participants mentioned that they had no hindrance in answering these as no identifiable information was being recorded, and these were important for them to know.

Quantitative Evaluation (Questionnaire)

Only 23 participants answered the postintervention questionnaire as 1 participant was called for work urgently. The details of the postintervention quantitative evaluation are shown in Table 3. The SUS score obtained (mean 76.4, SD 7.72) indicated that the ScreenMen had good usability (good usability score range: 71.4-85.5) [42]. All participants agreed that they understood more about their health risks and health screening after using ScreenMen, and would recommend it to others.

For intention to undergo screening, 8 out of 23 men (35%) planned to attend screening earlier than intended after using the ScreenMen (no intention to 2 years, n=1; 5 years to 1 year, n=1; 2 years to 1 month, n=1; 1 year to 6 months, n=3; and 6 months to 1 month, n=2); 14 out of 23 men (61%) did not change; whereas 1 out of 23 men (4%) planned to screen later (1 month to 6 months). In terms of stage of behavior change, 4 out of 12 (33%) men, who were in precontemplation stage, changed to either contemplation or preparation stage after using ScreenMen (Table 4). However, the change from precontemplation (more than 6 months) to either preparation or contemplation stage (6 months or less) after using ScreenMen was not statistically significant different as McNemar test revealed P=.13. Furthermore, 11 out of 23 men were already in contemplation/preparation stage before using ScreenMen.

Figure 5. Options of consultation or quick mode for health assessment.

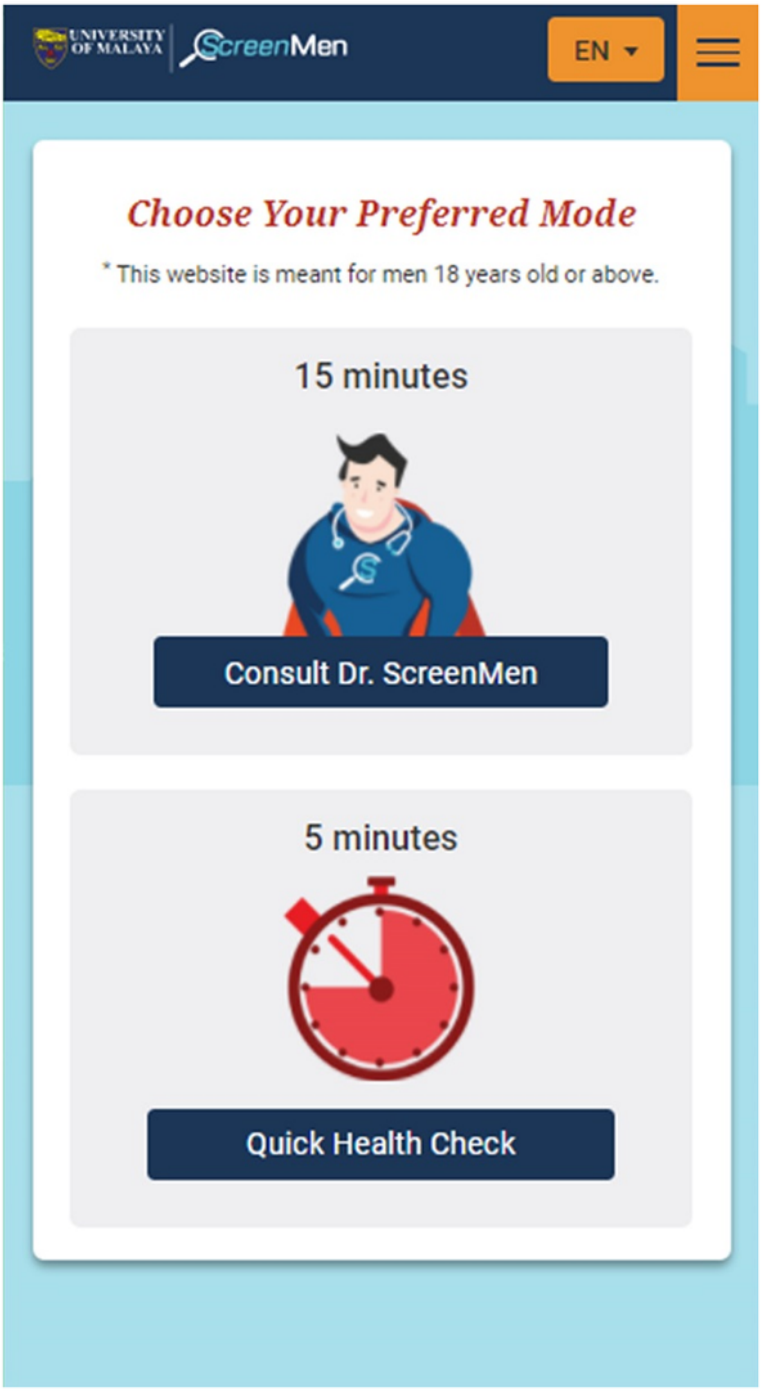


Table 3. Quantitative evaluation after using the ScreenMen (N=23).

Postintervention evaluation	Statistics
System Usability Scale score, mean (SD)	76.4 (7.72)
Understand more about their health risks, n (%)	23 (100)
Understand more about health screening, n (%)	23 (100)
Will recommend ScreenMen to others, n (%)	23 (100)
Change in the intended time to screen (before and after using ScreenMen), n (%)	
Earlier	8 (35)
No change	14 (61)
Later	1 (4)

Table 4. Intention to screen by stage of behavior change before and after using ScreenMen (N=23).

Stage of behavior change	Preintervention, n (%)	Postintervention, n (%)
Precontemplation (>6 months)	12 (52)	8 (35)
Preparation/contemplation (≤6 months)	11 (48)	15 (65)
Total	23 (100)	23 (100)

Discussion

Principal Findings

This study found that ScreenMen is acceptable to men in terms of its utility and usability. The participants found ScreenMen useful as they could learn more about their health risks and the evidence-based health screening they should go for. They also felt that ScreenMen was convenient to be used and may trigger men to undergo screening. The quantitative data showed that many men planned to undergo screening earlier than initially intended after using ScreenMen, although there was no significant difference when analyzed on the basis of the stage of behavior change using McNemar test. The participants also felt that ScreenMen was user-friendly and comfortable to look at. The SUS score also indicated that ScreenMen had good usability.

Tackling the issue of health screening is not an easy endeavor, and ScreenMen had to be further improved in terms of utility and usability. The key improvement in terms of utility was the addition of a reminder function. Past studies have shown that reminder interventions including those using a letter, email, and short message service were effective in increasing screening uptake [43,44]. ScreenMen also had to be framed in terms of *improving health* instead of *going for screening*, as men generally do not see the importance of health screening but would like to know more about staying healthy. As a result of the lack of interest in screening, ScreenMen had to be shortened and simplified to ensure that key messages were delivered in the shortest possible time. The most challenging part was advocating evidence-based health screening. Additional efforts and emphases needed to be placed for men to internalize the message to avoid unnecessary health screening. These were crucial for policy makers and researchers to consider when developing interventions in the future, particularly on topics that are surrounded by misconceptions and have low public interest.

This study has taken a male-sensitive approach to improve men's behavior. ScreenMen was developed specifically for men and incorporated with male-familiar contents such as the car maintenance concept and the *Superman-like* Dr ScreenMen figure. The car maintenance concept has been used in several programs globally to encourage health screening such as the Man MOT, which is a suite of Web-based health information and advice services, where men can chat with a National Health Services General Practitioner service anonymously on any health topic [45]. MOT stands for the Ministry of Transport, which was the responsible Ministry for the road worthiness test in the United Kingdom. It was used to name this program because of men's familiarity with it. Men reported that they felt empowered using it and were likely to use it again, especially as the first port of call for nonemergency health issues [45]. In addition, the use of [26,46-49] Dr ScreenMen avatar might be useful too, as there are preliminary studies that showed promising results of using avatars or embodied conversational agents (ECA) in health interventions, where an ECA is capable in engaging and motivating users in terms of learning and behavioral change [50-52].

The outcomes of this utility and usability testing may appear differently if ScreenMen was developed as a mobile app instead of a mobile Web app. The reminder function would be easily built, and more interesting functions such as alert, monitoring function, daily health messages, and integration with social media can be included. However, the research team decided to develop ScreenMen in the form of mobile Web app for the ease of dissemination. Though this hindered having more useful functions in ScreenMen, reaching out to men is seen as a more important factor, as a health screening mobile Web app or mobile app is not something being sought after by men as they do not see the importance of health screening, unlike for exercise or diet apps. A Web app has a broader dissemination than an app as it can be accessed instantly without needing to download and install, can be shared quickly among friends, and can be

viewed on a computer as well [53]. This factor is important to be considered by public health experts and researchers, especially when addressing health issues that are not seen to be important by the public.

The findings from this beta testing reinforced the importance of conducting testing with end users. Though many iterations of testing were done with experts during alpha testing, some of the issues were not captured. For example, the fruit intake per day question was not seen as a problem to experts but posed difficulties for users to answer. Other than that, the experts felt that the amount of text was just right; however, it was still too much for some participants. Apart from testing with experts, the development team has also considered many usability guidelines such as the International Organization for Standardization and the International Electrotechnical Commission (ISO/IEC) 25010 software product quality model and Nielsen usability diagram [29,54]. Nevertheless, many usability issues still emerged. The nuance of usability issues would only emerge during the in-depth beta testing with end users.

Strengths and Limitations

There are several strengths and limitations in this study. The strength of this study was that we managed to sample men from a wide demographic range, which gave rise to maximal variations of the qualitative findings. The multifaceted approach used (quantitative and qualitative, observation and retrospective think aloud with playback, IDI and FGD) allowed the study to gather a rich amount of data and enabled data triangulation. With regard to limitations, the design of the utility section in the questionnaire limited the data analysis. The questions on *understand more about health risks*, *understand more about health screening*, and *recommend this website to family and friends* should provide a Likert scale instead of *Yes* and *No* to enable a more meaningful analysis. For the intention to screen question, instead of fixing the options on the basis of the stage of behavior change, an open-ended field that allows participants to enter their actual number of months to screen would also allow better analysis. The sample size for quantitative analysis, though sufficient for the SUS as recommended by experts, was inadequate for the utility questions, especially the McNemar test for intention to screen. In addition, because of purposive sampling reason, about half of the participants were already in the contemplation or preparation stage even before using ScreenMen, which further diminished the analyzable sample size. Nevertheless, this study's primary focus was on the qualitative findings, which aimed to identify issues so that ScreenMen could be improved. The quantitative data were just the preliminary effectiveness findings, which will be measured more definitively in a trial. For the qualitative method, although ScreenMen was meant to be tested in a *real-world* setting, the researcher was present in the same room to observe and to assist the users, in case any technical issue occurred. This may affect how the users used ScreenMen, as they might feel being monitored and obliged to use ScreenMen properly, unlike at home. However, this gave more gain than loss, as observation on users' behavior provides very important data for probing during interview; nevertheless, solving technical issues is also important to prevent errors in the future.

Conclusions

This study showed that ScreenMen is acceptable to men in terms of its utility and usability. Men are able to learn more about their health risks and screening via ScreenMen. The preliminary data suggested that ScreenMen might increase men's intention to undergo screening and may potentially improve the actual uptake of health screening as well. Further evaluation in the form of randomized controlled trial should be conducted to determine the effectiveness of ScreenMen in improving the uptake of evidence-based health screening. Apart from that, this study also allowed further refinement of ScreenMen to improve its utility and usability. We have shared the key lessons learned from this beta testing, which might be useful for public health experts and researchers to develop user-centered mobile Web apps in the future.

Acknowledgments

The authors would like to thank all participants and the banking institution for approving this study to be conducted with the staff. This study was funded by the University of Malaya Research Programme (RP041A-15HTM) and partially funded by the Postgraduate Research Grant (PG265-2016A). Publication of this paper was cosponsored by the *Journal of Medical Internet Research*.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The framework used for data analysis.

[DOCX File, 13KB - mhealth_v7i4e10216_app1.docx]

References

1. GSM. 2018. The Mobile Economy 2018 URL: https://www.gsma.com/mobileeconomy/wp-content/uploads/2018/02/The-Mobile-Economy-Global-2018.pdf [accessed 2018-03-16] [WebCite Cache ID 6xxJlzuIV]

2. van den Berg MH, Schoones JW, Vliet VT. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
3. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of web-based vs non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
4. Fridriksdottir N, Gunnarsdottir S, Zoëga S, Ingadottir B, Hafsteinsdottir EJ. Effects of web-based interventions on cancer patients' symptoms: review of randomized trials. Support Care Cancer 2018 Feb;26(2):337-351. [doi: 10.1007/s00520-017-3882-6] [Medline: 28921391]
5. Välimäki M, Anttila K, Anttila M, Lahti M. Web-based interventions supporting adolescents and young people with depressive symptoms: systematic review and meta-analysis. JMIR Mhealth Uhealth 2017 Dec 8;5(12):e180 [FREE Full text] [doi: 10.2196/mhealth.8624] [Medline: 29222079]
6. Martorella G, Boitor M, Berube M, Fredericks S, Le MS, Gélinas C. Tailored web-based interventions for pain: systematic review and meta-analysis. J Med Internet Res 2017 Nov 10;19(11):e385 [FREE Full text] [doi: 10.2196/jmir.8826] [Medline: 29127076]
7. Sorgente A, Pietrabissa G, Manzoni GM, Re F, Simpson S, Perona S, et al. Web-based interventions for weight loss or weight loss maintenance in overweight and obese people: a systematic review of systematic reviews. J Med Internet Res 2017 Jun 26;19(6):e229 [FREE Full text] [doi: 10.2196/jmir.6972] [Medline: 28652225]
8. Birney AJ, Gunn R, Russell JK, Ary DV. MoodHacker mobile web app with email for adults to self-manage mild-to-moderate depression: randomized controlled trial. JMIR Mhealth Uhealth 2016 Jan 26;4(1):e8 [FREE Full text] [doi: 10.2196/mhealth.4231] [Medline: 26813737]
9. Hawkins SD, Barilla S, Feldman SR. Web app based patient education in psoriasis - a randomized controlled trial. Dermatol Online J 2017 Apr 15;23(4). [Medline: 28541882]
10. Irvine AB, Russell H, Manocchia M, Mino DE, Cox GT, Morgan R, et al. Mobile-web app to self-manage low back pain: randomized controlled trial. J Med Internet Res 2015;17(1):e1 [FREE Full text] [doi: 10.2196/jmir.3130] [Medline: 25565416]
11. Institute for Quality and Efficiency in Health Care. NCBI. 2016. Benefits and risks of screening tests URL: https://www.ncbi.nlm.nih.gov/books/NBK279418/ [accessed 2019-03-05] [WebCite Cache ID 76cetBdVo]
12. Dryden R, Williams B, McCowan C, Themessl-Huber M. What do we know about who does and does not attend general health checks? Findings from a narrative scoping review. BMC Public Health 2012 Aug 31;12:723 [FREE Full text] [doi: 10.1186/1471-2458-12-723] [Medline: 22938046]
13. Mao L, de Wit JB, Kippax SC, Prestage G, Holt M. Younger age, recent HIV diagnosis, no welfare support and no annual sexually transmissible infection screening are associated with nonuse of antiretroviral therapy among HIV-positive gay men in Australia. HIV Med 2015 Jan;16(1):32-37 [FREE Full text] [doi: 10.1111/hiv.12169] [Medline: 24889053]
14. Borkhoff CM, Saskin R, Rabeneck L, Baxter NN, Liu Y, Tinmouth J, et al. Disparities in receipt of screening tests for cancer, diabetes and high cholesterol in Ontario, Canada: a population-based study using area-based methods. Can J Public Health 2013 Jun 21;104(4):e284-e290. [doi: 10.17269/cjph.104.3699] [Medline: 24044467]
15. Institute for Public Health. National Health and Morbidity Survey 2011 Vol III: Healthcare Demand and Out-of-pocket Health Expenditure. Kuala Lumpur: National Institutes of Health, Ministry of Health Malaysia; 2011.
16. Institute for Public Health. National Health and Morbidity Survey 2015 Vol II: Non-Communicable Diseases, Risk Factors & Other Health Problems. Kuala Lumpur: National Institutes of Health, Ministry of Health Malaysia; 2015.
17. Teo CH, Ng CJ, Booth A, White A. Barriers and facilitators to health screening in men: a systematic review. Soc Sci Med 2016 Dec;165:168-176. [doi: 10.1016/j.socscimed.2016.07.023] [Medline: 27511617]
18. Teo CH, Ng CJ, White A. Factors influencing young men's decision to undergo health screening in Malaysia: a qualitative study. BMJ Open 2017 Dec 10;7(3):e014364 [FREE Full text] [doi: 10.1136/bmjopen-2016-014364] [Medline: 28283491]
19. Teo CH, Ling CJ, Ng CJ. Improving health screening uptake in men: a systematic review and meta-analysis. Am J Prev Med 2018 Jan;54(1):133-143. [doi: 10.1016/j.amepre.2017.08.028] [Medline: 29254551]
20. Bauermeister JA, Pingel ES, Jadwin-Cakmak L, Harper GW, Horvath K, Weiss G, et al. Acceptability and preliminary efficacy of a tailored online HIV/STI testing intervention for young men who have sex with men: the Get Connected! program. AIDS Behav 2015 Feb 1;19(10):1860-1874. [doi: 10.1007/s10461-015-1009-y] [Medline: 25638038]
21. Ilic D, Egberts K, McKenzie JE, Risbridger G, Green S. Informing men about prostate cancer screening: a randomized controlled trial of patient education materials. J Gen Intern Med 2008 Apr;23(4):466-471 [FREE Full text] [doi: 10.1007/s11606-007-0466-z] [Medline: 18373146]
22. Young SD, Cumberland WG, Nianogo R, Menacho LA, Galea JT, Coates T. The HOPE social media intervention for global HIV prevention in Peru: a cluster randomised controlled trial. Lancet HIV 2015 Jan;2(1):e27-e32 [FREE Full text] [doi: 10.1016/S2352-3018(14)00006-X] [Medline: 26236767]
23. United States Preventive Services Task Force. 2017. USPSTF A and B recommendations URL: https://www.uspreventiveservicestaskforce.org/Page/Name/uspstf-a-and-b-recommendations/ [accessed 2018-06-04] [WebCite Cache ID 6wx0iwOBz]

24. Teo CH, Ng CJ, White A. What do men want from a health screening mobile app? A qualitative study. PLoS One 2017;12(1):e0169435 [FREE Full text] [doi: 10.1371/journal.pone.0169435] [Medline: 28060953]
25. World Health Organization. Gender and Health Promising Practices Series: A Framework. Geneva, Switzerland; 2010.
26. WHO Europe. 2001. Madrid Statement. Mainstreaming gender equity in health: The need to move forward URL: http://www.euro.who.int/document/a75328.pdf [accessed 2018-05-12] [WebCite Cache ID 6wx0SgVUu]
27. Hong Y, Goldberg D, Dahlke D, Ory M, Cargill J, Coughlin R, et al. Testing usability and acceptability of a web application to promote physical activity (iCanFit) among older adults. JMIR Hum Factors 2014 Oct 13;1(1):e2 [FREE Full text] [doi: 10.2196/humanfactors.3787] [Medline: 27025254]
28. Schmiedl G, Seidl M, Temper K. Mobile phone web browsing: a study on usage and usability of the mobile web. In: Proceedings of the 11th international Conference on Human-Computer interaction with Mobile Devices and Services. 2009 Presented at: MobileHCI'09; September 15-18, 2009; Bonn, Germany.
29. Nielsen J. Usability Engineering. Boston: Morgan Kaufmann; 1994.
30. Lee YK, Lee PY, Ng CJ, Teo CH, Abu Bakar AI, Abdullah KL, et al. Usability and utility evaluation of the web-based "Should I Start Insulin?" patient decision aid for patients with type 2 diabetes among older people. Inform Health Soc Care 2018 Jan;43(1):73-83. [doi: 10.1080/17538157.2016.1269108] [Medline: 28139158]
31. Usability.gov. 2018. Running a Usability Test URL: https://www.usability.gov/how-to-and-tools/methods/running-usability-tests.html [accessed 2018-03-16] [WebCite Cache ID 6xxKj6AAa]
32. van den Haak M, De Jong M, Jan Schellens P. Retrospective vs concurrent think-aloud protocols: testing the usability of an online library catalogue. Behav Inf Technol 2003 Sep;22(5):339-351. [doi: 10.1080/0044929031000]
33. Giorgi G, Arcangeli G, Perminiene M, Lorini C, Ariza-Montes A, Fiz-Perez J, et al. Work-related stress in the banking sector: a review of incidence, correlated factors, and major consequences. Front Psychol 2017;8:2166 [FREE Full text] [doi: 10.3389/fpsyg.2017.02166] [Medline: 29312044]
34. Makhbul ZM, Alam SS, Azmi SM, Talib NA. Ergonomics and work stress issues in banking sector. Aust J Basic Appl Sci 2011;5(9):1301-1309 [FREE Full text]
35. Giga IS, Hoel H. Violence and Stress at Work in Financial Services. Geneva: International Labor Office; 2003.
36. Silva JL, Navarro VL. Work organization and the health of bank employees. Rev Lat Am Enfermagem 2012;20(2):226-234 [FREE Full text] [Medline: 22699722]
37. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/BF03195514] [Medline: 14587545]
38. Nielsen J. Nielsen Norman Group. 2006. Quantitative Studies: How Many Users to Test? URL: https://www.nngroup.com/articles/quantitative-studies-how-many-users/ [accessed 2018-02-03] [WebCite Cache ID 6wx0dB9PB]
39. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
40. Brooke J. SUS: a quick and dirty usability scale. In: Jordan P, Thomas B, McClelland I, Weerdmesster B, editors. Usability Evaluation in Industry. Philadelphia: Taylor and Francis; 1996.
41. 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]
42. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
43. Muller D, Logan J, Dorr D, Mosen D. The effectiveness of a secure email reminder system for colorectal cancer screening. AMIA Annu Symp Proc 2009 Nov 14;2009:457-461. [doi: 10.6083/M4KS6PJR] [Medline: 20351899]
44. Bourne C, Knight V, Guy R, Wand H, Lu H, McNulty A. Short message service reminder intervention doubles sexually transmitted infection/HIV re-testing rates among men who have sex with men. Sex Transm Infect 2011 Apr;87(3):229-231. [doi: 10.1136/sti.2010.048397] [Medline: 21296796]
45. Baker P. Man MOT: a new approach to primary care for men. Trends Urol Mens Health 2017;8(1):13-16 [FREE Full text] [doi: 10.1002/tre.559]
46. Baker P. Men's health: a global problem requiring global solutions. Trends Urol Mens Health 2016 May 25;7(3):11-14. [doi: 10.1002/tre.519]
47. Baker P. Men's health: an overlooked inequality. Br J Nurs 2016 Oct 27;25(19):1054-1057. [doi: 10.12968/bjon.2016.25.19.1054] [Medline: 27792439]
48. Rovito MJ, Leonard B, Llamas R, Leone JE, Talton W, Fadich A, et al. A call for gender-inclusive global health strategies. Am J Mens Health 2017 Nov;11(6):1804-1808 [FREE Full text] [doi: 10.1177/1557988317723424] [Medline: 28884639]
49. The European Men's Health Forum. The Vienna declaration on the health of men and boys in Europe. J Mens Health 2005 Dec;2(4):446-447 [FREE Full text] [doi: 10.1016/j.jmhg.2005.10.004]
50. Martínez-Miranda J. Embodied conversational agents for the detection and prevention of suicidal behaviour: current applications and open challenges. J Med Syst 2017 Sep;41(9):135. [doi: 10.1007/s10916-017-0784-6] [Medline: 28755270]
51. Provoost S, Lau H, Ruwaard J, Riper H. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res 2017 Dec 9;19(5):e151 [FREE Full text] [doi: 10.2196/jmir.6553] [Medline: 28487267]

52. Scholten M, Kelders S, Van Gemert-Pijnen JE. Self-guided web-based interventions: scoping review on user needs and the potential of embodied conversational agents to address them. J Med Internet Res 2017 Dec 16;19(11):e383 [FREE Full text] [doi: 10.2196/jmir.7351] [Medline: 29146567]
53. Budiu R. Nielsen Norman Group. 2013. Mobile: Native Apps, Web Apps, and Hybrid Apps URL: https://www.nngroup.com/articles/mobile-native-apps/ [accessed 2019-03-05] [WebCite Cache ID 76cn3fGPi]
54. International Organization for Standardization. 2017. ISO/IEC 25010: Usability URL: http://iso25000.com/index.php/en/iso-25000-standards/iso-25010?limit=3&start=3 [WebCite Cache ID 6wx0bQqMH]

Abbreviations

ECA: embodied conversational agent
FGD: focus group discussion
IDI: in-depth interview
MYR: Malaysian Ringgit
NHMS: National Health and Morbidity Survey
SUS: System Usability Scale
USPSTF: United States Preventive Services Task Force

Edited by G Eysenbach; submitted 19.03.18; peer-reviewed by M Anttila, A Panzeri, L Marciano; comments to author 06.10.18; revised version received 31.12.18; accepted 25.01.19; published 15.04.19.

Please cite as:
Teo CH, Ng CJ, Lo SK, Lim CD, White A
A Mobile Web App to Improve Health Screening Uptake in Men (ScreenMen): Utility and Usability Evaluation Study
JMIR Mhealth Uhealth 2019;7(4):e10216
URL: https://mhealth.jmir.org/2019/4/e10216/
doi:10.2196/10216
PMID:30985280

Original Paper

A Noninvasive, Economical, and Instant-Result Method to Diagnose and Monitor Type 2 Diabetes Using Pulse Wave: Case-Control Study

Yiming Hao[1*], MD; Feng Cheng[2*], PhD; Minh Pham[3], MA; Hayley Rein[4], BSc; Devashru Patel[5], UG; Yuchen Fang[1], UG; Yiyi Feng[1], UG; Jin Yan[6], BMed; Xueyang Song[1], BMed; Haixia Yan[1], MD; Yiqin Wang[1], MD

[1]Shanghai Key Laboratory of Health Identification and Assessment/Laboratory of Traditional Chinese Medicine Four Diagnostic Information, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[2]Department of Pharmaceutical Sciences, College of Pharmacy, University of South Florida, Tampa, FL, United States

[3]Department of Mathematics and Statistics, College of Arts and Sciences, University of South Florida, Tampa, FL, United States

[4]Department of Cellular and Molecular Biology, College of Arts and Sciences, University of South Florida, Tampa, FL, United States

[5]College of Computing, Georgia Institute of Technology, Atlanta, GA, United States

[6]Jiangsu Province Hospital of Chinese Medicine, Nanjing, China

[*]these authors contributed equally

Corresponding Author:
Yiqin Wang, MD
Shanghai Key Laboratory of Health Identification and Assessment/Laboratory of Traditional Chinese Medicine Four Diagnostic Information
Shanghai University of Traditional Chinese Medicine
1200 Cailun Road
Shanghai,
China
Phone: 86 21 51322447
Email: wangyiqin2380@sina.com

Abstract

Background: We should pay more attention to the long-term monitoring and early warning of type 2 diabetes and its complications. The traditional blood glucose tests are traumatic and cannot effectively monitor the development of diabetic complications. The development of mobile health is changing rapidly. Therefore, we are interested in developing a new noninvasive, economical, and instant-result method to accurately diagnose and monitor type 2 diabetes and its complications.

Objective: We aimed to determine whether type 2 diabetes and its complications, including hypertension and hyperlipidemia, could be diagnosed and monitored by using pulse wave.

Methods: We collected the pulse wave parameters from 50 healthy people, 139 diabetic patients without hypertension and hyperlipidemia, 133 diabetic patients with hypertension, 70 diabetic patients with hyperlipidemia, and 75 diabetic patients with hypertension and hyperlipidemia. The pulse wave parameters showing significant differences among these groups were identified. Various machine learning models such as linear discriminant analysis, support vector machines (SVMs), and random forests were applied to classify the control group, diabetic patients, and diabetic patients with complications.

Results: There were significant differences in several pulse wave parameters among the 5 groups. The parameters height of tidal wave (h_3), time distance between the start point of pulse wave and dominant wave (t_1), and width of percussion wave in its one-third height position (W) increase and the height of dicrotic wave (h_5) decreases when people develop diabetes. The parameters height of dominant wave (h_1), h_3, and height of dicrotic notch (h_4) are found to be higher in diabetic patients with hypertension, whereas h_5 is lower in diabetic patients with hyperlipidemia. For detecting diabetes, the method with the highest out-of-sample prediction accuracy is SVM with polynomial kernel. The algorithm can detect diabetes with 96.35% accuracy. However, all the algorithms have a low accuracy when predicting diabetic patients with hypertension and hyperlipidemia (below 70%).

Conclusions: The results demonstrated that the noninvasive and convenient pulse-taking diagnosis described in this paper has the potential to become a low-cost and accurate method to monitor the development of diabetes. We are collecting more data to improve the accuracy for detecting hypertension and hyperlipidemia among diabetic patients. Mobile devices such as sport bands,

smart watches, and other diagnostic tools are being developed based on the pulse wave method to improve the diagnosis and monitoring of diabetes, hypertension, and hyperlipidemia.

(JMIR Mhealth Uhealth 2019;7(4):e11959) doi:10.2196/11959

KEYWORDS

type 2 diabetes; hypertension; hyperlipidemia; pulse wave analysis; diagnosis

Introduction

Background

Diabetes is becoming one of the most severe health problems in China. The World Health Organization indicated that type 2 diabetes accounts for around 90% of all the diabetes cases worldwide. In 2016, 9.4% of the population of China was diabetic, and 2% of all deaths in the country were because of diabetes and its chronic complications [1,2]. Type 2 diabetes and a variety of chronic complications in the middle and late stages of the disease develop over a long time. Therefore, early diagnosis, prevention, and treatment of type 2 diabetes are vital for reducing the medical burden and mortality rate. Diabetes is a long-standing example of a disease whose patients have been positively impacted by traditional Chinese medicine (TCM) [3]. Accurate diagnosis through the noninvasive, convenient, and economical techniques of TCM allows effective prevention and treatment of type 2 diabetes.

In TCM diagnosis, pulse taking is an important skill for diagnosing diseases by touching and sensing radial pulsations, but it frequently depends on the subjective consciousness and experience accumulation of the doctors. It is the lack of objective criteria that reduces the accuracy and repeatability of diagnosis. To overcome the shortcomings of subjectivity of traditional pulse diagnosis, in recent years, the importance of objective pulse-taking diagnosis has gained more and more attention. In addition, the study of diagnosis of some chronic diseases such as coronary heart disease and lung cancer by using the objective pulse parameters had some progress [4,5]. In the study of type 2 diabetes, some researchers have found that the changes of radial artery pulse wave are related to elevated blood glucose levels and major adverse cardiovascular events caused by diabetes mellitus [6,7]. Therefore, we hope that the noninvasive, convenient, and objective pulse information detection can be used to assist conventional methods to diagnose and monitor the occurrence and development of type 2 diabetes.

Objectives

The purpose of this study was to use a *TCM pulse informatics analysis system* to measure people's objective pulse information and use that data to develop a model to diagnose type 2 diabetes. This instrument uses a pressure sensor to record pulse beats and display this information within a pulse wave. The pressure sensor is the most commonly used device for recording pulse wave of people's radial artery at present. The parameters of the pulse wave are extracted and analyzed using statistical methods. This approach can quantitatively analyze pulse signals and provide more objective results than traditional pulse-taking diagnosis by TCM doctors.

This is the first study to use this method to analyze the differences in pulse wave parameters between healthy individuals and diabetic patients with hypertension and hyperlipidemia. The purpose of our research was (1) to find the association of the objective pulse information with type 2 diabetes and the disease with hypertension or hyperlipidemia and (2) to establish the prediction models of diabetes and its complications. We hypothesized that diabetes and its combination with hypertension and hyperlipidemia can be diagnosed reliably by using the pulse wave parameters.

Methods

Patients

This is a case-control study. Patients with type 2 diabetes (referred to as diabetes from now on) were recruited from outpatient services in 4 hospitals in Shanghai, including Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shuguang Hospital, Shanghai Municipal Hospital of TCM, and Shanghai Qiangsheng Worker's Hospital, between April 2012 and December 2018. The individuals in the healthy control group (group 1, n=50) were healthy staff and graduate students from the Shanghai University of TCM. The subjects were of Han ethnicity.

On the basis of blood pressure and serum lipid levels, diabetic patients were divided into 4 groups, including diabetic patients without hypertension and hyperlipidemia (group 2, n=139), diabetic patients with hypertension (group 3, n=133), diabetic patients with hyperlipidemia (group 4, n=70), and diabetic patients with hypertension and hyperlipidemia (group 5, n=75).

There was no significant difference in the age (P=.13) or gender (P=.59) between healthy control individuals (group 1) and all diabetic patients (group 2-5) (Table 1).

Table 1. Summary of demographics and clinic characteristics of each group.

Group	Cases	Ratio of male to female	Average age (years)	Average course of disease (years)
Group 1	50	1:1.27	61.40 (SD 10.08)	NA[a]
Group 2-5	417	1:1.40	63.67 (SD 11.15)	7.36 (SD 6.44)
Group 2	139	1:1.62	61.10 (SD 11.34)	6.28 (SD 5.99)
Group 3	133	1:1.15	67.97 (SD 10.22)	8.04 (SD 6.44)
Group 4	70	1:1.26	59.20 (SD 11.29)	6.22 (SD 5.26)
Group 5	75	1:1.68	64.96 (SD 9.37)	8.63 (SD 7.55)

[a]NA: not applicable.

Ethics Approval

The study was approved by the Ethics Committee of Shanghai University of TCM in China in January 2012 and performed in accordance with the Declaration of Helsinki. All subjects had signed informed consent agreements.

Criteria

Diagnostic Criteria

The diagnostic criteria of diabetes were referred to as the *Standards of medical care in diabetes* [8].

Inclusion Criteria

The inclusion criteria were as follows:

1. Meet the diagnostic standard of diabetes
2. Age range is from 40 to 75 years
3. Hypertension or hyperlipidemia occur after diabetes.

Exclusion Criteria

The exclusion criteria were as follows:

1. Women during their pregnancy and lactation period
2. Mentally ill individuals
3. Those complicated with other diseases
4. Individuals who had acute metabolic disorders such as diabetic ketoacidosis or inflammatory complications.

Elimination Criteria

Patients without complete clinical data either because of incomplete collection or missing data were eliminated.

Collecting Methods for the Objective Parameters of Traditional Chinese Medicine Diagnosis

All objective parameters of TCM pulse diagnosis were collected by 2 MD doctors using the TCM pulse informatics analysis system (type: Smart TCM-I, product by: Shanghai Asia & Pacific Computer Information System CO, Ltd, Shanghai, China; Figure 1). In the analysis system, the instrument of detecting pulse wave information is a wristband acquisition terminal (Figure 2). The device consists of a pulse sensor, adapter, and acquisition software. The pulse sensor is attached to the arm by a wristband and connected to the adapter by a cable. The adapter has a universal serial bus (USB) connector, which can be directly connected to the computer through the USB interface. The acquisition software runs on the computer and realizes the acquisition and data management function of the pulse wave image by cooperating with the hardware (Figure 3).

The indoor temperature was 18°C to 25°C during pulse wave information collection. The subjects were either sitting or lying down for at least 3 min before they were tested. For pulse-taking diagnosis data collection, the forearms of the patients were extended forward naturally and at the same height as the heart. Their wrists were kept straight, with palms upward and fingers slightly bent. A small, soft pillow was placed under the wrist joint for better data collection. The pulse was measured at the radial artery corresponding to the inside of the styloid process of the radius in the left hand, the best position to feel the pulse. The frequency of the pulse acquisition device was 720 Hz. The sensor was tied to the position of the radial artery, and the pressure was adjusted by a knob. The acquisition was realized through observation and operation on the software interface. The software interface displayed the dynamic pulse waveform and pressure value in real time. The software recorded the pulse wave parameters for 1 min when the amplitude of the pulse wave reaches the maximum value. Afterward, the parameters of the pulse wave were extracted through the built-in software.

The data collected could be presented as a time-domain pulse wave (Figure 4), which includes the height of dominant wave (h_1), height of tidal wave (h_3), height of dicrotic notch (h_4), height of dicrotic wave (h_5), time distance between the start point of pulse wave and dominant wave (t_1), time distance between the start point of pulse wave and dicrotic notch (t_4), time distance between dicrotic notch and the end point of pulse wave (t_5), and width of percussion wave in its one-third height position (W).

The pulse wave parameters, including h_1, h_3, h_4, h_5, t_1, t_4, t_5, and W, have different physiological significances (Table 2).

Disease Prediction Methods and Statistical Analysis

The first objective was to analyze the differences in pulse wave parameters between healthy individuals and diabetic patients, and among the 4 subgroups of diabetic patients. One-way analysis of variance (ANOVA) was applied to the pulse wave parameters that follow the normal distribution, and non-parametric methods including the Mann-Whitney U test (for 2 groups) and the Kruskal-Wallis one-way ANOVA test (for 4 groups) were used for non-normal data.

The second objective was to classify diabetic patients from healthy individuals and to detect hypertension and

hyperlipidemia among diabetic patients by using the pulse wave parameters. The following machine learning algorithms were applied:

- Logistic regression using all the pulse wave parameters as predictors
- Linear discriminant analysis (LDA) using all the pulse wave parameters as predictors
- Random forests with 500 decision trees, each tree samples two-thirds of the data, and each split in a decision tree randomly samples 3 out of 11 pulse wave parameters.
- Support vector machines (SVMs): We applied SVM with linear kernel (SVM-Linear) and SVM with 3-degree polynomial kernel (SVM-Poly). The cost and gamma parameters are determined through cross-validation, with trial values for cost being 0.001, 0.01, 0.1, 1, 5, 10, and 100 and trial values for gamma being 0.001, 0.01, and 0.1.

To estimate the prediction accuracy of the models (out-of-sample accuracy), we applied the 10-fold cross-validation technique.

Figure 1. Traditional Chinese medicine pulse life informatics analysis system.

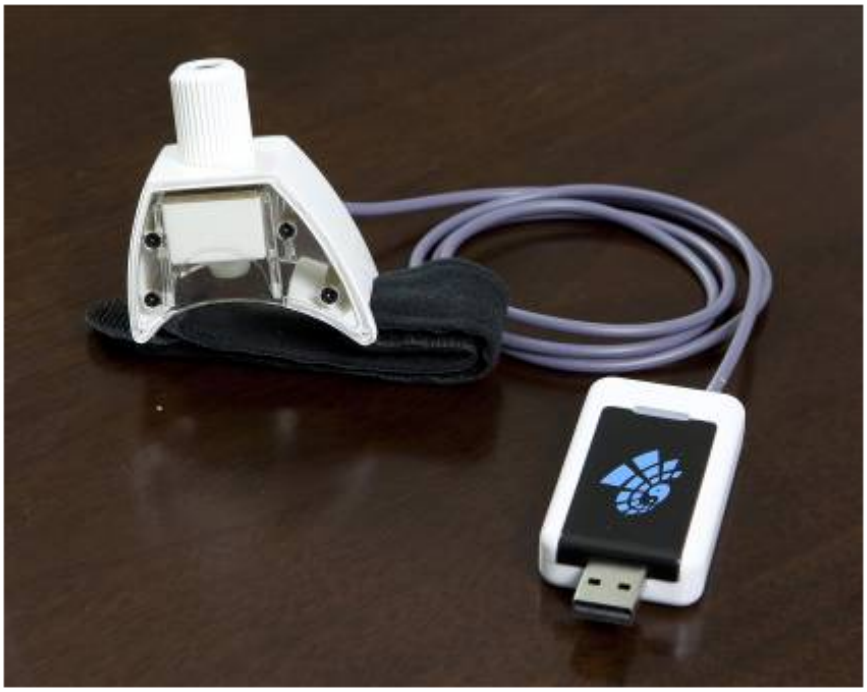


Figure 2. The wristband pulse wave information acquisition terminal.

Figure 3. The pulse wave image collected.

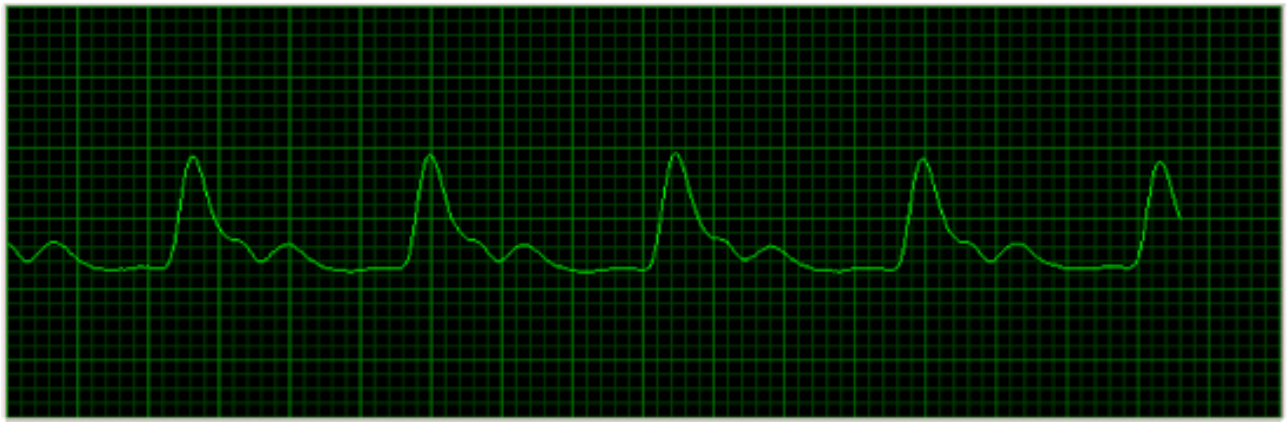


Figure 4. A typical single pulse wave. h: height; h_1: height of dominant wave; h_3: height of tidal wave; h_4: height of dicrotic notch; h_5: height of dicrotic wave; mm: millimeter; t: time; s: second; t_1: time distance between the start point of pulse wave and dominant wave; t_4: time distance between the start point of pulse wave and dicrotic notch; t_5: time distance between dicrotic notch and the end point of pulse wave; W: width of percussion wave in its one-third height position.

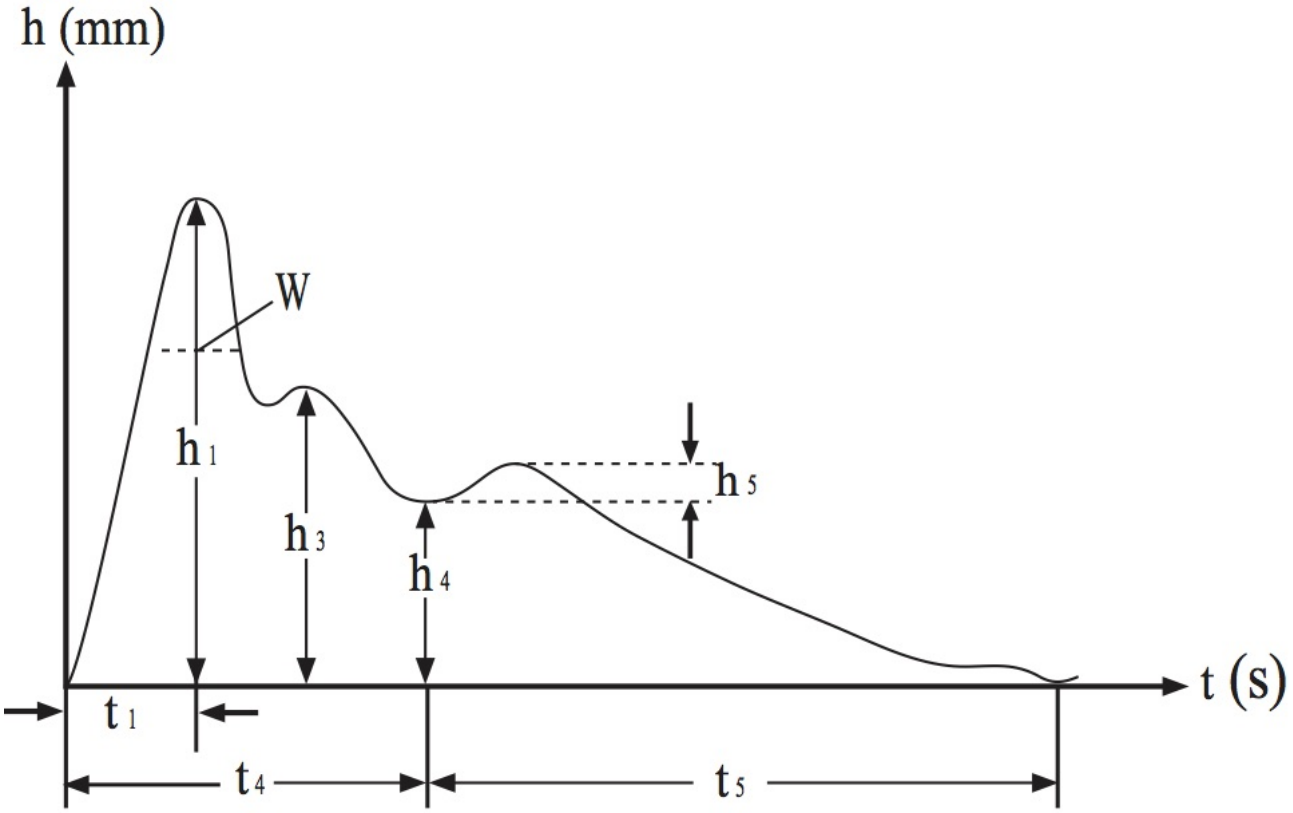


Table 2. Physiological significance of pulse wave parameters.

Pulse wave parameters	Physiological significance
Height parameters of pulse wave	
Height of dominant wave (h_1)	High value of h_1 reflects the strong elasticity of large artery and good ejection function of the left ventricle
Height of tidal wave (h_3)	High value of h_3 reflects the weak elasticity and/or high peripheral resistance of the artery
Height of dicrotic notch (h_4)	High value of h_4 reflects the high diastolic blood pressure, high peripheral resistance of the artery, and/or weak closing function of the aortic valve
Height of dicrotic wave (h_5)	High value of h_5 reflects the strong elasticity of the large artery and good closing function of the aortic valve
Time parameters of pulse phase	
Time distance between the start point of pulse wave and dominant wave (t_1)	High value of t_1 reflects the weak elasticity of large vessels and/or high tensioning of small vessels
Time distance between the start point of pulse wave and dicrotic notch (t_4)	High value of t_4 reflects the good systolic function of the heart
Time distance between dicrotic notch and the end point of pulse wave (t_5)	High value of t_5 reflects the weak elasticity and/or high peripheral resistance of large vessels
Width of percussion wave in its one-third height position (W)	High value of W reflects the long time maintained by the high pressure in the artery

Results

Difference in Pulse Wave Parameters Between Groups

Tables 3 and 4 show the ANOVA test and nonparametric test results on the pulse wave parameters between groups.

The results indicated that all diabetic patients (groups 2-5) had significantly higher h_3, t_1, and W than those in the healthy control group, whereas h_5 was significantly lower in diabetics when compared with healthy individuals. The parameters of pulse wave height, h_1, h_3, h_4, and h_5, were significantly different in groups 2 to 5. On the contrary, there were no significant differences in parameters of time distance, t_1, t_4, t_5, and W, among these groups. Pairwise comparisons indicated that the h_1 values of diabetic patients with hypertension (groups 3 and 5) were significantly higher than those of diabetic patients with hyperlipidemia (group 4). In addition, the h_1 values of diabetic patients with hypertension (group 3) were also significantly higher than those of diabetic patients without hypertension and hyperlipidemia (group 2). Diabetic patients without hypertension (groups 2 and 4) showed a significantly lower h_3 parameter than those of diabetic patients with hypertension (groups 3 and 5). The h_4 parameter of group 2 was significantly lower than that of groups 5. Similar to h_1 and h_3, h_4 values of groups 3 and 5 were also significantly higher than those of group 4. The value of parameter h_5 in group 5 was significantly lower than that of parameter h_5 in groups 2, 3, and 4.

Table 3. Comparison of pulse wave parameters between group 1 and group 2-5.

Parameter	Group 1	Groups 2-5	*P* value
Height of dominant wave (h_1) (mm)	11.58 (9.65-13.68)	11.98 (8.25-16.60)	.48
Height of tidal wave (h_3) (mm)	7.74 (6.54-9.96)	9.83 (6.71-13.95)	.001
Height of dicrotic notch (h_4) (mm)	4.69 (4.24-6.25)	5.47 (3.32-7.36)	.44
Height of dicrotic wave (h_5) (mm)	5.48 (4.66-6.70)	2.48 (0.24-5.00)	<.001
Time distance between the start point of pulse wave and dominant wave (t_1) (s)	0.11 (0.10-0.12)	0.14 (0.12-0.18)	<.001
Time distance between the start point of pulse wave and dicrotic notch (t_4) (s)	0.33 (0.30-0.35)	0.33 (0.31-0.36)	.24
Time distance between dicrotic notch and the end point of pulse wave (t_5) (s)	0.45 (0.42-0.54)	0.49 (0.42-0.59)	.15
Width of percussion wave in its one-third height position (W) (s)	0.17 (0.13-0.22)	0.21 (0.18-0.23)	<.001

Table 4. Comparison of pulse wave parameters between groups 2, 3, 4, and 5.

Parameter	Group 2	Group 3	Group 4	Group 5	*P* value
Height of dominant wave (h_1) (mm)	11.56 (8.07-16.17)	13.04 (9.82-18.21)[a]	10.80 (7.39-14.22)[b]	12.65 (8.50-18.77)[c]	.01
Height of tidal wave (h_3) (mm)	8.87 (6.55-13.35)	10.89 (7.33-14.99)[a]	8.98 (6.02-12.18)[b]	10.74 (7.78-15.42)[b,c]	.01
Height of dicrotic notch (h_4) (mm)	4.66 (3.29-6.97)	5.84 (3.48-7.77)	4.78 (2.65-6.68)[b]	5.87 (3.70-8.60)[a□]	.02
Height of dicrotic wave (h_5) (mm)	2.86 (0.88-5.05)	2.80 (0.22-5.60)	2.42 (1.00-4.51)	1.17 (-0.60-3.85)[a,b,c]	.004
Time distance between the start point of pulse wave and dominant wave (t_1) (s)	0.14 (0.12-0.18)	0.14 (0.12-0.18)	0.14 (0.12-0.19)	0.14 (0.12-0.16)	.66
Time distance between the start point of pulse wave and dicrotic notch (t_4) (s)	0.33 (0.31-0.35)	0.33 (0.30-0.36)	0.33 (0.31-0.35)	0.33 (0.32-0.36)	.67
Time distance between dicrotic notch and the end point of pulse wave (t_5) (s)	0.48 (0.41-0.60)	0.48 (0.42-0.61)	0.51 (0.43-0.59)	0.49 (0.43-0.55)	.87
Width of percussion wave in its one-third height position (W) (s)	0.21 (0.18-0.23)	0.21 (0.18-0.24)	0.21 (0.18-0.24)	0.21 (0.18-0.22)	.87

[a]Means compared with group 2; *P*<.05.

[b]Means compared with group 3; *P*<.05.

[c]Means compared with group 4; *P*<.05.

Table 5. Physiological significance of pulse wave parameters.

Method	Accuracy to detect diabetes	Accuracy to detect hypertension	Accuracy to detect hyperlipidemia
Logistic regression	0.9293	0.5920	0.6500
Linear discriminant analysis	0.9037	0.5944	0.6500
Random forests	0.9294	0.5697	0.6977
SVM[a] with linear kernel	0.9421	0.5780	0.6572
SVM with polynomial kernel	0.9635	0.5858	0.6821

[a]SVM: support vector machine.

Table 6. Algorithm statistics.

Diagnosis of	Method used	Accuracy	Sensitivity	Specificity
Diabetes	SVM[a] with polynomial kernel	0.9635	0.8571	0.9535
Hypertension	Linear discriminant analysis	0.5944	0.7419	0.5429
Hyperlipidemia	Random forests	0.6977	0.7333	0.6190

[a]SVM: support vector machine.

Classification Algorithms

The accuracies of the classification algorithms are presented in Table 5.

Table 4 indicates that SVM with polynomial kernel (tuned parameters are cost=100 and gamma=0.01) has the highest prediction accuracy for diabetes, LDA has the highest prediction accuracy for hypertension among diabetic patients, and random forests has the highest prediction accuracy for hyperlipidemia among diabetic patients. The maximum accuracy, sensitivity, and specificity for each diagnosing task are presented in Table 6.

Discussion

Principal Findings

Pulse wave is the track of radial artery pulsation. It integrates a large amount of useful information about heart ejection activity and how the pulse wave travels along the vascular tree [9]. In this paper, time-domain analysis, one of the most frequently used techniques for pulse research [9], was applied to study the association between pulse parameters (including the height of pulse wave and the time of pulse phase) and type 2 diabetes to predict the attack of type 2 diabetes and its complications (hypertension and hyperlipidemia) for the first time.

Hyperglycemia is a leading cause of cardiovascular disease [10]. Therefore, the incidence of atherosclerosis is increased 2 to 4 folds in diabetic patients compared with nondiabetic individuals [11]. These diabetic patients experience hardened arterial blood vessels and decreased arterial elasticity, which increases peripheral resistance and the duration of high pressure in blood vessels. In this paper, we found that the parameters of h_3, t_1, and W increased significantly in diabetic patients compared with healthy individuals. From the result, we know that the high value of h_3 gives rise to weak elasticity and high peripheral resistance of arterial blood vessels. The high t_1 value reflects the weak elasticity of large vessels and high tensioning of small vessels, and the high W value denotes that high pressure in the artery is maintained for a longer period. Diversely, the value of h_5 was significantly lower in diabetic patients when compared with their healthy counterparts. The high h_5 value reflects the strong elasticity of the large artery. Our results from pulse time-domain analysis agreed well with the findings on patients and rats with diabetes [12,13].

Our analysis revealed that diabetic patients with hyperlipidemia had lower h_5 values compared with patients without hyperlipidemia, indicating that the arterial blood vessels were more rigid and less elastic in hyperlipidemia patients. These results are consistent with previous findings of high arteriosclerosis incidence and reduced hemodynamic functions in hyperlipidemia patients [14].

Our results also showed that h_1, h_3, and h_4 were relatively high in diabetic patients with hypertension compared with patients without hypertension, which can be explained by elevated arterial pressure because of high systolic and diastolic blood pressure in hypertension patients.

The prediction accuracy for diabetes is 96.35% (SVM-Poly). The algorithm can be used in mobile devices to conveniently and reliably diagnose diabetes. Predicting hypertension and hyperlipidemia among diabetic patients has low accuracy, 59.44% (LDA) and 69.77% (random forests), respectively. This may be explained by 2 reasons. First, the sample size for classifying hypertension and hyperlipidemia is relatively small (n=417). Second, hyperglycemia, hypertension, and hyperlipidemia can aggravate the microvascular and macrovascular lesions in diabetic patients [15], such as the decrease in vascular elasticity and the increase of intravascular pressure. Therefore, the differences between diabetic patients with hypertension/hyperlipidemia and patients without hypertension/hyperlipidemia may not be obvious. We plan to collect more samples, and we hope the accuracy of the models for identifying diabetic patients with hypertension/hyperlipidemia may increase with the larger sample size.

Strengths and Limitations

With the growing maturity of sensors, chips, mobile internet, and other technologies, people's health awareness is increasing and the demand for health services is greatly improved. The demand for wearable technology is becoming more and more high. Wearable and portable devices have shown great potential in the field of medical health. Pulse-taking diagnosis has several advantages such as its noninvasive nature and convenience. In our research, the pulse wave information acquisition element is a pressure sensor, which is the most commonly used device. Although there is another new type of sensor called an *ultrasonic sensor* used to detect pulse, the traditional pressure sensor is more cost-effective. In recent years, mobile phone apps are becoming more and more useful for the self-management of diabetes [16-18]. The wristband acquisition terminal we used also provides the possibility of integrating pulse parameter collection to mobile devices (such as sport bands and watches) and data analysis through mobile apps in the future.

The results from our study have confirmed the findings from previous studies [10-15]. Our methods showed a very high diagnostic accuracy for type 2 diabetes, whereas the diagnosis of hypertension and hyperlipidemia is not yet reliable. If we continuously improve the accuracy of using objective pulse information to detect type 2 diabetes in future research, the noninvasive and convenient pulse-taking diagnosis technique may have the potential to become a low-cost and accurate method to help people monitor the occurrence and development of type 2 diabetes and its complications more conveniently in daily life. It is important to note that in our research, when we used this device to collect pulse wave information, the subjects were asked to keep quiet and relaxed to measure the pulsation of the radial artery of their left hands under the appropriate temperature (as mentioned in the section Collecting Methods for the Objective Parameters of Traditional Chinese Medicine Diagnosis). Therefore any mobile device that uses our model will rely on the user being able to replicate the same condition. In the future, we hope to improve the convenience and stability of the device so that people can use it on both the left and right hands when moving. We are currently developing such devices and programs for the improvement of diabetes monitoring and remote diagnosis.

Acknowledgments

This work was supported by the Scientific Research Project of Shanghai Municipal Commission of Health and Family Planning (No. 20164Y0279), School Sports Scientific Research Project of Shanghai Municipal Education Commission (No. HJTY-2018-D12), and National Natural Science Foundation of China (No. 81673880).

Authors' Contributions

YH wrote the whole manuscript text, which was revised by FC, MP, HR, DP, and JY. YW and FC helped in the ideas of study. MP and DP helped in analyzing the data and plotting the results. YF, YF, XS, and HY helped in the sample collection. All authors (YH, FC, MP, HR, DP, YF, YF, JY, XS, HY, and YW) reviewed the manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2016. 10 facts about diabetes URL: http://www.who.int/campaigns/world-health-day/2016/en/ [accessed 2019-03-30] [WebCite Cache ID 77FroF8KW]
2. World Health Organization. 2016. Diabetes country profiles-China URL: http://www.who.int/diabetes/country-profiles/chn_en.pdf?ua=1 [accessed 2019-03-30] [WebCite Cache ID 77Fs2au0k]
3. Xu M. Diabetes mellitus (Second Edition). Shanghai: Shanghai science and Technology Press; 2010:369.
4. Wang Y, Xu J, Guo R, Xu C, Hao Y, Chen C, et al. Therapeutic effect in patients with coronary heart disease based on information analysis from traditional Chinese medicine four diagnostic methods. J Tradit Chin Med 2014 Feb;34(1):34-41 [FREE Full text] [Medline: 25102688]
5. Zhang Z, Zhang Y, Yao L, Song H, Kos A. A sensor-based wrist pulse signal processing and lung cancer recognition. J Biomed Inform 2018 Dec;79:107-116 [FREE Full text] [doi: 10.1016/j.jbi.2018.01.009] [Medline: 29428411]
6. Zhang M, Bai Y, Ye P, Luo L, Xiao W, Wu H, et al. Type 2 diabetes is associated with increased pulse wave velocity measured at different sites of the arterial system but not augmentation index in a Chinese population. Clin Cardiol 2011 Oct;34(10):622-627 [FREE Full text] [doi: 10.1002/clc.20956] [Medline: 21994082]
7. Chang C, Liao K, Chang Y, Wang S, Chen Y, Wang G. The effect of radial pulse spectrum on the risk of major adverse cardiovascular events in patients with type 2 diabetes. J Diabetes Complications 2019 Feb;33(2):160-164. [doi: 10.1016/j.jdiacomp.2018.10.009] [Medline: 30381150]
8. American Diabetes Association. Standards of medical care in diabetes--2014. Diabetes Care 2014 Jan;37(Suppl 1):S14-S80. [doi: 10.2337/dc14-S014] [Medline: 24357209]
9. Fei Z. Study on pulse-taking in China. Shanghai: Shanghai Institute of Traditional Chinese Medicine Press; 1991:101-105.

10. Roussel R, Steg PG, Mohammedi K, Marre M, Potier L. Prevention of cardiovascular disease through reduction of glycaemic exposure in type 2 diabetes: a perspective on glucose-lowering interventions. Diabetes Obes Metab 2018 Dec;20(2):238-244. [doi: 10.1111/dom.13033] [Medline: 28597588]
11. Khaleeli E, Peters S, Bobrowsky K, Oudiz R, Ko J, Budoff M. Diabetes and the associated incidence of subclinical atherosclerosis and coronary artery disease: implications for management. Am Heart J 2001 Apr;141(4):637-644. [doi: 10.1067/mhj.2001.113224] [Medline: 11275932]
12. Harper E, Forde H, Davenport C, Rochfort KD, Smith D, Cummins PM. Vascular calcification in type-2 diabetes and cardiovascular disease: integrative roles for OPG, RANKL and TRAIL. Vascul Pharmacol 2016 Dec;82:30-40. [doi: 10.1016/j.vph.2016.02.003] [Medline: 26924459]
13. Li T, Ni L, Liu X, Wang Z, Liu C. High glucose induces the expression of osteopontin in blood vessels in vitro and in vivo. Biochem Biophys Res Commun 2016 Dec 11;480(2):201-207. [doi: 10.1016/j.bbrc.2016.10.027] [Medline: 27743892]
14. Sharma A, Tate M, Mathew G, Vince JE, Ritchie RH, de Haan JB. Oxidative stress and NLRP3-inflammasome activity as significant drivers of diabetic cardiovascular complications: therapeutic implications. Front Physiol 2018 Feb 20;9:114 [FREE Full text] [doi: 10.3389/fphys.2018.00114] [Medline: 29515457]
15. Khunti K, Kosiborod M, Ray K. Legacy benefits of blood glucose, blood pressure and lipid control in individuals with diabetes and cardiovascular disease: Time to overcome multifactorial therapeutic inertia? Diabetes Obes Metab 2018 Dec;20(6):1337-1341. [doi: 10.1111/dom.13243] [Medline: 29405543]
16. Desveaux L, Shaw J, Saragosa M, Soobiah C, Marani H, Hensel J, et al. A mobile app to improve self-management of individuals with type 2 diabetes: qualitative realist evaluation. J Med Internet Res 2018 Mar 16;20(3):e81. [doi: 10.2196/jmir.8712]
17. 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]
18. Hou C, Xu Q, Diao S, Hewitt J, Li J, Carter B. Mobile phone applications and self-management of diabetes: a systematic review with meta-analysis, meta-regression of 21 randomized trials and GRADE. Diabetes Obes Metab 2018 Dec;20(8):2009-2013. [doi: 10.1111/dom.13307] [Medline: 29582538]

Abbreviations

ANOVA: analysis of variance
h_1: height of dominant wave
h_3: height of tidal wave
h_4: height of dicrotic notch
h_5: height of dicrotic wave
LDA: linear discriminant analysis
SVM: support vector machine
TCM: traditional Chinese medicine
t_1: time distance between the start point of pulse wave and dominant wave
t_4: time distance between the start point of pulse wave and dicrotic notch
t_5: time distance between dicrotic notch and the end point of pulse wave
USB: universal serial bus
W: width of percussion wave in its one-third height position

Edited by G Eysenbach; submitted 19.08.18; peer-reviewed by J Goris, D López López; comments to author 14.10.18; revised version received 08.03.19; accepted 24.03.19; published 23.04.19.

Please cite as:
Hao Y, Cheng F, Pham M, Rein H, Patel D, Fang Y, Feng Y, Yan J, Song X, Yan H, Wang Y
A Noninvasive, Economical, and Instant-Result Method to Diagnose and Monitor Type 2 Diabetes Using Pulse Wave: Case-Control Study
JMIR Mhealth Uhealth 2019;7(4):e11959
URL: http://mhealth.jmir.org/2019/4/e11959/
doi:10.2196/11959
PMID:31012863

Original Paper

Effect of Serial Anthropometric Measurements and Motivational Text Messages on Weight Reduction Among Workers: Pilot Randomized Controlled Trial

Renee Chan[1], MBBS; Matthew Nguyen[1], MBBS; Rachel Smith[1], MBBS; Sarah Spencer[1], MBBS; Sabrina Winona Pit[1,2], PhD, MSc, Dip (OHS)

[1]School Of Medicine, University Centre for Rural Health, Western Sydney University, Lismore, Australia

[2]Rural Clinical School, University of Sydney, Lismore, Australia

Corresponding Author:
Sabrina Winona Pit, PhD, MSc, Dip (OHS)
School Of Medicine
University Centre for Rural Health
Western Sydney University
62 Uralba Street
Lismore, 2480
Australia
Phone: 61 266207570
Email: sabrina.pit@sydney.edu.au

Abstract

Background: Obesity is an endemic problem with significant health and financial consequences. Text messaging has been shown to be a simple and effective method of facilitating weight reduction. In addition, waist-to-hip ratio (WHR) has emerged as a significant anthropometric measure. However, few studies have examined the effect of serial anthropometric self-measurement combined with text messaging.

Objective: The primary aim of this study was to assess whether an 8-week program, consisting of weekly serial self-measurements of waist and hip circumference, combined with motivational text messages, could reduce WHR among Australian workers.

Methods: This was a community-based, participant-blinded, staggered-entry, parallel group study. Adult workers with access to mobile phones were eligible and recruited through an open access Web-based survey. Participants were randomly allocated to receive intervention or control messages for 8 weeks. Outcome data were self-assessed through a Web-based survey.

Results: A total of 60 participants were randomized with 30 participants each allocated to a control and an intervention group. There was no significant change in WHR (P=.43), and all secondary outcome measures did not differ between the intervention group and the control group at the end of the 8-week intervention. Both groups, however, showed a significant decrease in burnout over time (mean [SE]: pre 4.80 [0.39] vs post 3.36 [0.46]; P=.004). The intervention uptake followed a downward trend. Peak participant replies to weekly self-measurements were received in week 3 (14/23, 61%) and the least in week 8 (8/23, 35%). No harm was found to result from this study.

Conclusions: This study is an innovative pilot trial using text messaging and serial anthropometric measurements in weight management. No change was detected in WHRs in Australian workers over 8 weeks; therefore, it could not be concluded whether the intervention affected the primary outcome. However, these results should be interpreted in the context of limited sample size and decreasing intervention uptake over the course of the study. This pilot trial is useful for informing and contributing to the design of future studies and the growing body of literature on serial self-measurements combined with text messaging.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12616001496404; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=371696&isReview=true (Archived by WebCite at http://www.webcitation.org/73UkKFjSw)

(JMIR Mhealth Uhealth 2019;7(4):e11832) doi:10.2196/11832

KEYWORDS

text messages; obesity; waist-hip ratio; weight reduction programs; mHealth

Introduction

The Problem of Obesity

Obesity is an endemic problem worldwide with significant health consequences to the individual [1] and financial burden on the community. The economic cost of obesity in Australia was estimated at Aus $52.8 billion in 2008 alone, including productivity loss, costs to the health system, and impact on well-being [2]. In 2014-15, 28% of Australian adults were obese, showing an increase of 19% from 1995 [3]. Obesity is commonly measured by weight and body mass index (BMI); however, waist-to-hip ratio (WHR) is also increasingly being recognized as being an important measure of obesity [4].

Text Messaging in Weight Loss

Exercise is a key element in achieving weight loss goals; however, there are many influences on an individual's level of exercise, including perceptions of support [5]. Text messaging has proven to be a cheap, simple, and an effective support strategy in encouraging weight reduction [6-8]. Text messages used to promote health messages currently have limited evidence but show promising potential as an effective health promotion tool [9]. In 1 study, tailored text messages and multimedia messaging service with tips, suggestions, and positive reinforcement over 4 months led to an additional loss of 2 kg in the intervention compared with the control group that received monthly printed materials over a period of 4 months [10]. Moreover, there is some evidence that 1 text message a day is able to improve motivation toward weight loss behaviors without adding extra burden, but this requires further testing [9].

There is no consensus with respect to the most effective text message content. Interventions to increase physical activity and healthy eating vary widely. Despite this, there are few systematic reviews evaluating factors that influence the effectiveness of text message interventions. There is a taxonomy of behavioral change techniques created to improve the effectiveness of interventions aiming at increasing physical activity and healthy eating [11]. It was found in 2 systematic reviews using the taxonomy that interventions where participants engaged in self-monitoring were more effective in achieving goals of behavior change [11].

Potential harms of text messaging are generally limited but could potentially depend on the context and frequency of text messages. These might include, for example, perceptions of privacy invasion, and emotional trauma as a result of negative body image. This can be addressed by providing participant information before consent and access to services and resources designed to assist individuals in these issues.

Serial Body Measurements in Weight Loss

Some reviews found that regular self-weighing was associated with weight loss. Despite variations in the frequency and size of correlation, the association with weight loss was consistent [12-14]. It has been suggested that the frequency for self-weighing to achieve successful outcomes is weekly [15]. However, current evidence does not conclude what the ideal frequency for self-weighing is despite most studies evaluating daily or weekly self-weighing [7]. There are other body measurements that can be taken serially in weight loss programs including BMI, waist and hip measurements, and WHRs. However, there is a significant gap in the literature regarding the effectiveness of these other measurements. Few studies conducted serial anthropometric measurements, and although text messaging was identified as an effective intervention [6,8,16], few combined it with anthropometric measurements. Research has shown the importance of WHRs in relation to obesity and health risks. For example, 1 study demonstrated that WHR is significantly associated with the risk of incident cardiovascular disease events and is a simple measure of abdominal obesity [17]. Similarly, another study found that WHR is associated with a higher risk of major adverse cardiovascular events among females, but not in males, with established coronary artery disease [18]. To our knowledge, there are no studies that have used WHR as a primary outcome measure in motivational text messaging studies. This study will therefore use WHR as an outcome measure.

The primary aim of this pilot study was to assess whether an 8-week program, consisting of weekly serial self-measurements of waist and hip circumference, combined with motivational text messages, could reduce WHR among Australian workers. Secondary aims were to examine the effects of the program on weight loss, exercise, eating behavior, and work-related well-being measures.

Methods

Trial Design

This pilot study evaluated the impact of an 8-week program consisting of motivational text messages and serial anthropomorphic measurements on reducing the WHR, other anthropometric measurements, health behaviors, and occupational health-related outcomes. It was a community-based, participant-blinded, staggered entry, parallel group study with balanced randomization (1:1) conducted in Australia using convenience sampling.

Participants and Recruitment

Eligibility criteria were being above 18 years or older, being employed, and having access to a mobile phone in Australia. Exclusion criteria were people receiving weight-altering medications or participating in other weight loss programs. Participants were provided with a participant information sheet providing them with the length of the study, purpose, and affiliations of the study before enrollment into the study. Ethics approval was received from the Western Sydney Human Research Ethics Committee H11327.

Study recruitment ran from October 2016 to January 2017 via a Facebook page, emails to the researchers' contacts, flyers to public notice boards and local businesses in the Northern Rivers, New South Wales, and information in councils' newsletters in the Northern Rivers and Western Sydney region. Flyers and emails contained a link to an open a Web-based survey for participant enrollment and baseline data collection. The initial contact with the potential participants was thus made via the internet. Institutional affiliation to Western Sydney University was indicated in our materials. The recruitment materials

advertised the study as a weight loss program, but there were no incentives offered to participants and participation was voluntary.

The survey was pretested on 18 volunteers to assess usability and technical functionality. Each participant completed identical baseline surveys, which consisted of 29 items over 7 pages. Items were not randomized or alternated. Adaptive questioning was not used. Only submitted surveys were considered as participant consent to the study. Respondents were able to review and change their answers while completing the survey but not after submission. No participant submitted more than 1 survey. No identifying information was linked to the data. ID numbers were used to analyze the data on password-protected computers.

Intervention Group

Participants were randomly assigned to receive intervention or control messages for 8 weeks using a Web-based short message service SMS company. The intervention was a composite of motivational and self-monitoring messages. The 25 motivational messages sent every second day were based on promotion messages from another text-based study regarding nutrition [16], exercise, and monitoring (Multimedia Appendix 1). The self-monitoring messages were identical weekly requests for waist and hip circumference, aimed at providing self-feedback on progress (Multimedia Appendix 1) [15]. All messages were sent at 12 pm.

Participants were able to opt out of the intervention anytime by texting *STOP*.

Control Group

Fortnightly control messages were sent with health information from the national guidelines on physical activity, diet, and nutrition [19]. Control group participants did not require to report their anthropometric measurements weekly. They only were requested to provide their WHR measurements at baseline and 8week follow-up.

Outcomes

Participants were followed up with a final Web-based survey 8 weeks after the start of their intervention. Final surveys were identical for all participants except for those completing it after Jan 23, 2017, when 2 open-ended questions were added to gather a more in-depth understanding about the pilot study. The final survey consisted of 23 items over 5 pages. Reminder emails and text messages were sent a week and a fortnight after completion. Two sets of participants were asked to complete the final survey outside of this protocol as a result of researcher error. This affected 10 participants; 6 participants were able to complete the survey 3 weeks after program completion instead of 2 weeks, whereas 4 participants were invited to complete the final survey 2 weeks before the completion of their intervention.

Outcome data was self-assessed and collected on the Web using SurveyMonkey at the beginning and end of the study. Questions were mainly derived from existing scales. The invitation to the final survey was sent in the last text message and email.

The primary outcome was WHR change from baseline to 8 weeks collected by participants measuring their waist and hip circumference in centimeters with help from instructive pictures and videos. Secondary outcomes were changes in anthropometric measurements, health behaviors, and occupational health-related outcomes. Self-reported health was measured with the global health question from the Short Form-36 (*In general, would you say your health is;* rated on a 5-point Likert scale ranging from *poor* to *excellent*). This item has consistently been found to possess strong psychometric properties compared with validated multi-item measures [20]. Parts of the widely used and validated Work Ability Index were used to measure occupational health [21]. Specifically, the 3-item version of the Work Ability Index, extensively validated by Mykletun and Furunes [22] was used and asked participants to self-report their current work ability; (1) on a scale from 0-10 compared with lifetime best, (2) in relation to physical demands, and (3) in relation to mental demands. Work ability at its lifetime best was measured through asking: *Work ability is an indication of how well your health, skills and experience match your current job demands. Assume that your work ability at its lifetime best has a value of 10 points. How many points would you give your current work ability? (0 means that you currently cannot work at all)*. The physical demands of work ability were measured by asking: *How do you rate your current work ability with respect to the physical demands of your work.* Answer categories ranged from *very good, rather good, moderate, rather poor, and very poor.* The same question was asked for the mental demands of the job. The commonly used and validated 9-item Emotional Exhaustion subscale from the Maslach Burnout Inventory (Human Services Survey) was used to keep the survey short and this often being regarded to be the core component of the Maslach Burnout Inventory [23]. The reliability coefficients for emotional exhaustion were 0.89 (frequency) and 0.86 (intensity) [24]. The Single Item Burnout scale was measured through the question*: On a scale from 0 to 10, how would you rate your current level of burnout?* where 1 represents *Not at all burnt out* and 10 represents *Extremely burnt out.* A previous study has demonstrated the validity of the Single Item Burnout scale. The item was highly and positively correlated with MBI-EE scores (r=0.8, *P*<.001) and was significantly associated with various outcome measures [25]. Productivity was measured through a self-developed 1-item question asking: *On a scale of 1-10, could you rate, how productive you were at work in the last week?* where 1 represents *Not at all productive* and 10 represents *Extremely productive.* Questions from a large cohort study, the 45 and Up Study, were used to measure healthy eating behavior, exercise, and total sitting hours per day [26].

Process measures were also assessed to measure levels of engagement and intervention uptake. These were as follows:

- The number of replies the intervention group made to the weekly request messages for self-measurement
- The time between finishing the study and completing the final survey

In addition, to elicit feedback on the program, the following 2 open-ended questions were added:

1. Do you feel that taking part in the study made you live or feel healthier? Can you explain?
2. Do you have any further comments on the program?

Sample Size

To detect a difference in our participants' WHR of 0.03 with a 5% significance level and assuming an SD of 0.064, 72 participants per group were required to provide the study with a power of 80%. However, our study included 30 participants at baseline per group because of unexpected difficulties in recruitment.

Randomization

Participants were randomized in blocks of 10 to intervention or control through a computer-generated random number list on Excel created by a researcher (SWP) not involved in allocation. All other researchers were involved in allocation. Although there was no allocation concealment, enrollment of the participants occurred automatically during the baseline survey with no direct contact from the researchers.

Participants were allocated an ID number based on the order in which they completed the baseline survey. They were allocated to control or intervention on the Sunday after enrollment and started the intervention on the Monday in either the intervention or control group.

Blinding

Data analysts and researchers undertaking randomization were not blinded during the trial; however, there was no direct contact between participants and researchers throughout the entirety of the trial, and ID numbers were used for participant anonymity during analysis. Participants were blinded to group allocation by concealing the frequency and content of text messages.

Statistical Analysis

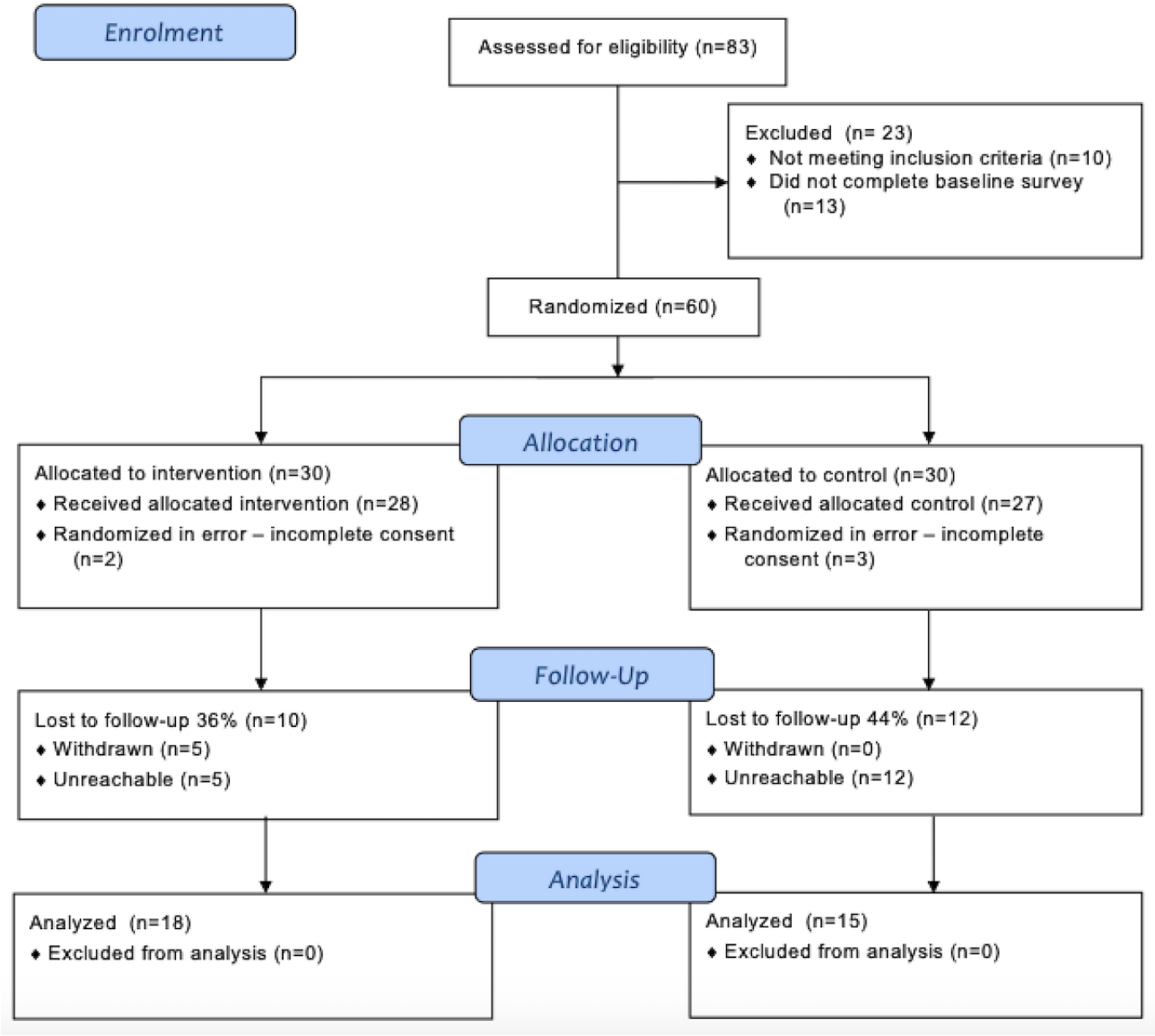
Baseline descriptive analyses examined variable distribution for sample characteristics. Continuous variables were presented as mean and SDs, and nonparametric data as median and interquartile ranges. Binomial and categorical variables were reported as proportions. Statistical analyses were performed on SPSS Version 22.0 (SPSS Inc, Chicago, IL, USA) and SAS version 9.4 (SAS Institute, Cary, NC, USA). *P* values less than .05 were considered significant. A mixed-model repeated measures analysis was performed to compare the effect of the intervention on the primary outcome measure, WHR, and other continuous secondary outcome measures at 8-week follow-up, compared with control. The interaction effect between *group* and *time* indicates whether the intervention successfully reduced WHR overtime compared with the control group. Compound symmetry was the specified covariance structure. This means that all the variances and covariances are equal for all participants. Healthy eating behaviors, exercise, and sitting hours per day were skewed to the right. These variables were analyzed to detect the presence of a negative or positive change from baseline to 8-week follow-up. The intervention effect was then compared by using the Fisher exact test because of the small numbers in each category. Atypical data such as impossible body measurement values were considered as missing values, and these numbers were excluded from analysis.

Results

Recruitment

This study recruited members between October 2016 and January 2017. There were 9 weekly sets of participants who entered the trial throughout this recruitment period, with the final set of participants completing the intervention in March 2017. Participant numbers each week varied from 1 to 18. The trial was ended as per the scheduled date of closure. The participant flow is summarized in Figure 1.

Figure 1. Participant flowchart.



Baseline

A total of 55 participants entered the trial, 33 completed the trial and were included in the final analyses. Baseline characteristics appear to vary between the intervention and control group (Tables 1 and 2). The intervention group was older (median age 45 vs 33 years) and overall less healthy than the control group, with higher mean hip circumference (106.1 vs 95.7 cm), mean weight (76.9 vs 69.9 kg), and proportion of overweight and obese BMI 52% (11/21) versus 39% (9/23); and lower moderate physical activity levels (median: 30 vs 60 min/week).

Numbers Analyzed

Of the 60 participants randomized, 5 participants were randomized in error (3 intervention and 2 control group participants), 5 participants withdrew during the program and 22 participants were lost to follow-up (ee Figure 1). Of these participants lost to follow-up, 5 provided partially completed final surveys at 8 weeks. Data from partially completed surveys were used in the analysis where possible. These participants completed the first page of the final survey only. All 5 participants did not enter weight and waist or hip circumference measurements. Therefore, only the number of measurements for sitting time in the analyses was higher than the other variables. Completed final surveys were provided by 33 participants and used in primary outcome analysis.

Table 1. Baseline participant characteristics.

Characteristics	Intervention (n=23)	Control (n=27)	Total (n=50)
Sociodemographics			
Age (years), median (Q1, Q3)	45 (30, 53)	33 (25, 46)	38 (27,51)
Female, n (%)	19 (83)	23 (85)	42 (84)
Hours spent doing paid work, median (Q1, Q3)	32 (20, 40)	35 (30, 40)	33.5 (24, 40)
Relationship status, n (%)			
In a relationship or married or engaged or de facto[a]	11 (50)	14 (52)	25 (51)
Single or divorced or widowed or separated[a]	11 (50)	13 (48)	24 (49)
Anthropometric measures			
Height, cm[b], mean (SD)	164.9 (8.3)	166.0 (9.2)	165.5 (8.7)
Weight, kg[b], mean (SD)	76.9 (20.4)	69.9 (21.4)	73.2 (21.0)
Body mass index[b], kg/m^2, mean (SD)	28.2 (6.7)	25.1 (6.1)	26.6 (6.5)
Overweight status[b], n (%)	11 (52)	9 (39)	20 (45.5)
Waist circumference, cm[c], mean (SD)	89.4 (16.1)	80.4 (17.2)	85 (17.0)
Hip circumference, cm[c], mean (SD)	106.1 (13.4)	95.7 (13.5)	101.0 (14.3)
Waist-to-hip ratio[c], mean (SD)	0.84 (0.08)	0.84 (0.10)	0.83 (0.09)
Health behaviors			
Number of serves of cooked vegetables, median (Q1, Q3)	2 (1, 3)	1 (1, 2)	2 (1, 3)
Number of serves of raw vegetables, median (Q1, Q3)	1 (1, 1)	2 (1, 2)	1 (1, 2)
Met vegetables requirement [27], n (%)	4 (18)	5 (19)	9 (18)
Number of serves of fruit, median (Q1, Q3)	2 (1, 2)	1 (1, 2)	1 (1, 2)
Number of glasses of fruit juice[d], median (Q1, Q3)	0 (0, 0)	0 (0, 0)	0 (0, 0)
Met fruit requirement [27], n (%)	12 (55)	13 (48)	25 (51)
Hours spent sitting per day, mean (SD)[e]	8 (7, 10)	8 (6, 12)	8 (7, 10)
Mild physical activity per week (min)[f], median (Q1, Q3)	60 (14, 240)	60 (10, 120)	60 (10, 180)
Moderate physical activity per/week (min)[g], median (Q1, Q3)	60 (10, 180)	30 (20, 60)	60 (10, 120)
Vigorous physical activity per/week (min)[h], median (Q1, Q3)	2 (0, 60)	20 (0, 60)	7 (0, 60)
Proportion meeting physical activity guidelines [27], n (%)	12 (55)	14 (52)	26 (53)
Accountability partner, n (%)	6 (27)	7 (26)	13 (27)

[a]n=49 (control: 27 vs intervention: 22).
[b]n=44 (intervention: 21 vs control: 23).
[c]n=39 (intervention: 20 vs control: 19).
[d]n=47 (intervention: 20 vs control: 27).
[e]n=50.
[f]n=47 (intervention: 21 vs control: 26).
[g]n=46 (intervention: 21 vs control: 25).
[h]n=48 (intervention: 21 vs control: 27).

Table 2. Baseline occupational health-related outcomes.

Outcomes	Intervention (n=23)	Control (n=27)	Total (n=50)
General health, n (%)[a]			
Excellent or very good	6 (27)	8 (31)	14 (29)
Good	10 (46)	13 (50)	23 (48)
Fair or poor	6 (27)	5 (19)	11 (23)
Work ability lifetime best (0 to 10), median (Q1, Q3)	8 (7, 9)	8 (7, 9)	8 (7, 9)
Work ability mental demands, n (%)			
Very good	3 (14)	7 (26)	10 (20)
Rather good	12 (55)	14 (52)	26 (53)
Moderate or rather poor or very poor	7 (32)	6 (22)	13 (27)
Work ability physical demands, n (%)			
Very good, n (%)	9 (41)	14 (52)	23 (47)
Rather good, n (%)	10 (46)	10 (37)	20 (41)
Moderate or rather poor or very poor, n (%)	3 (14)	3 (11)	6 (12)
Productivity (0 to 10), median (Q1, Q3)	7 (6, 8)	8 (6, 9)	7 (6, 8)
Burnout Score (0 to 10), mean (SD)	4.8 (2.6)	4.7 (2.6)	4.8 (2. 7)
Maslach burnout inventory—emotional exhaustion, mean (SD)	31.0 (13.6)	30.4 (11.5)	30.7(12.4)

[a]n=48 (intervention: 22 vs control: 26).

Table 3. Comparison between control group and intervention group anthropometric measures and work-related health.

Outcomes	Intervention		Control		Time effect	Group × time effect	Group effect
	Baseline, mean (SE)	At 8 weeks, mean (SE)	Baseline, mean (SE)	At 8 weeks, mean (SE)	*P* value	*P* value	*P* value
Primary outcome							
Waist-to-hip ratio	0.84 (0.02)	0.86(0.02)	0.83(0.02)	0.82(0.03)	.68	.30	.43
Anthropometric measures							
Body mass index (kg/m^2)	28.25 (1.34)	28.14 (1.34)	25.11 (1.29)	25.26 (1.29)	.85	.22	.12
Waist circumference (cm)	89.95 (3.73)	90.28 (3.83)	80.76 (3.88)	77.00 (4.26)	.45	.36	.04
Hip circumference (cm)	107.08 (3.24)	104.12 (3.34)	96.39 (3.36)	93.31 (3.76)	.17	.97	.02
Well-being and work-related health							
Burnout score (0 to 10)	4.81 (0.57)	3.11 (0.63)	4.77 (0.52)	3.60 (0.67)	.004	.57	.75
Maslach burnout inventory—emotional exhaustion	25.50 (2.72)	30.82 (2.74)	30.37 (2.37)	31.00 (2.62)	.08	.11	.39
Work ability (0 to 10)	7.91 (0.36)	7.65 (0.45)	8.11 (0.32)	8.31 (0.41)	.91	.41	.35
Productivity (0 to 10)	7.18 (0.39)	6.65 (0.45)	7.25 (0.35)	7.92 (0.47)	.86	.15	.14

Post Results

The changes in anthropometric measures and work-related health are summarized in Table 3. The group by time effect was not significant indicating the pattern over time did not differ by group for the primary outcome (Table 3, group by time). The main effect of time was also not significant for all variables, indicating no change over time for either group, except for the single item burnout where there was a significant decrease in burnout over time (mean [SE]: pre 4.80 [0.39] vs post 3.36 [0.46]; *P*=.004). Only waist and hip circumferences had a significant group effect, indicating that these measurements averaged over time were significantly greater in the intervention group.

Table 4. Comparison between control group and intervention group for health behaviors, representing the number (%) of participants that reported changes in health behaviors from baseline to 8-week follow-up.

Health behaviors	Less, n (%)	Same, n (%)	More, n (%)	*P* value
Number of serves of cooked vegetables[a]				
Intervention	4 (24)	10 (59)	3 (18)	.26
Control	2 (13)	6 (40)	7 (47)	—[a]
Number of serves of raw vegetables[a]				
Intervention	0 (0)	7 (39)	11 (61)	.08
Control	3 (20)	7 (47)	5 (33)	—
Number of serves of fruit				
Intervention	2 (11)	10 (56)	6 (33)	.41
Control	0 (0)	7 (47)	8 (53)	—
Number of glasses of fruit juice				
Intervention	1 (6)	14 (88)	1 (6)	.79
Control	1 (7)	12 (80)	2 (13)	—
Total minutes of moderate or vigorous exercise				
Intervention	6 (33)	3 (17)	9 (50)	.38
Control	6 (40)	0 (0)	9 (60)	—
Total hours sitting per day				
Intervention	6 (33)	3 (17)	9 (50)	.63
Control	7 (40)	2 (0)	8 (60)	—

[a]Not applicable.

No statistically significant differences were detected for health behaviors between the 2 groups (Table 4). However, it is interesting to note there was an increase, from baseline to 8-week follow-up, in the number of serves of raw vegetables consumed by 11 out of 18 participants (11/18, 61%) in the intervention group with no decrease of serves, whereas 5 out of 15 participants (5/15, 33%) in the control group increased their servings of raw vegetables and 3 out of 15 (3/15, 20%) decreased. No participant had significantly poorer health behaviors at the end of the study or other harms found as an outcome of this study.

Process Measures

The pattern of intervention uptake is shown in Table 1 in Multimedia Appendix 2. Intervention group participants' responses to weekly self-measurements requests followed an overall downward trend with the peak percentage of 61% (14/23) in week 3 and the lowest percentage of 35% (8/23) in week 8. It was also found that participants in the control group took longer to complete the final survey compared with the intervention group after the 8 weeks with a mean of 8.80 days compared with 3.17 days (Table 2, Multimedia Appendix 2).

Participant Feedback

A total of 15 people responded to the qualitative question (intervention=8 vs control=7).

Overall Findings

A total of 11 participants felt that the study did not make them live or feel healthier and had *no impact*. Feedback included:

... text came through whilst at work so didn't influence exercise pattern, also over Xmas period so exercised less and ate more and *I think more frequent texts, even daily, with a list of exercises for the day or positive affirmations for regular exercising and why you should do the days 30 min exercise plan (for example).*

Intervention Group

In total, 5 of 8 participants who responded commented that the study did not make them live or feel healthier. They commented:

> *Unfortunately no. Although I found the texts very informative, I didn't make an effort to put them into practice*
>
> *No, not really. The texts were too easy for me to ignore, or forget about. I'm not in the habit of checking my phone regularly*

In total, 3 participants thought the intervention had a positive impact, for example:

> *Yes as I thought about my health more often*
>
> *It has been positive to have daily texts and reminders even if I took little action from them*

Control Group

In total, 6 of 7 responses were negative. Some felt that:

> *...it would have been better to have more frequent texts*
>
> *I don't know think one quick fact every couple of weeks can change a whole attitude. For me, it requires more regular reminders and having someone like an accountability partner who can follow you up often works best*

In total, 1 had a positive response to the study:

> *yes it was a good reminder to eat healthy and exercise*

Discussion

Principal Findings

No significant changes over time were found between the WHR of the intervention and control group over time. Single item burnout showed a significant decrease over time. No weight gain or other anthropometric measurements, health behaviors, and occupational health measures showed significant changes over time.

A number of mechanisms might account for the results of this study. First, the study ran over the holiday season in Australia, which is traditionally a time in which individuals gain weight [28]. One study on diabetes prevention advised dietitians that the goal for patients in this period should be weight stabilization [29]. No weight gain or WHR change was found at the 8-week follow-up, suggesting that weight stabilization might have occurred.

The decrease in burnout independent of allocation might similarly be a reflection of upcoming major holidays. There is little evidence to suggest that upcoming holidays decrease burnout; however, a study found that if an individual had a trip planned for their holiday time, they were more likely to report being happier [30].

The diversity of interventions makes direct numerical comparisons difficult in technological weight loss studies. However, text messaging use for reminders, such as those we used to induce self-monitoring or to promote behavior change, does have an evidence basis. One systematic review found that although a relatively new area of research, text messaging as a lifestyle intervention was promising in its feasibility and acceptability [9]. However, of the 10 studies reviewed measuring weight or BMI as an outcome, only 5 showed a statistical and clinical difference after the intervention. The sole study in the review that used WHR found no difference when using text messages to remind participants of physical exercise goals they had set [31]. In reviewing the effect of text messaging on physical activity, the systematic review found that 3 of the 6 trials that used physical activity as an outcome showed a statistically significant increase in the frequency or duration [9]. With regard to diet, 3 out of the 4 studies with dietary outcomes showed a statistically significant improvement using text messaging. More recently, a pilot study examining the use of text messaging to improve health among African American women, an at-risk group for obesity, found it to be effective [32]. Self-monitoring as an intervention has been shown to be effective with 1 study finding that self-weighing was associated with weight loss [12], a result we did not establish with self-measuring WHR.

Burnout reduction using text messages among workers has, to the researchers' knowledge, never been investigated before in a randomized controlled trial (RCT) setting. This suggests that this maybe an area of further research to further explore its impact given that it significantly reduced overtime in both groups. This might have been a type 2 error though. Other intervention studies have tested the efficacy of guided Web-based and mobile-based stress management training for employees and found that emotional exhaustion was reduced in the intervention group [33]. Similarly, a 2018 RCT evaluated the efficacy of an internet-based, app-supported stress management intervention for college students and also found a reduction in emotional exhaustion among intervention group participants [34]. Along with this, most studies excluded participants with normal BMIs. One trial that included these participants, similar to our study, found that though text messages increased physical exercise significantly, there was no impact on BMI [35]. In addition, another study found that their intervention had significantly less effect on those with lower BMIs when compared with those in the obese range [36]. Finally, a 2017 study among obese adolescents that used a mobile phone–based intervention consisting of 3 parts—use of the Fitbit Flex, delivery of an online educational program, and biweekly text messages during the maintenance phase—also found that the program significantly improved BMI, physical activity days per week, and servings of fruits and vegetables per day [37]. This could offer further explanation for our nonsignificant findings.

Limitations

This study had several limitations which must be considered when reviewing the results and in the development of future research. The restricted recruitment time limited the number of participants and contributed to a low power. Participant enrollment in the study was also limited by access to tools such as a tape measure. However, we compensated for this by the provision of videos explaining how to use string and a ruler to measure WHR. Another limitation of this study was the differing baseline characteristics between groups. The intervention group had a higher proportion of overweight and obese BMIs than the control group (52% vs 39%), a higher weight (76.9 kg vs 69.9 kg), and a wider hip circumference (106.1 cm vs 95.7 cm) in baseline characteristics. Furthermore, there was a significant difference between the waist and hip measurements averaged over time for the 2 groups, with the intervention group having significantly greater waist and hip measurements (P=.04 and P=.02, respectively).

Although our lost to follow-up rates were relatively high 40% (22/55), a review focusing on Web-based weight loss interventions showed most had an attrition rate that was higher [38]. Another study showed that 55% was a usual rate in obesity trials [39], which brings our lost to follow-up into perspective**.** However, there was a difference in the rates of participants who completed the intervention but not the final survey. There was

a lower proportion in the intervention group (n=5 [18%]) than control (n=12 [44%]). These differing unreachable rates could be because of the different frequencies of messages between the groups and indicates that fortnightly contact was not enough to maintain engagement in the control group. This was confirmed in the feedback of the control group forgetting they were in the study. It is interesting to note, however, that all the participants who withdrew from the study were in the intervention group, despite their high levels of contact. This might be due to the fact that some participants found the frequency of text messages intrusive, though this was not reflected in the feedback.

Another reason for the unequal attrition rates was suggested by 1 study that examined the reasons behind dropout [40]. They found that lower initial weight loss was associated with attrition [40]. Therefore, the lost to follow-up rates might have been higher in the control group because of the poor efficacy of the control text messages and lack of weight loss. This might provide an explanation for the results, as those in the control group who achieved weight reduction were more likely to complete the final survey and be included in the analysis.

All partial completers of the final survey (n=5) dropped out when asked to provide their weight and waist and hip circumference. It might be useful for future studies to consider placing questions known to lead to low response rates at the end of the survey. However, in our case, it was the primary outcome measure, so in future trials other methods of collecting weight and WHR data might need to be considered to ensure valid and complete data are collected.

Process measures to test engagement in the study were used to assist with the development of further research into this area. Follow-up response times were a part of these, and we found the intervention group had a shorter response time than the control group (mean days [SD]: 3.17 [3.50] vs 8.80 [6.27]). This presents an important aspect to consider when running similar trials in the future as it suggests that altering the content rather than the frequency of the messages between the groups might be a more effective option.

The other process measure of this study was the replies from the intervention group to the request messages. Serial self-measurement was a key intervention in this study, and strong uptake would be needed to measure its efficacy. However, replies to these messages were shown to decrease over time, revealing a decreasing level of engagement. Reasons for this could be similar to reasons for lost to follow-up events as discussed previously. This decreasing intervention uptake might have contributed to our negative findings and should be considered when interpreting results.

Some aspects of this study limit its application to the wider Australian context. First, 84% of the population was female, which, although a common problem with many weight loss studies [41], restricts generalizability if the target population is dissimilar. However, 1 systematic review found that lifestyle interventions, like ours, are effective in both men and women [42].

This study also had a limited, primarily young, age range. This might be because of the nature of our recruitment via social media and by restricting this study population to workers. Younger participants might also feel more comfortable participating in a study involving a relatively new aspect of technology. Another study limitation is selfreported measurements. We included several strategies for accuracy for the primary outcome measure. First, before starting the survey, we advised people we would ask them to measure their waist and hip and asked them to be in a comfortable place to measure themselves. Second, we showed them a picture and a video during the Web-based survey on how to measure hip and waist circumference to assist people in completing their measurements. Nonetheless, it is likely that some people will have estimated their hip and waist circumference, which might have biased the results. Secondary outcome measures were mainly based on validated scales for work ability [21], emotional exhaustion [24], single item burnout [25], or questions derived from a large cohort study for sitting behavior, healthy eating, and exercise [26].

Conclusions

This study is an innovative pilot trial using text messaging and serial self-measurement in weight management. The results did not detect a change in WHR ratio in Australian workers over 8 weeks. However, these results should be interpreted in the context of limited sample size and decreasing intervention uptake over the course of the study. We are unable to conclude this intervention is not effective. A larger sample would be necessary to see if the combination of these interventions is effective. The findings around study design and participant interaction with the interventions are useful for informing and contributing to the design of future studies and the growing body of literature on serial self-measurements combined with text messaging.

Acknowledgments

This study was self-funded. The research team would like to thank Western Sydney University and the University Centre for Rural Health for their practical support. The authors also thank all the study participants.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Text messages.

[DOCX File, 30KB - mhealth_v7i4e11832_app1.docx]

Multimedia Appendix 2

Process measures.

[DOCX File, 47KB - mhealth_v7i4e11832_app2.docx]

Multimedia Appendix 3

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i4e11832_fig.pdf]

References

1. World Health Organization. 2018. Obesity and overweight URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2019-03-17] [WebCite Cache ID 76vWqPVyH]
2. Australian Government. Canberra: The Department of Health, Australian Government; 2009. Obesity in Australia: a need for urgent action URL: http://www.fairfieldcity.nsw.gov.au/download/downloads/id/379/national_health_prevention_obesity.pdf [accessed 2019-03-18] [WebCite Cache ID 76xiUkk46]
3. Australian Institue of Health and Welfare.: Australian Institue of Health and Welfare; 2018. Overweight and Obesity URL: https://www.aihw.gov.au/reports-data/behaviours-risk-factors/overweight-obesity/overview [accessed 2019-03-17] [WebCite Cache ID 76vyl8KtU]
4. Ashwell M, Gunn P, Gibson S. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. Obes Rev 2012 Mar;13(3):275-286. [doi: 10.1111/j.1467-789X.2011.00952.x] [Medline: 22106927]
5. Berghöfer A, Pischon T, Reinhold T, Apovian CM, Sharma AM, Willich SN. Obesity prevalence from a European perspective: a systematic review. BMC Public Health 2008 Jun 05;8:200 [FREE Full text] [doi: 10.1186/1471-2458-8-200] [Medline: 18533989]
6. Ahn A, Choi J. A one-way text messaging intervention for obesity. J Telemed Telecare 2016 Apr;22(3):148-152. [doi: 10.1177/1357633X15591129] [Medline: 26169349]
7. Siopis G, Chey T, Allman-Farinelli M. A systematic review and meta-analysis of interventions for weight management using text messaging. J Hum Nutr Diet 2015 Feb;28(Suppl 2):1-15. [doi: 10.1111/jhn.12207] [Medline: 24480032]
8. Wendel-Vos W, Droomers M, Kremers S, Brug J, van Lenthe F. Potential environmental determinants of physical activity in adults: a systematic review. Obes Rev 2007 Sep;8(5):425-440. [doi: 10.1111/j.1467-789X.2007.00370.x] [Medline: 17716300]
9. Shaw R, Bosworth H. Short message service (SMS) text messaging as an intervention medium for weight loss: a literature review. Health Informatics J 2012 Dec;18(4):235-250 [FREE Full text] [doi: 10.1177/1460458212442422] [Medline: 23257055]
10. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009;11(1):e1 [FREE Full text] [doi: 10.2196/jmir.1100] [Medline: 19141433]
11. 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]
12. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
13. Zheng YG, Klem ML, Sereika SM, Danford CA, Ewing LJ, Burke LE. Self-weighing in weight management: a systematic literature review. Obesity (Silver Spring) 2015 Feb;23(2):256-265. [doi: 10.1002/oby.20946] [Medline: 25521523]
14. Shieh C, Knisely MR, Clark D, Carpenter JS. Self-weighing in weight management interventions: a systematic review of literature. Obes Res Clin Pract 2016;10(5):493-519. [doi: 10.1016/j.orcp.2016.01.004] [Medline: 26896865]
15. Vanwormer JJ, French SA, Pereira MA, Welsh EM. The impact of regular self-weighing on weight management: a systematic literature review. Int J Behav Nutr Phys Act 2008;5:54 [FREE Full text] [doi: 10.1186/1479-5868-5-54] [Medline: 18983667]
16. Shaw RJ, Bosworth HB, Hess JC, Silva SG, Lipkus IM, Davis LL, et al. Development of a theoretically driven mHealth text messaging application for sustaining recent weight loss. JMIR Mhealth Uhealth 2013;1(1):e5 [FREE Full text] [doi: 10.2196/mhealth.2343] [Medline: 25100678]
17. de Koning L, Merchant AT, Pogue J, Anand SS. Waist circumference and waist-to-hip ratio as predictors of cardiovascular events: meta-regression analysis of prospective studies. Eur Heart J 2007 Apr;28(7):850-856. [doi: 10.1093/eurheartj/ehm026] [Medline: 17403720]

18. Medina J, Batsis J, Supervia M, Somers V, Thomas R, Jenkins S, et al. The association between waist-hip ratio with long-term cardiovascular events in patients with coronary artery disease: a population-based historical cohort study. J Am Coll Cardiol 2017 Mar 21;69(11):1843. [doi: 10.1016/S0735-1097(17)35232-4]
19. Australian Government: Department of Health. AUSTRALIA'S PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR GUIDELINES FACT SHEET: ADULTS (18-64 YEARS) URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-phys-act-guidelines/$File/FS-Adults-18-64-Years.PDF [accessed 2019-03-17] [WebCite Cache ID 71PS9D4rk]
20. DeSalvo KB, Fisher WP, Tran K, Bloser N, Merrill W, Peabody J. Assessing measurement properties of two single-item general health measures. Qual Life Res 2006 Mar;15(2):191-201. [doi: 10.1007/s11136-005-0887-2] [Medline: 16468076]
21. Ilmarinen J. The Work Ability Index (WAI). Occup Med 2007 Mar 1;57(2):160 [FREE Full text] [doi: 10.1093/occmed/kqm008]
22. Mykletun RJ, Furunes T. Relationships between work ability and "The Work Ability House"—towards a multi-dimensional work ability model. : American Psychological Association; 2011 Presented at: The 9th International Conference on Occupational Stress and Health; May 19–22, 2011; Orlando URL: https://www.apa.org/wsh/past/2011/conference-program.pdf
23. Maslach C, Jackson SE, Leiter M. Maslach Burnout Inventory. Palo Alto, CA: Consulting Psychologists Press; 1996.
24. Maslach C, Jackson SE. The measurement of experienced burnout. J Organiz Behav 1981 Apr;2(2):99-113 [FREE Full text] [doi: 10.1002/job.4030020205]
25. Hansen V, Pit SW. The single item burnout measure is a psychometrically sound screening tool for occupational burnout. Health Scope 2016 Jan 03;5(2):e32164 [FREE Full text] [doi: 10.17795/jhealthscope-32164]
26. Sax Institute. Sydney; 2008 Jun 20. 45 and Up Study Questionnaire for Women URL: https://www.saxinstitute.org.au/wp-content/uploads/45Up-Study-Female-v3b_ba7a.pdf [accessed 2018-08-04] [WebCite Cache ID 71PT3RLfy]
27. Government of New South Wales. North Sydney: NSW Ministry of Health; 2013. NSW Healthy Eating and Active Living Strategy: Preventing overweight and obesity in New South Wales 2013-2018 URL: https://www.health.nsw.gov.au/heal/publications/nsw-healthy-eating-strategy.pdf [accessed 2019-03-18] [WebCite Cache ID 76xmC31Zd]
28. Yanovski JA, Yanovski SZ, Sovik KN, Nguyen TT, O'Neil PM, Sebring NG. A prospective study of holiday weight gain. N Engl J Med 2000 Mar 23;342(12):861-867 [FREE Full text] [doi: 10.1056/NEJM200003233421206] [Medline: 10727591]
29. Wylie-Rosett J, Delahanty L. An integral role of the dietitian: implications of the Diabetes Prevention Program. J Am Diet Assoc 2002 Aug;102(8):1065-1068 [FREE Full text] [Medline: 12171450]
30. Nawijn J, Marchand MA, Veenhoven R, Vingerhoets AJ. Vacationers happier, but most not happier after a holiday. Appl Res Qual Life 2010 Mar;5(1):35-47 [FREE Full text] [doi: 10.1007/s11482-009-9091-9] [Medline: 20234864]
31. Prestwich A, Perugini M, Hurling R. Can implementation intentions and text messages promote brisk walking? A randomized trial. Health Psychol 2010 Jan;29(1):40-49. [doi: 10.1037/a0016993] [Medline: 20063934]
32. Lee S, Schorr E, Chi CL, Treat-Jacobson D, Mathiason MA, Lindquist R. Peer group and text message-based weight-loss and management intervention for African American women. West J Nurs Res 2018 Dec;40(8):1203-1219. [doi: 10.1177/0193945917697225] [Medline: 28335711]
33. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: a randomized controlled trial. J Med Internet Res 2016;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline: 26818683]
34. Harrer M, Adam SH, Fleischmann RJ, Baumeister H, Auerbach R, Bruffaerts R, et al. Effectiveness of an internet- and app-based intervention for college students with elevated stress: randomized controlled trial. J Med Internet Res 2018 Apr 23;20(4):e136 [FREE Full text] [doi: 10.2196/jmir.9293] [Medline: 29685870]
35. McCoy P, Leggett S, Bhuiyan A, Brown D, Frye P, Williams B. Text messaging: an intervention to increase physical activity among African American participants in a faith-based, competitive weight loss program. Int J Environ Res Public Health 2017 Dec 29;14(4):326 [FREE Full text] [doi: 10.3390/ijerph14040326] [Medline: 28353650]
36. Unick JL, Beavers D, Bond DS, Clark JM, Jakicic JM, Kitabchi AE, Look AHEAD Research Group. The long-term effectiveness of a lifestyle intervention in severely obese individuals. Am J Med 2013 Mar;126(3):236-242 [FREE Full text] [doi: 10.1016/j.amjmed.2012.10.010] [Medline: 23410564]
37. Chen JL, Guedes CM, Cooper BA, Lung AE. Short-term efficacy of an innovative mobile phone technology-based intervention for weight management for overweight and obese adolescents: pilot study. Interact J Med Res 2017 Aug 02;6(2):e12 [FREE Full text] [doi: 10.2196/ijmr.7860] [Medline: 28768612]
38. Douketis JD, Macie C, Thabane L, Williamson DF. Systematic review of long-term weight loss studies in obese adults: clinical significance and applicability to clinical practice. Int J Obes (Lond) 2005 Oct;29(10):1153-1167. [doi: 10.1038/sj.ijo.0802982] [Medline: 15997250]
39. Elobeid MA, Padilla MA, McVie T, Thomas O, Brock DW, Musser B, et al. Missing data in randomized clinical trials for weight loss: scope of the problem, state of the field, and performance of statistical methods. PLoS One 2009;4(8):e6624 [FREE Full text] [doi: 10.1371/journal.pone.0006624] [Medline: 19675667]
40. Colombo O, Ferretti VV, Ferraris C, Trentani C, Vinai P, Villani S, et al. Is drop-out from obesity treatment a predictable and preventable event? Nutr J 2014 Feb 03;13:13 [FREE Full text] [doi: 10.1186/1475-2891-13-13] [Medline: 24490952]

41. Pagoto SL, Schneider KL, Oleski JL, Luciani JM, Bodenlos JS, Whited MC. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. Obesity (Silver Spring) 2012 Jun;20(6):1234-1239. [doi: 10.1038/oby.2011.140] [Medline: 21633403]
42. Williams RL, Wood LG, Collins CE, Callister R. Effectiveness of weight loss interventions--is there a difference between men and women: a systematic review. Obes Rev 2015 Feb;16(2):171-186 [FREE Full text] [doi: 10.1111/obr.12241] [Medline: 25494712]

Abbreviations

BMI: body mass index
RCT: randomized controlled trial
WHR: waist-to-hip ratio

Edited by G Eysenbach; submitted 12.08.18; peer-reviewed by S Lee; comments to author 14.10.18; revised version received 19.01.19; accepted 21.01.19; published 24.04.19.

Please cite as:
Chan R, Nguyen M, Smith R, Spencer S, Pit SW
Effect of Serial Anthropometric Measurements and Motivational Text Messages on Weight Reduction Among Workers: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(4):e11832
URL: https://mhealth.jmir.org/2019/4/e11832/
doi:10.2196/11832
PMID:31017585

Original Paper

Text Messaging to Enhance Behavioral Health Treatment Engagement Among Justice-Involved Youth: Qualitative and User Testing Study

Marina Tolou-Shams[1,2], PhD; Juliet Yonek[1,2], PhD; Katharine Galbraith[1,2], BA; Eraka Bath[3], MD

[1]Division of Infant, Child and Adolescent Psychiatry, Zuckerberg San Francisco General Hospital, University of California, San Francisco, San Francisco, CA, United States

[2]Department of Psychiatry, Weill Institute of Neurosciences, University of California, San Francisco, San Francisco, CA, United States

[3]Department of Psychiatry and Biobehavioral Sciences, Jane and Terry Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Marina Tolou-Shams, PhD
Department of Psychiatry
Weill Institute of Neurosciences
University of California, San Francisco
401 Parnassus Ave
San Francisco, CA,
United States
Phone: 1 415 206 2212
Email: marina.tolou-shams@ucsf.edu

Abstract

Background: Mental health and substance use disorders are highly prevalent in justice-involved youth, yet only 8% of court-involved, nonincarcerated (CINI) youth in need of treatment receive it. Dual diagnosis (co-occurring psychiatric and substance use disorders) in justice-involved youth is highly predictive of recidivism. Identifying novel approaches, such as the use of mobile health (mHealth) technologies, to close this gap between need and receipt of behavioral health treatment for the CINI population could potentially offset rates of reoffending into adulthood. Text-messaging (short message service, SMS) interventions have demonstrated efficacy in improving treatment adherence and other associated outcomes in other vulnerable youth populations, but development and testing of mHealth interventions to improve behavioral health treatment rates and outcomes for CINI youth are lacking.

Objective: This study aimed to collect qualitative data from key stakeholders to inform the development of a theoretically grounded, family-based text-messaging (SMS) intervention targeting CINI youth's behavioral health treatment engagement; additionally, the aim was to conduct end-user testing over 6 months with CINI youth and caregivers to determine intervention feasibility and acceptability.

Methods: CINI youth and caregivers were referred from a California-based Juvenile Probation Department and community-based provider organizations providing services for justice-involved youth. Eligibility criteria included the following: being a justice-involved youth or a caregiver of a justice-involved youth, English speaking, youth aged 13 to 17 years old and either referred to or currently attending mental health or substance use treatment, and youth and caregiver have access to a cell phone with text-messaging capability.

Results: Overall, 28 individuals participated in focus groups and interviews—8 youth, 5 caregivers, and 15 juvenile justice (JJ) personnel. Three major themes emerged: (1) texting among JJ personnel and CINI youth and caregivers in their caseload is common but not systematic, (2) stigma and privacy are perceived as barriers to texting youth about behavioral health treatment appointments, and (3) messages should be short, simple, relatable, positive, and personalized. In total, 9 participants (7 youth and 2 caregivers) participated in end-user testing and rated the intervention as useful, helpful, and supportive.

Conclusions: Text messaging (SMS) is an acceptable and feasible means of reminding CINI youth to attend behavioral health treatment appointments. Future implementation challenges include making text messaging (SMS) personalized and tailored but not resource intensive (eg, requiring one-to-one, 24/7 human contact) and identifying which systems will deliver and sustain the intervention. Text messaging (SMS) among justice personnel, youth, and their caregivers is already widespread, but lack of clear

guidelines about privacy, confidentiality, and information sharing poses ethical conundrums. Future hybrid-type research designs that explore the efficacy of the intervention while also studying ethical, system, and policy-level factors associated with using digital health interventions to improve CINI youth outcomes is a key next step.

(JMIR Mhealth Uhealth 2019;7(4):e10904) doi:10.2196/10904

KEYWORDS

juvenile delinquency; treatment adherence and compliance; mental health; short message service text messaging

Introduction

Background

Mental health and substance use disorders are significantly more prevalent among justice-involved youth than their nonoffending counterparts [1-3]. Among detained youth, it is estimated that approximately 70% to 90% of these individuals have psychiatric symptoms [2,4,5]. Court-involved, nonincarcerated (CINI) youth—those who are legally involved but living in the community—are also more likely to have psychiatric problems compared with general community adolescent samples; between one-third and one-half of this population has a diagnosable psychiatric condition [6]. In addition, 73% have some form of lifetime trauma exposure [7], and approximately 50% of first-time offending youth endorse lifetime marijuana use [8]. Despite high rates of behavioral health (mental health and substance use) disorders among justice-involved adolescents, only 15% of detained youth receive behavioral health treatment; this number falls to 8% once these youth reenter the community [2]. These statistics regarding treatment receipt among justice-involved youth are important to consider from a health care perspective and in terms of public health significance and policy. Dual diagnosis (ie, cooccurring psychiatric and substance use disorders) in justice-involved youth is one of the most significant predictors of recidivism [9], and as such, closing the gap between need and receipt of behavioral health treatment for justice-involved youth could potentially offset rates of reoffending into adulthood [10].

Several barriers contribute to the gap between need and receipt of behavioral health treatment. CINI youth are commonly diverted at various points within the system (eg, arrest, intake, and probation) and referred to different providers and systems throughout their community while still under court supervision. Thus, screening and further assessment of their initial behavioral health needs are not implemented as easily, systematically, or consistently as might be implemented in juvenile detention settings, where all youth come through a central intake. Furthermore, formal assessment does not guarantee referral to treatment, and unfortunately, many of those referred to treatment do not link to providers to initiate treatment (ie, attend the first appointment), particularly when referrals require complex navigation through the community. Even among those who do successfully initiate mental health treatment, only some remain engaged in treatment (defined as at least two visits within 60 days of the first appointment) and continue care (defined as minimum of 3 months of treatment) [11]. From initial intake into the system through the continuum of care, there are a host of barriers that contribute to low rates of treatment receipt among justice-involved youth; these include youths' developmentally appropriate lack of insight, lesser problem recognition and motivation to engage in treatment [12-14], gaps in communication between the justice system and the community-based organizations that provide mental health services, overburdened systems with excessive waiting periods of appointment times, and staff turnover and burnout [11,15]. Racial and ethnic minority youth with fewer individual, family and neighborhood resources are also disproportionately represented in the juvenile justice (JJ) system, thereby presenting additional barriers to behavioral health services access and engagement [16,17]. Justice-involved youth and caregivers also have the additional context of justice system oversight and involvement (eg, mandated treatment, punitive sanctions-based approach to noncompliance) that may or may not affect treatment engagement as compared with nonjustice-involved populations. A review of studies among adult justice-involved populations suggests that mandated or sanction-based approaches overall are not effective in improving substance use or criminal justice outcomes [18]; thus, the field should identify ways to reduce reliance on compulsory approaches to enhancing treatment engagement and outcomes.

Mental health treatment engagement is predictive of improved behavioral health outcomes [19], and thus working to improve rates of treatment receipt by enhancing treatment engagement is critical. The use of digital mobile health (mHealth) technologies has been shown to be a low-cost, efficacious way of reaching vulnerable populations to facilitate treatment engagement [20]. For example, a recent meta-analysis (N=14 studies) concluded that short message service (SMS) text messaging is a promising tool for effective substance use prevention (including relapse) for nonoffending adolescents and young adults [21]. In another study, the use of bidirectional SMS text messaging with caregivers to enhance adolescents' receipt of vaccine and well-care services improved adolescents' utilization of both services [22], suggesting that caregiver involvement in adolescent-focused mHealth interventions may also be effective in improving other outcomes (eg, engagement in mental health treatment). Studies show that only some, but not all, caregivers are ready for electronic messaging support for health care [23] and that depending on caregiver race (eg, Latino), socioeconomic status (eg, low), and age (eg, younger), SMS text messaging may be more or less appealing as a tool for their adolescent's health care engagement [19,24].

The use of mHealth technology presents a promising approach for closing the gap between CINI youth's need and receipt of behavioral health treatment. Mobile phone usage among justice-involved youth, particularly those supervised in the community, is also widespread [25]. However, to our knowledge, an empirically supported mHealth technology

intervention specifically tailored to justice-involved youth and their caregivers does not currently exist.

Objectives

This study developed and conducted end-user testing of a dyadic (youth and caregiver) SMS text-messaging intervention that included sending appointment reminders and motivational messages to enhance likelihood of the youth attending face-to-face community-based treatment, as referred by probation staff. Our aim was to understand what key system stakeholders (clinicians and probation staff) and end users (youth and their families) thought was feasible and acceptable regarding the use of SMS text messaging to enhance treatment engagement.

Methods

Study Overview, Population, and Recruitment

Pilot study data collection was completed in 2 phases: (1) Development (focus groups and interviews with CINI youth, caregivers, and JJ personnel to inform the SMS text-messaging intervention) and (2) End-User Testing (with CINI youth and caregivers). For both phases, youth and caregivers were referred from a California-based Juvenile Probation Department and community-based provider organizations that served justice-involved youth and their caregivers. Probation staff and community-based providers referred interested youth and families to the study staff, and the study staff screened referred youth and caregivers for eligibility. Youth eligibility criteria included being English speaking, between 13 and 17 years old, justice involved, either referred to or currently attending mental health or substance use counseling, and have a personal cell phone with SMS text-messaging capability. Caregiver eligibility included being English speaking and have a personal cell phone with SMS text-messaging capability. JJ personnel included any probation staff and providers (eg, case managers, behavioral health clinicians, and social workers) serving justice-involved youth and their families in the same geographical region as youth and caregivers in the study. JJ personnel were recruited through emails and follow-up phone calls, with assistance from JJ administrators. In the Development phase, caregivers provided written informed consent for their or their youth's participation and youth provided separate assent. JJ personnel provided written informed consent for their participation. In the End-User Testing phase, because of challenges with reaching caregivers in person, parental consent was waived for youth participation and youth completed in-person written consent. For youth who had involved caregivers and gave permission to the study staff to contact them, interested caregivers provided verbal consent (by phone) for their separate End-User Testing phase participation. Institutional Review Board approval for the study was obtained from the Principal Investigator's (PI's) institution before any data collection.

Study Procedures: Development Phase

In total, 4 focus groups (1 JJ probation staff, 1 JJ providers, and 2 youth) and 5 individual caregiver phone interviews were conducted between October 2016 and February 2018. Before starting focus groups and interviews, participants completed a brief demographics questionnaire. The focus group and interview guides were developed by the study PI (MTS), a child psychologist with expertise in designing and implementing behavioral health interventions targeting justice-involved youth as well as qualitative methodology, and a health services researcher with expertise in qualitative methodology (JY). Focus groups and interviews were conducted by the PI and 2 research staff members, all of whom were trained in both qualitative data collection methods. Focus groups were either conducted in a private conference room in the probation department (probation staff) or community partner settings (providers and youth). All interviews and focus group sessions were audio recorded with consent. The youth and JJ staff focus groups each lasted for approximately 90 min, and the individual caregiver phone interviews lasted for 60 min. Youth and caregivers were each compensated US $25 for the focus group and individual interviews, respectively. JJ personnel were compensated US $25 for their participation, if allowable by their organization.

Focus Groups: Youth

Groups began with an *ice breaker* section regarding participants' general cell phone use and texting patterns with JJ personnel and caregivers. Next, youth were asked about the types of messages they would find most helpful or effective in increasing their attendance and engagement in mental health and substance use treatment. For example, participants were asked about messages serving as appointment reminders, messages about the benefits of completing treatment and other probation-related requirements on time, and messages providing positive reinforcement for attending appointments. Facilitators also asked participants for specific feedback regarding SMS text message structure, including message frequency, language, level of interactivity, and possible concerns, such as privacy and participant burden.

Focus Groups: Juvenile Justice Personnel

Groups began by asking personnel to provide their perspective on the acceptability, practicality, and feasibility of administering an SMS text message–based system within a probation department or a treatment setting. Participants were asked about their current texting practice with youth or caregivers as well as to describe the individual, family, and system-level benefits to and challenges associated with using an SMS text-messaging system to send systematic appointment reminders and motivational messages to youth.

Interviews: Caregivers

Caregivers were first asked to describe their level of involvement in their youth's court and treatment-related appointments, including their communication with youth about appointments. They were also asked to describe their preferred mode of communication with their child (eg, phone, email, and text) as well as perceived barriers and benefits to using SMS text messaging to remind youth of their appointments and keep them engaged in treatment. Caregivers were also asked to describe their interaction and communication with their youth's probation officers or clinical providers, including preferred forms of communication (texting, phone, and in-person visits).

Study Procedure: End-User Testing Phase

SMS text message development and content were informed by results from the Development phase. Participants directly received SMS text messages for a period of 6 months. Automated appointment reminders were sent to both CINI youth and caregivers 3 days before, 1 day before, and on the day of the appointment. A follow-up message was sent after the appointment to find out whether the youth attended their appointment, and if not, why not. Prescripted motivational messages were sent twice a week (eg, on Monday and Friday), and they did not include any words related to mental health or substance use (per phase 1 participant feedback).

Perceived usefulness, acceptability, and recommendations for improvement were assessed via repeat Web-based surveys over 6 months (administered at 1, 3, and 5 months). CINI youth and caregivers were asked for their general opinions about the intervention (ie, clear, helpful, and user-friendly), the motivational messages (ie, interesting, motivating, and boring), and the reminders (ie, helpful). Youth and caregivers received US $25 for each survey completed.

Qualitative Data Analysis

Interviews, focus group recordings, and written notes were reviewed by research team members for accuracy and completeness. This information was used to construct an executive summary of the main discussion points and topics within 24 hours of conducting focus groups or individual interviews, and the information was used to identify commonly reported themes. Themes were refined on the basis of group discussions by the research team, led by the PI (MTS). Illustrative quotes were then extracted for each theme.

Survey Data Analysis

Given the pilot nature of this intervention, participants completed surveys at 1, 3, and 5 months (within the 6-month intervention period) to inform iterative refinement and obtain feedback on changes over time. Descriptive statistics were used to summarize 5-month (final) follow-up survey results for CINI youth and caregivers who participated in user testing.

Results

Development Phase

There was a total of 28 study participants—8 youth, 5 caregivers, and 15 JJ personnel (8 probation staff and 7 providers). There was a single youth and caregiver dyad; the remaining youth and caregiver participants were not related. Sample youth and caregiver demographic characteristics are in Table 1.

Key Development Phase Themes

In total, 3 major themes (Figure 1) emerged from the analysis of focus groups and interviews: (1) texting among JJ personnel and youth and caregivers in their caseload is common but not systematic, (2) stigma and privacy are perceived as barriers to texting youth about mental health and substance use treatment appointments, and (3) messages should be short, simple, relatable, positive, and personalized.

Theme 1: Texting Among Juvenile Justice Personnel, Youth, and Caregivers Is Common but Not Systematic

Youth, caregivers, and JJ personnel were universally enthusiastic about an SMS text message–based system to help remind and encourage youth to attend mental health or substance use treatment appointments and complete treatment. All youth said that they already use SMS text messaging to communicate with their JJ personnel contact (eg, probation officer, clinician, or case manager) primarily to check in, such as "They check up on me," and "They ask how I'm doing in school," but also to schedule appointments and to obtain general advice and support, such as:

> *One time I was at work and got really mad at a co-worker. I didn't want to get in trouble so I texted my case manager for help. I got a text back helping me…telling me what to do.*
>
> *I'll text my case manager on a daily basis, if I have a question on something.*

All caregivers (n=5) were highly interested and invested on obtaining SMS text message reminders of their youth's responsibilities and mandates to support or monitor; however, 3 caregivers noted that they were not as facile as the youth in texting and that their youth often ignore their messages, which then results in a phone call (that some caregivers preferred over texting to begin with). One caregiver noted:

> *I'm not really a texter.*
>
> *I'd rather talk [to my daughter] and get an answer right there instead of having to wait for the answer [via text].*

JJ personnel said they commonly text youth to let them know they are trying to get a hold of them, to *check in* and to remind them of their appointments. However, the frequency of texting for these reasons varies for each child and thus is not systematic. JJ personnel commented, for example:

> *Some kids respond best to [a text message that says], "every Tuesday at 2 pm I will be here at your school." But for some kids that's not going to work. Everything needs to be very individualized for each kid. Their situations are so up and down that it could be one way for 2 months and then change in a completely different way.*
>
> *They may meet Tuesday this week and Thursday next week. There's a lot of fluidity in the scheduling because they're all over the place.*

In particular, the probation staff stated that they frequently use texting to communicate with caregivers about their teen's appointments as they are more likely to respond to a text than a phone call. One probation officer explained:

> *I can depend on a text message to the parent to remind the kid more than a voicemail because I don't know if they're going to listen to it.*

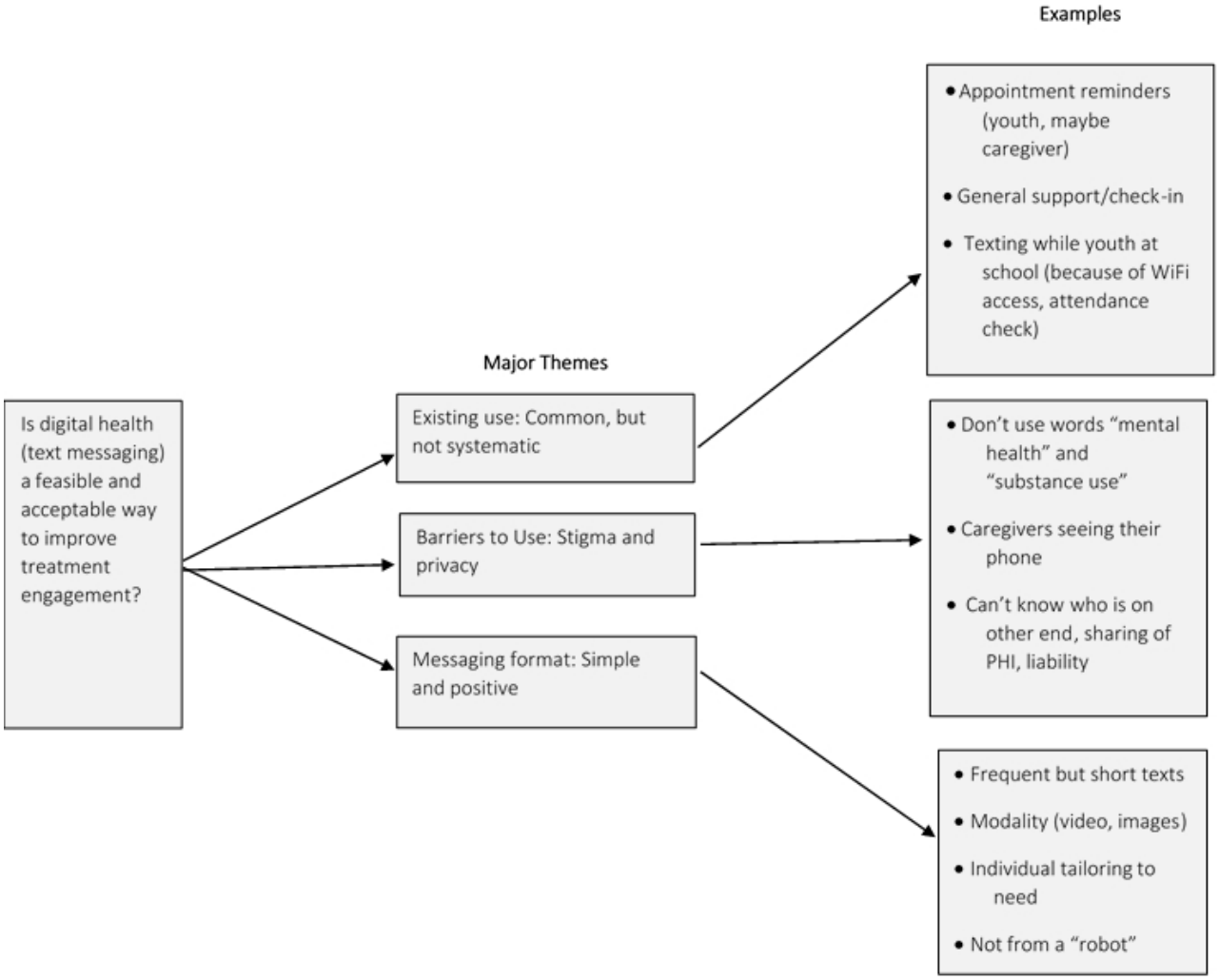
Table 1. Demographic characteristics of court-involved, nonincarcerated youth and caregivers.

Demographics	Development phase	End-user testing
Youth (n)	8	7
Median age (years)	17	16
Gender, n (%)		
Female	6 (75)	1 (14)
Male	2 (25)	8 (8)
Race, n (%)		
White	0 (0)	0 (0)
African American/black	3 (38)	2 (29)
Asian	1 (13)	1 (14)
Multiracial	2 (25)	0 (0)
Native Hawaiian or other Pacific Islander	1 (13)	0 (0)
American Indian or Alaska native	0 (0)	1 (14)
Other	1 (13)	3 (43)
Hispanic ethnicity, n (%)	3 (38)	6 (86)
Owned a cell phone (yes), n (%)	6 (75)[a]	7 (100)
Caregiver (n)	5	2
Relationship to youth		
Biological caregiver	4 (80)	2 (100)
Age range, n (%)		
35-44	1 (20)	N/A[b]
45-54	2 (40)	1 (50)
55-64	1 (20)	1 (50)
65+	1 (20)	N/A
Gender, n (%)		
Female	4 (80)	2 (100)
Race, n (%)		
White	0 (0)	1 (50)
African American/black	4 (80)	1 (50)
Asian	0 (0)	0 (0)
Multiracial	0 (0)	0 (0)
Native Hawaiian/other Pacific Islander	1 (20)	0 (0)
American Indian or Alaska native	0 (0)	0 (0)
Other	0 (0)	0 (0)
Hispanic ethnicity, n (%)	0 (0)	0 (0)
Owned a cell phone (yes), n (%)	5 (100)	2 (100)

[a]Owning a personal cell phone was only required for teens participating in the End-User Testing phase.

[b]N/A: not applicable.

Figure 1. Primary themes on the feasibility of short message service text messaging for justice-involved youth. PHI: Personal Health Information.



JJ providers specifically expressed that automating text-based appointment reminders would be very helpful, particularly if reminders could be generated directly from the calendar on their cell phone or computer. Examples of suggestions made included:

> *With the Google calendar you can do a reminder and that's very helpful. But if there was also a way to set that calendar up to text or send a text to whomever – as you put it in your calendar there's an option to send a text – that would be very useful.*

> *What would be helpful is being able to input all of my appointments scheduled for that week. Being able send those reminders weekly, when I'm setting it up in a calendar and then those messages with reminders are getting sent out so I don't even have to think about it.*

Entering appointment dates and times into a separate system felt like *additional work* and less efficient than sending appointment reminders themselves, especially given the frequency of 1-time (nonreoccurring) appointments. One JJ provider stated:

> *Going into a system that has to also be altered constantly…that feels like more work than sending out the text [directly]. I would have to figure out a system, what's the schedule, what do I need to remind them of? It feels like almost an additional step. The large majority of the work we do is not "at every Tuesday at 2 PM we do this."*

Theme 2: Barriers to Use—Stigma and Privacy

Youth, caregivers, and JJ personnel were adamant that neither appointment reminders nor motivational messages should mention mental health or substance use. For example, one youth stated:

> *You have an appointment at such and such, but don't say what it's for.*

A caregiver separately commented:

> *Instead say "hey, you have a session" or "hey, you have an appointment."*

JJ providers explicitly confirmed that sending youth a text directly asking about substance use or mental health would not be well received, as illustrated by statements such as:

> *That wouldn't be really comfortable…because those are sensitive subjects and that sensitivity isn't going to come across well through text. It is going to be really dry.*

> *I haven't thought of a time where I would have done that. I can't even imagine texting like 'are you okay?'*

> *It would be more like, if I am concerned, I am calling. I am not going to text.*

Although youth said they use the password feature to lock their phone, some reported sharing their password with certain friends and dating partners. Most also said that their caregivers tried to look at their phones at least once, even when they were not given permission to do so. One youth described when her parent tried to take her phone while sleeping:

> *One time my mom came in there when I was sleeping and was going through my phone, SnapChat, Twitter, she was just being nosy. I was so mad.*

Another youth described why privacy was important (ie, that their phone held all of their sensitive information):

> *You're not getting my phone. I have too much information up in here.*

JJ personnel also expressed several concerns about youth privacy when sending texts, including that there is no way to know whether the phone is in the youth's possession or who is on the other side of the text and whether it is the youth whom they are intending to contact:

> *Texts are there forever. If youth doesn't have a passcode, then anyone can see it.*

Concerns were also raised by probation staff, specifically that text conversations can also be subpoenaed, and there is currently no system-level policy or protocol on how to manage this. One probation officer stated:

> *Some of these kids have warrant searches on their phones…So a lot of times they don't want to communicate over text.*

Probation officers expressed concern over what aspects of SMS text messages are part of the legal record and what constitutes protected health information when shared over texting. JJ providers separately expressed similar concern over how to handle potentially self-incriminating information sent to them by the youth or caregiver. JJ providers additionally raised uncertainty and questions regarding whether text conversations with youth should be documented in the treatment chart as part of clinical care; this might be akin to written collateral contact logs often included in client charts as part of standard clinical practice to document any contact outside of scheduled appointments. For example, 1 provider asked the question:

> *Since the communication [via texting] is happening during work hours, what responsibility do we have to maintain those records of that communication?*

Another JJ provider expressed concerns that *not* documenting text conversations might diminish perceptions of their actual workload:

> *Are we really keeping track of how much outreach and how much that we're doing [via texting]? I can say no, we are not writing our [text conversations] in the file every time, and so then it doesn't really look like we're doing all of the stuff that we're doing…yeah it might be on your phone, but that doesn't do us any good when we're talking about the work that we've done.*

Finally, JJ personnel shared concern about how their own privacy can be compromised when using their personal cell phones to text youth, which all JJ personnel reported occurs by virtue of necessity and lack of system resources to provide work cell phones for probation or provider personnel.

Theme 3: Messages Should Be Short, Simple, Relatable, Positive, and Personalized

Youth and JJ personnel emphasized that messages should be short and conceptually straightforward. One youth raised the option of visual, Multimedia Messaging Service (MMS), or video messaging indicating that:

> *A video would be cool. Not longer than a minute though.*

Multiple youth mentioned that SMS text language should be conversational and friendly, noninvasive, and worded in a way that sounds like something a teenager would say (eg, language that they can relate to), such as:

> *It has to be short and sweet*
>
> *How they say it is important. Don't just pile it on like you got this this this and this. Let me know what I have to do.*

Another youth stated:

> *It would be cool if it's an app that reminds you of your appointments. Not something that asks you questions all about your life.*

Positively worded messages were universally endorsed by youth and caregivers. Youth stated that they would prefer texts emphasizing that it is their choice to attend mental health or substance use treatment appointments rather than being told what to do, such as:

> *It's your choice. Get what you have to do to get things done so you can do what you want to do.*

Caregivers provided examples of SMS text messages for appointments that they perceived might be motivating for their youth, including messages that reminded their youth of longer-term goals (eg, getting back to *normal life*, not having to worry as much), such as:

> *Get it over with so you don't have to keep worrying, instead of making it last another 3 months.*
>
> *The sooner we can get this behind us the sooner we can get back to our normal life.*

Other suggestions made by youth and JJ probation staff specifically included having positive, nonjudgmental words in the texts, such as "Keep up the positive work. You're really handling things well," and 1 probation staff suggested, "Remind them of the positive… feeding to their strengths." Similarly, JJ providers stressed the importance of *tone* to get kids' attention and keep them engaged. Some staff use emojis (ie, ideograms and smileys used in electronic messages and Web pages) to convey a positive tone or use a humorous character to relay the message (eg, Mickey Mouse). For example:

> *The winky eye is probably the one I use the most, and then the one with the big old giant smile.*

> *I use the hand waving one. It's like giving them praise.*

Youth universally stated that they wanted messages to be individualized and bidirectional rather than generic, even if automated. They expressed a desire for messages to come from an actual person, not a *robot*:

> *I would rather have somebody actually texting me on the other side and they would reply when they got my texts.*

Caregivers (n=4) and JJ providers expressed a similar viewpoint, indicating that messages need to be more personal for youth to pay attention to them. One caregiver explicitly stated:

> *If you personalize it through an automated program [rather than a real person], it might feel kind of fake.*

End-User Testing Phase

On the basis of Development phase findings, the bidirectional SMS text-messaging intervention comprised 2 major components: (1) mental health or substance use treatment appointment reminders and (2) short motivational messages to enhance engagement and retention (Textboxes 1 and 2). Messages were intended to be simple and positive, and the theme of privacy, for example, was addressed by not including any language in texts, such as *mental health* or *substance use*.

A total of an additional 8 youth and 2 caregivers participated in the 6-month user-testing phase; 5-month survey data were available for 7 youth and 2 caregivers (1 youth only completed the 1-month survey). All youth received SMS text-messaging services for an average of 180.25 days (approximately 6 months). Of the 7 youth who completed a 5-month survey, all reported that the SMS text-messaging intervention helped them attend their treatment appointments. The feature that youth liked the most was being able to read messages at their convenience (n=6), followed by the ability to save reminders on their phone (n=4). The most helpful part of the SMS text-messaging intervention was that it reminded youth of appointments *they had completely forgotten about* (n=5). Most youth liked the motivational SMS text messages (n=5), felt they were supportive (n=5), and felt they were relevant (n=5). The majority of youth (n=6) felt that the motivational messages made them want to attend their counseling appointments. Youth did not feel any changes to the SMS text-messaging intervention were needed, and all would recommend it to a friend who is in counseling.

Both caregivers found the messaging system easy to use. Like youth, caregivers appreciated being able to read messages at their convenience and felt that the messages helped them remind their child of their upcoming appointments. They also liked the wording or tone of the SMS text messages, and they liked the message frequency. They agreed that the motivational messages made them want to encourage or support their child in attending their appointments. Neither caregiver felt that any aspect of the SMS text-messaging intervention should be changed.

Textbox 1. Sample motivational messages sent to youth.

- Taking care of yourself includes making it to your appointments. Be sure to take care of YOU!
- Staying on top of your appointments is a great first step to keeping healthy habits
- Keep your head up—stay positive
- Always remember, you are braver than you believe, stronger than you seem
- Each appt. you attend will bring you closer to reaching your goals
- Your success in the future=making it to your appts. in the present

Textbox 2. Sample motivational messages sent to caregivers.

- Remember to take care of you in the midst of taking care of your child
- Encouraging your child to stay on top of their appointments will help them keep healthy habits
- One small positive thought can change your whole day
- The struggle you are in today is developing the strength you need for tomorrow. Keep moving forward!
- Helping your child get to their appts. Will bring them closer to their goals
- Your child's future success=making it to their appts. in the present

Discussion

Principal Findings

Our small first-time pilot exploration led to 2 overarching findings. First, the bidirectional use of informal SMS text messaging between JJ personnel and youth (and caregivers) in their caseloads is acceptable and commonly occurring; however, current SMS text-messaging practices are nonsystematic and idiosyncratic, depending on a multitude of individual, family, or system-level factors. Second, a formal SMS text-messaging service provided for 6 months to youth on probation, which comprises behavioral health appointment text reminders and motivational SMS text messages, appears feasible to implement and acceptable to end users.

JJ personnel, caregivers, and youth shared information that not only informed the development and content of our pilot SMS text-messaging service but also provided ideas for future research. For example, JJ personnel and caregiver stakeholders

shared that utilizing SMS text messaging for quick *check-ins* as well as brief reminders for upcoming important appointments (including treatment and counseling, court, school, and job) was frequent and generally accepted by all stakeholders (including youth) as helpful, supportive, and much easier than utilizing phone or in-person strategies for reminders and engagement. Youth were also clear that messages should be brief and not contain any stigmatizing language or reveal sensitive information. Some youth verbalized a preference for visual (eg, MMS) messaging content as part of reminders and motivational check-ins. Youth and JJ personnel also wanted the messaging to be tailored and bidirectional and not rote or automatic. Incorporating some of these elements into the piloted SMS text-messaging service may have been responsible for youth and caregiver's high acceptability and satisfaction with the SMS text-messaging program. The challenge for future research and implementation may be to identify how to create chatbots or other highly tailored ways of engaging youth that do not require 1:1 24/7 human resources (as our service did) but still provide the experience of tailoring and personalization. Similarly, identifying whether MMS content is more engaging and effective than SMS content and whether there is certain content that is more engaging for caregivers than for youth are all important areas for future research.

Perhaps of most notable significance to informing future digital health research and practice in this area was JJ personnel's recognition of the multiple systems-level factors they confront when informally texting with youth. Without clear policy, guidance, and protocol on acceptable SMS text-messaging practices with justice-involved youth and within justice settings, such as probation, JJ personnel are unsure about critical issues related to privacy, confidentiality, and stigma. Examples include the following: (1) what aspects of current information-sharing (eg, through email, phone, and release of records) practice apply to texting, (2) what constitutes self-incriminating information via text (eg, if youth inadvertently admit to using a substance and communicating it via an SMS text message to their probation officer), (3) what and when can texts be subpoenaed, (4) how can one confirm that the private information being shared via text is arriving to the intended recipient, (5) how to or whether to document texting and information contained within texts in legal or clinical care records (as collateral), and (6) how to exert boundaries around texting (eg, what responsibility does a probation officer have to respond back to youth texting late at night about being in an emergent, unsafe, or illegal situation?). This pilot study utilized an external SMS text-messaging platform with aspects such as timing, frequency, and content standardized across participants. Whether the decision is made by JJ systems and partnering providers to utilize external SMS text-messaging service supports or to develop internal policies and procedures regarding the use of informal SMS text messaging with youth and families, these complex issues must be addressed. Future research that explores these ethical questions and other system- and policy-level factors associated with barriers to and facilitators of using a digital health intervention to improve justice-involved youth's behavioral health outcomes is a key next step in the field.

Limitations and Future Directions

This first, small pilot study starts an important scientific conversation about use of digital health technology to improve mental health and substance use treatment engagement (and subsequently outcomes) for this vulnerable youth population, but the study is not without its limitations. Data collection was limited to a small number of stakeholders and end users in 1 small geographical region of the United States; however, incorporation of perspectives from youth, caregivers, *and* JJ personnel allowed us to triangulate across key stakeholders and helped us begin to identify commonalities and differences in various stakeholders' perception of what would be acceptable and feasible. JJ personnel included frontline probation and provider staff but not higher-level administration and decision makers with respect to policy- and system-level changes; given our preliminary findings, future research will want to consider how to add this level of stakeholder perspective, given they would be key to implementation. Overall, our data suggest that it is not clear that, in day-to-day practice, all youth are getting the texting and reminders and support, and our preliminary findings suggest that text messaging is not done in any sort of systematic or consistent way, which will ultimately, differentially impact outcomes. The user testing data provide some initial demonstration that it is feasible and acceptable for youth and caregivers to receive these SMS text messages (as part of an outside service) to remind and motivate them to attend treatment appointments, but what remains unknown is whether such an intervention can lead to improved youth outcomes. Following youth to measure (in more detail) current texting practices and the preliminary impact of messaging on youth treatment engagement was not a part of this pilot study. Ultimately, the impact of SMS text messaging on rates of appointment compliance, treatment attendance, and school absenteeism remains an empirical question. Do these SMS text-messaging reminders actually lead to improved adherence to scheduled appointments, such as with a case manager, a court hearing, or mandated substance use counseling? Is that outcome moderated by who sends the SMS text message (ie, probation officer vs treatment provider) or is there an additive, enhanced effect of getting coordinated reminders from all those involved in the youth's care (eg, to demonstrate communication across systems and strong care coordination)? When the caregiver is also included in these SMS text messages, does this increase the likelihood of more positive youth attendance and engagement outcomes? Examples of future directions therefore include the following: (1) an efficacy trial of this SMS text-messaging intervention with appointment reminders and motivational messages to enhance treatment engagement to examine actual impact on outcomes (eg, does the SMS text messaging actually lead to increased treatment attendance, fewer symptoms, and lower rates of recidivism), (2) identifying the systems and organizational factors associated with uptake of such intervention, and (3) study of the ethical issues associated with the potential for self-incrimination, information-sharing, and other privacy concerns associated with sharing information with youth and caregivers over texting. These are all important next steps for this nascent area of highly significant digital mental health research.

Acknowledgments

The authors are grateful to the JJ collaborators and court-involved youth and families for their participation and contributions to this study. This publication was supported by the National Center for Advancing Translational Sciences, National Institutes of Health (NIH), through University of California San Francisco, Clinical and Translational Science Institute Grant Number UL1 TR001872. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Conflicts of Interest

None declared.

References

1. Dembo R, Jainchill N, Turner C, Fong C, Farkas S, Childs K. Levels of psychopathy and its correlates: a study of incarcerated youths in three states. Behav Sci Law 2007;25(5):717-738. [doi: 10.1002/bsl.784] [Medline: 17899531]
2. Teplin LA, Abram KM, Washburn JJ, Welty LJ, Hershfield JA, Duncan MK. US Department of Justice. Washington, DC; 2013 Feb. The northwestern juvenile project: overview URL: https://www.ojjdp.gov/pubs/234522.pdf [accessed 2019-02-24] [WebCite Cache ID 76QeJm8Ou]
3. Vermeiren R, Jespers I, Moffitt T. Mental health problems in juvenile justice populations. Child Adolesc Psychiatr Clin N Am 2006 Apr;15(2):333-51, vii-viii. [doi: 10.1016/j.chc.2005.11.008] [Medline: 16527659]
4. Cauffman E. A statewide screening of mental health symptoms among juvenile offenders in detention. J Am Acad Child Adolesc Psychiatry 2004 Apr;43(4):430-439. [doi: 10.1097/00004583-200404000-00009] [Medline: 15187803]
5. Plattner B, Aebi M, Steinhausen HC, Bessler C. [Psychopathological and comorbid disorders of incarcerated adolescents in Austria]. Z Kinder Jugendpsychiatr Psychother 2011 Jul;39(4):231-240. [doi: 10.1024/1422-4917/a000113] [Medline: 21667448]
6. Gavazzi SM, Yarcheck CM, Sullivan JM, Jones SC, Khurana A. Global risk factors and the prediction of recidivism rates in a sample of first-time misdemeanant offenders. Int J Offender Ther Comp Criminol 2008 Jun;52(3):330-345. [doi: 10.1177/0306624X07305481] [Medline: 17684119]
7. Hoskins D, Marshall BD, Koinis-Mitchell D, Galbraith K, Tolou-Shams M. Latinx youth in first contact with the justice system: trauma and associated behavioral health needs. Child Psychiatry Hum Dev 2018 Nov 27:1-14. [doi: 10.1007/s10578-018-0855-z] [Medline: 30483922]
8. Hirschtritt ME, Dauria EF, Marshall BD, Tolou-Shams M. Sexual minority, justice-involved youth: a hidden population in need of integrated mental health, substance use, and sexual health services. J Adolesc Health 2018 Oct;63(4):421-428. [doi: 10.1016/j.jadohealth.2018.05.020] [Medline: 30077548]
9. Tolou-Shams M, Rizzo CJ, Conrad SM, Johnson S, Oliveira C, Brown LK. Predictors of detention among juveniles referred for a court clinic forensic evaluation. J Am Acad Psychiatry Law 2014;42(1):56-65. [Medline: 24618520]
10. Colins O, Vermeiren R, Vahl P, Markus M, Broekaert E, Doreleijers T. Psychiatric disorder in detained male adolescents as risk factor for serious recidivism. Can J Psychiatry 2011 Jan;56(1):44-50. [doi: 10.1177/070674371105600108] [Medline: 21324242]
11. Belenko S, Knight D, Wasserman GA, Dennis ML, Wiley T, Taxman FS, et al. The Juvenile Justice Behavioral Health Services Cascade: a new framework for measuring unmet substance use treatment services needs among adolescent offenders. J Subst Abuse Treat 2017 Mar;74:80-91 [FREE Full text] [doi: 10.1016/j.jsat.2016.12.012] [Medline: 28132705]
12. Cornelius T, Earnshaw VA, Menino D, Bogart LM, Levy S. Treatment motivation among caregivers and adolescents with substance use disorders. J Subst Abuse Treat 2017 Dec;75:10-16 [FREE Full text] [doi: 10.1016/j.jsat.2017.01.003] [Medline: 28237049]
13. Fisher JH, Lichvar E, Hogue A, Dauber S. Perceived need for treatment and engagement in mental health services among community-referred racial/ethnic minority adolescents. Adm Policy Ment Health 2018 Dec;45(5):751-764. [doi: 10.1007/s10488-018-0863-0] [Medline: 29525929]
14. Sylwestrzak A, Overholt CE, Ristau KI, Coker KL. Self-reported barriers to treatment engagement: adolescent perspectives from the national comorbidity survey-adolescent supplement (NCS-A). Community Ment Health J 2015 Oct;51(7):775-781. [doi: 10.1007/s10597-014-9776-x] [Medline: 25326732]
15. Seave PL. Evidence-based practices reduce juvenile recidivism: can state government effectively promote implementation among probation departments? Am J Community Psychol 2011 Sep;48(1-2):138-140. [doi: 10.1007/s10464-010-9416-8] [Medline: 21264682]
16. Malhotra K, Shim R, Baltrus P, Heiman HJ, Adekeye O, Rust G. Racial/ethnic disparities in mental health service utilization among youth participating in negative externalizing behaviors. Ethn Dis 2015;25(2):123-129. [Medline: 26118137]
17. Saloner B, Carson N, Lê Cook B. Explaining racial/ethnic differences in adolescent substance abuse treatment completion in the United States: a decomposition analysis. J Adolesc Health 2014 Jun;54(6):646-653 [FREE Full text] [doi: 10.1016/j.jadohealth.2014.01.002] [Medline: 24613095]

18. Werb D, Kamarulzaman A, Meacham MC, Rafful C, Fischer B, Strathdee SA, et al. The effectiveness of compulsory drug treatment: a systematic review. Int J Drug Policy 2016 Feb;28:1-9 [FREE Full text] [doi: 10.1016/j.drugpo.2015.12.005] [Medline: 26790691]
19. Garnick DW, Lee MT, O'Brien PL, Panas L, Ritter GA, Acevedo A, et al. The Washington circle engagement performance measures' association with adolescent treatment outcomes. Drug Alcohol Depend 2012 Aug 01;124(3):250-258 [FREE Full text] [doi: 10.1016/j.drugalcdep.2012.01.011] [Medline: 22364777]
20. Batastini AB. Improving rehabilitative efforts for juvenile offenders through the use of telemental healthcare. J Child Adolesc Psychopharmacol 2016 Apr;26(3):273-277. [doi: 10.1089/cap.2015.0011] [Medline: 26288198]
21. Mason M, Ola B, Zaharakis N, Zhang J. Text messaging interventions for adolescent and young adult substance use: a meta-analysis. Prev Sci 2015 Feb;16(2):181-188. [doi: 10.1007/s11121-014-0498-7] [Medline: 24930386]
22. O'Leary ST, Lee M, Lockhart S, Eisert S, Furniss A, Barnard J, et al. Effectiveness and cost of bidirectional text messaging for adolescent vaccines and well care. Pediatrics 2015 Nov;136(5):e1220-e1227 [FREE Full text] [doi: 10.1542/peds.2015-1089] [Medline: 26438703]
23. Rand CM, Blumkin A, Vincelli P, Katsetos V, Szilagyi PG. Parent preferences for communicating with their adolescent's provider using new technologies. J Adolesc Health 2015 Sep;57(3):299-304. [doi: 10.1016/j.jadohealth.2015.06.006] [Medline: 26299557]
24. Smaldone A, Stockwell MS, Osborne JC, Cortes Y, Bekele E, Green NS. Adolescent and parent use of new technologies for health communication: a study in an urban latino community. J Public Health Res 2015 Feb 20;4(1):376 [FREE Full text] [doi: 10.4081/jphr.2015.376] [Medline: 25918691]
25. Bath E, Tolou-Shams M, Farabee D. Mobile health (mHealth): building the case for adapting emerging technologies for justice-involved youth. J Am Acad Child Adolesc Psychiatry 2018 Dec;57(12):903-905. [doi: 10.1016/j.jaac.2018.08.012] [Medline: 30522732]

Abbreviations

CINI: court-involved, nonincarcerated
JJ: juvenile justice
mHealth: mobile health
MMS: Multimedia Messaging Service
NIH: National Institutes of Health
PI: principal investigator
SMS: short message service

Edited by G Eysenbach; submitted 27.04.18; peer-reviewed by L McGoron, M Krausz; comments to author 31.08.18; revised version received 16.11.18; accepted 26.01.19; published 05.04.19.

Please cite as:
Tolou-Shams M, Yonek J, Galbraith K, Bath E
Text Messaging to Enhance Behavioral Health Treatment Engagement Among Justice-Involved Youth: Qualitative and User Testing Study
JMIR Mhealth Uhealth 2019;7(4):e10904
URL: https://mhealth.jmir.org/2019/4/e10904/
doi:10.2196/10904
PMID:30950808

Original Paper

Text Message Responsivity in a 2-Way Short Message Service Pilot Intervention With Adolescent and Young Adult Survivors of Cancer

Alexandra M Psihogios[1], PhD; Yimei Li[2], PhD; Eliana Butler[1], BA; Jessica Hamilton[3], PhD; Lauren C Daniel[4], PhD; Lamia P Barakat[2], PhD; Christopher P Bonafide[2], MD, MSCE; Lisa A Schwartz[2], PhD

[1]The Children's Hospital of Philadelphia, Philadelphia, PA, United States

[2]The Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA, United States

[3]University of Pittsburgh, Pittsburgh, PA, United States

[4]Rutgers University Camden, The Children's Hospital of Philadelphia, Camden, NJ, United States

Corresponding Author:
Alexandra M Psihogios, PhD
The Children's Hospital of Philadelphia
3501 Civic Center Boulevard
Philadelphia, PA, 19104
United States
Phone: 1 267 314 0164
Email: psihogiosa@email.chop.edu

Abstract

B**ackground:** Text message interventions hold promise for adolescents and young adults (AYAs) with chronic health conditions, including childhood cancer survivors; however, engagement is often suboptimal. Limited research has studied mobile health intervention outcomes beyond efficacy. Understanding responsivity to different types of text messages (ie, when a participant texts back) can provide practical, actionable information to optimize engagement in future projects.

Objective: Within a 2-way text messaging study in AYAs who recently completed treatment for cancer, we sought to evaluate text message responsivity across different types of text messages.

Methods: AYAs who recently completed treatment for cancer (n=26; mean age=16 years; 62% female, 16/26 participants) received 2-way text messages about survivorship health topics over a 16-week period. Using participants' text message log data, we coded responsivity to text messages and evaluated trends in responsivity to unprompted text messages and prompted text messages of varying content (eg, medication reminders, appointment reminders, and texts about personal experiences as a cancer survivor).

Results: Across prompted and unprompted text messages, responsivity rapidly decreased ($P \leq .001$ and $=.01$, respectively) and plateaued by the third week of the intervention. However, participants were more responsive to prompted text messages (mean responsivity=46% by week 16) than unprompted messages (mean responsivity=10% by week 16). They also demonstrated stable responsivity to certain prompted content: medication reminders, appointment reminders, goal motivation, goal progress, and patient experience texts.

Conclusions: Our methodology of evaluating text message responsivity revealed important patterns of engagement in a 2-way text message intervention for AYA cancer survivors.

(JMIR Mhealth Uhealth 2019;7(4):e12547) doi:10.2196/12547

KEYWORDS

mHealth; adolescents; young adults; cancer; chronic illness; self-management

Introduction

Background

Adolescents and young adults (AYAs) send and receive text messages more often than any other age demographic [1,2]. They also represent a cohort with an elevated risk for a number of maladaptive health behaviors, including poor disease self-management and nonadherence among those with chronic health conditions [3]. Mobile health (mHealth) interventions involving text messaging are particularly appealing to AYAs [4-6] and offer practical and scalable solutions for improving their health behaviors within real-world environments [7-9]. Indeed, text messaging interventions demonstrate modest but significant improvements in health knowledge and behaviors across AYAs with a variety of chronic health conditions [7,10]. Unfortunately, despite AYAs' enthusiasm about text messaging and the proliferation of text message interventions, their objective engagement with text messaging interventions is often low [11,12]. There is no simple formula for designing interactive and engaging text messages, and continuous user engagement or *stickiness* represents a pervasive challenge [13,14]. To promote sustained AYA engagement in text message interventions, further research is needed to determine whether patterns of text message responsivity vary across prompted, unprompted, and content-specific (such as medication reminders) text messages. Using a 2-way short message service (SMS) intervention with AYA survivors of childhood cancer, we capitalized on an opportunity to evaluate trends in responsivity to different types of text messages over a 16-week period.

To date, limited mHealth research has moved beyond efficacy evaluations to rigorously examine user engagement [15], such as text message responsivity. We define text message responsivity as instances in which a participant sends a response text after receiving a prompted (requests a response) or unprompted (does not explicitly request a response) text message from the research team. Outside of traditional health care settings, AYA engagement with text messages must contend with other aspects of their daily life and their fluctuating motivation to engage in purposeful health behaviors [15]. Many texting interventions require repeated exposure to health information and reminders to promote health behavior change. However, AYAs may habituate to this information and experience competing demands from other frequently used mobile apps, leading to disengagement with the intervention [16-18]. Further investigation into AYAs' responsivity to different types of text messages is critical, as the strength and durability of mHealth intervention effects are intricately connected to participant engagement [19-22].

Currently, the mHealth literature offers limited methodological guidance on how to analyze and optimize user engagement [15], including responsivity to text messages. Little attention has been paid to dismantling and evaluating text message components, such as the content that is most likely to elicit a text back from participants, which would deepen our understanding of how text messages function to improve health and well-being [15]. As such, to our knowledge, this paper represents the first to operationalize the term text message responsivity. One text message intervention for adolescents with type 1 diabetes (*Sweet Talk*) evaluated unprompted messages that were sent from participants to the research team [11]. In this study, participants were more likely to send certain types of unprompted text messages (eg, submission of blood glucose values) compared with others (eg, questions about their diabetes). These authors concluded that participants valued 2-way text message capabilities but were most motivated to submit disease monitoring information. A more fine-grained evaluation of responsivity to different types of prompted and unprompted text messages will help to further delineate the content that may engage (or reengage) the end user in an intervention over time.

AYA cancer survivors represent an exemplar group for SMS interventions. AYA survivors must continue lifelong follow-up care, establish preventative health behaviors, and adhere to medication regimens to manage a host of secondary morbidities of their curative cancer treatment (ie, late effects) [23,24]. At the same time, they are often transient during young adulthood and live great distances from their treating hospital, leaving them with difficulties with access to appropriate follow-up care [25]. As such, digital health interventions (eg, mHealth, Web-based, and social media) have been evaluated as methods of overcoming traditional barriers to AYA survivors' engagement in follow-up care and to promote healthy behaviors such as exercise [26,27]. There is a growing body of literature supporting the feasibility, acceptability, and initial efficacy of mHealth interventions in this population, including 1 study of a text message system that delivered survivorship information and resources [28] and a separate study of an app-based symptom management intervention [29]. Yet, engaging AYA cancer survivors can be challenging [30], and research has yet to thoroughly investigate their engagement in mHealth interventions [26].

Objectives

This paper addresses a literature gap by assessing AYA childhood cancer survivors' responsivity to specific text message content over the course of a 16-week 2-way pilot SMS intervention called *T*exting *H*ealth *R*esources to *I*nform, moti*V*ate, and *E*ngage (THRIVE). Consistent with past text messaging interventions, we expected a significant trend for decreasing responsivity over the course of the intervention. However, given that tailored text message interventions yield larger effect sizes than generic ones [31,32], we hypothesized that AYAs would demonstrate greater responsivity to prompted text messages that were tailored. By evaluating AYA responsivity to different types of text messages, we contribute to the literature by (1) illustrating a potentially generalizable method of analyzing objective text message engagement data and (2) providing guidance about maximizing the engagement of AYA childhood cancer survivors in an SMS intervention.

Methods

Intervention Development

THRIVE text messages were designed to support theory-informed categories of inform (eg, information about

health promotion, late effects, and resources), motivate (eg, providing encouragement and monitoring of health goal attainment), and engage (eg, engagement in follow-up care, sharing experiences, autonomy promotion, and psychosocial support). THRIVE was grounded in the Social-ecological Model of AYA Readiness to Transition to Adult Care (SMART) [33]. SMART is a validated model of AYA self-management that emphasizes multilevel influences on AYA self-management and transition to adult-centered care, including knowledge, self-efficacy, supportive relationships, and goals. The health belief model [34] and social cognitive theory [35] also influenced THRIVE text message development. Consistent with these health behavior theories, text messages were intended to enhance awareness of health vulnerability and importance of continued engagement in follow-up care and to motivate and reinforce positive health behaviors.

Text messages were developed by a research team comprising clinical psychologists, nurse practitioners, oncologists, and student (young adult age) research assistants and volunteers. The content, limited to 160 characters and written at a sixth-grade reading level, was similar to health-related information in the hospital-based AYA Survivor Handbook provided to controls. All team members reviewed text messages individually and in weekly team meetings for a period of approximately 2 months to ensure clarity of messages and consistency with influential theories and study aims. The pilot text messages were then sent to all team members for a trial period to test the system and further review the content.

This process resulted in the creation of 210 text messages that included content on healthy eating, exercise, sleep, sun safety, risk-taking behavior (drugs, alcohol, and sexual activity), academic and social life after cancer treatment, engagement in follow-up medical care, and connecting with other survivors and content relevant to 1 of the following health goals selected by the participant: (1) healthy eating, (2) smoking cessation, (3) reengage in school, (4) reengage in social activities, (5) increase physical activity, and (6) improve sleep/fatigue. Text messages were tailored by age or goal (23% of messages) and interactive (41%) in that they prompted the participant to text back a response to receive additional information, answer a survey item, or answer a trivia question; the remaining messages were unprompted and/or generic. Participants with ongoing medication regimens and upcoming clinic appointments also received weekly medication adherence texts and appointment reminder texts. Participants could also spontaneously send an unprompted text to the study team. A separate manuscript of our proof-of-concept randomized controlled trial describes intervention development in more detail (Schwartz, LA, unpublished data, February 2019). Participants who were randomized to THRIVE reported high acceptability.

Participants

This manuscript analyzed data from the THRIVE intervention group only. The inclusion criteria were as follows: (1) must be within 1 year of completing cancer treatment; (2) must be in cancer remission; (3) must be aged between 12 and 25 years; (4) must be able to read and speak English; and (5) for AYAs aged under 18 years, must have a parent/caregiver provide consent for participation. Of the 31 patients in the intervention group, 26 patients received text messages and were included in this analysis. Exclusion reasons included never turning on the phone/received text messages (n=3), relapsed on study (n=1), and determined history of nonmalignant, genetic tumor diagnosis (n=1).

Procedure

After obtaining institutional review board approval, eligible participants were identified using the cancer center's patient registry and upcoming clinic appointment schedules. Before recruitment, the patients' primary oncology providers were contacted to confirm eligibility. Participants were approached and invited to participate during outpatient oncology visits. After participants and caregivers (when applicable) provided informed consent and assent, participants were asked to select a health-related goal to pursue over the course of the study (eg, increase physical activity, improve healthy eating, and improve sleep/fatigue). Participants randomized to the intervention group were provided with an iPhone (an older version to reduce cost) and received 1 to 2 daily text messages over the course of 16 weeks. At the time of study initiation, our institution requested that we provide participants with a secured iPhone to ensure privacy and protect personal data (eg, cell phone numbers) from a third-party vendor. The frequency and duration of the text messages were informed by similar SMS interventions for AYAs with other chronic diseases [32,36,37]. Participants were able to select the time of the day to receive text messages and additional appointment reminders. Text messages were automatically sent by a tailored text message platform designed by an outside vendor (Reify Health), who was contracted to provide technical infrastructure and support. A research assistant monitored the delivery of the texts and incoming texts. Log data of 2-way text communication were securely stored and downloaded from the Reify Health platform.

Table 1. Text message examples by category.

Text message category	Example message
Prompted messages	
Trivia[a]	Does smoking help you lose weight? Text back 1 for yes or 2 for no.
More information[a]	Losing your hair during treatment is hard for most patients with cancer. If you lost hair, text back 1 for info on your hair after cancer.
Patient experience[a]	Your treatment & some medications can make skin more sensitive to sun than it used to be. Did you apply sunscreen today? Text back 1 if yes, or 2 if no.
Goal-tailored	Think about your physical activity level now compared to 3 months ago - text back if you feel 1 better, 2 the same, or 3 worse
Age-tailored	Managing the healthcare system & learning about insurance can be tricky. Text 1 to learn more about health insurance as a young adult & survivor.
Goal motivation[a]	How motivated are you to be more active? Text back: 1 - Not at all, 2 - Slightly, 3 - Somewhat, 4 - Very, 5 - Extremely
Goal progress[a]	How much progress have you made on your goal to be more active? Text back: 1-None, 2-small amount, 3-moderate amount, 4-a lot, 5-huge amount
Medication reminder[a]	Have you been taking your medication? Text 0 - I am not taking medication anymore, 1 - No, not really, 2 - sometimes, 3 - always.
Appointment reminder[a]	You have an appointment to come to CHOP in 2 days, on 11/28/2014 at 3:00PM. Do you know how you are getting there? Text back 1 for Yes or 2 for No.
Unprompted messages	
Information	Feeling tired right after treatment is totally normal. Healing takes energy & time. Make sure to listen to your body & pace yourself throughout the day.
Goal-tailored	Put physical activity in your schedule - setting aside a specific time of the day will help prevent you from putting it off until later.
Age-tailored	Feeling connected when you go back to school can be tough. Join a club or sport that interests you to find peers who enjoy similar things.

[a]Analyzed responsivity to this type of prompted text message.

Data Analytic Plan

We exported SMS raw log data for each participant in Microsoft Excel format, which delineated the date, time, and content of every text message the participant received from the research team and every message the participant sent to the research team. Moreover, 2 trained study staff coded responsivity to text messages by week, that is, how many times each participant responded to a text message by text message category. After coding the text message responsivity, for each participant at each week, we calculated the overall percentage responsivity (ie, a continuous variable of the total number of responses to text messages divided by the total number of texts received). Similarly, for each participant at each week, we calculated percentage responsivity to prompted (ie, messages that requested a response) and unprompted (ie, messages that did not request a response) text messages. We further analyzed select prompted text messages about (1) medication reminders (as applicable), (2) appointment reminders (as applicable), (3) goal motivation, (4) goal progress, (5) health knowledge *trivia*, (6) patient experiences, and (7) requests for *more information* (see Table 1). We did not analyze goal-tailored or age-tailored texts as the prompted text messages in these categories were accounted for in patient experience and more information categories (ie, these categories were not mutually exclusive).

For text message responsivity outcomes measured weekly (overall, prompted, unprompted, and more information text messages), we constructed a longitudinal piecewise linear regression model, with the participant percentage responsivity at each week as the outcome. The piecewise model assumed 1 slope from weeks 1 to 3 and another slope after week 3. We specified this piecewise model based on descriptive data and graphs about responsivity over time, which suggest a rapid decrease in responsivity in weeks 1 to 3 and attenuated decrease afterward. Testing the significance of each slope suggests whether there was a significant decreasing or increasing trend of text message responsivity in each of the 2 periods. We also tested the significance of difference in the slopes for after versus before week 3 to evaluate if a reduced model of the longitudinal linear model should be used. The generalized estimation equation method with exchangeable correlation structure was used to account for the potential within-subject correlations among the repeated outcomes over weeks. For the other responsivity outcomes with less frequent measures, longitudinal linear regression models were constructed to test the significance of the slope in terms of whether there was a significant decreasing or increasing trend of text message responsivity over time. All analyses were performed in SAS 9.4.

Results

Demographic Information

Means, SDs, ranges, and percentages for demographic and medical information are presented in Table 2. The average age of participants was 16 years and the majority had completed cancer treatment for a leukemia or lymphoma diagnosis (14/26 participants, 54%). On average, participants had completed cancer therapy 5 months before their enrollment in the study.

Overall Responsivity

Means and SDs of percentage responsivity by week for each type of text message (ie, overall, prompted, unprompted, and prompted content) are provided in Table 3. When evaluating text message responsivity across all types of prompted and unprompted text messages, there was a rapid decrease in responsivity in the first 3 weeks of the intervention, with a 7% decrease per week (95% CI −10.7 to −3.0; *P*<.001), and a much attenuated decrease after week 3, with a 0.4% decrease per week (95% CI −0.7 to −0.1; *P*=.003; Figure 1). The slopes before and after week 3 were significantly different (*P*=.002), supporting the use of the piecewise linear model over the linear model. Descriptively, at week 1, the mean responsivity across all text messages was 36% (SD 22.4%), and by week 16, the mean responsivity decreased to 17% (SD 17.9%; see Table 3).

Responsivity to Prompted and Unprompted Messages

In terms of responsivity to prompted text messages, there was again a significant and rapid decrease in responsivity in the first 3 weeks of the intervention, with a 13% decrease per week (95% CI −19.5 to −6.3; *P*<.001), which plateaued after week 3 (decrease 0.1% per week; 95% CI −0.8 to 0.6; *P*=.79; see Figure 1). The slopes before and after week 3 were significantly different (*P*<.001). Participants demonstrated relatively high engagement with prompted text messages at week 1 but with significant variability (mean responsivity=78%, SD 31%). By week 16, the average responsivity to prompted text messages decreased to 46% (SD 37%; see Table 3).

Similarly, there was a rapid decrease in responsivity to unprompted text messages in the first 3 weeks, with a 7% decrease per week (95% CI −11.9 to −1.5; *P*=.01), and an attenuated decrease after week 3, with 0.4% decrease per week (95% CI −0.6 to −0.1; *P*=.01; see Figure 1). The slopes before and after week 3 were significantly different (*P*=.02). The mean responsivity at week 1 was 28% (SD 29.3%), which decreased to 10% by week 16 (SD 26.3%; see Table 3).

Table 2. Demographic and disease information (N=26).

Variable	Statistics, n (%)	Mean (SD)	Minimum-maximum
Current age (years)	26 (100)	16.42 (2.87)	12.00-20.00
Age at cancer diagnosis (years)	26 (100)	15.48 (3.01)	8.68-19.05
Time off treatment (months)	26 (100)	5.18 (3.58)	0.39-11.96
Gender (female)	16 (62)	—[a]	—
Race			
White	17 (65)	—	—
African American	6 (23)	—	—
Other	2 (8)	—	—
Asian	1 (4)	—	—
Ethnicity, Hispanic	3 (12)	—	—
First cancer type			
Leukemia/lymphoma	14 (54)	—	—
Solid tumor	7 (27)	—	—
Brain tumor	5 (19)	—	—
Had relapse	2 (8)	—	—
Had second cancer	1 (4)	—	—

[a]For categorical variables, mean, SD, and minimum-maximum have not been listed.

Table 3. Means and SDs for text message responsivity (%) across categories of text messages received weekly during the intervention.

Week	Total, n; mean (SD)	Unprompted, n; mean (SD)	Prompted, n; mean (SD)	More information, n; mean (SD)
Week 1	26; 36% (22%)	25; 28% (29%)	26; 78% (31%)	24; 71% (44%)
Week 2	26; 32% (18%)	26; 13% (20%)	26; 65% (36%)	25; 54% (43%)
Week 3	26; 17% (23%)	26; 14% (25%)	26; 50% (51%)	25; 52% (51%)
Week 4	26; 23% (18%)	26; 9% (20%)	26; 53% (36%)	26; 44% (43%)
Week 5	26; 25% (23%)	26; 12% (26%)	26; 51% (42%)	26; 46% (45%)
Week 6	26; 21% (23%)	26; 12% (26%)	26; 42% (40%)	26; 35% (42%)
Week 7	26; 22% (23%)	26; 9% (23%)	26; 54% (47%)	26; 51% (43%)
Week 8	26; 21% (24%)	26; 10% (27%)	26; 65% (49%)	26; 58% (50%)
Week 9	26; 21% (20%)	26; 8% (22%)	26; 48% (40%)	26; 47% (39%)
Week 10	26; 21% (18%)	26; 7% (17%)	26; 58% (40%)	25; 48% (49%)
Week 11	26; 22% (20%)	26; 9% (20%)	26; 55% (42%)	26; 47% (48%)
Week 12	26; 21% (18%)	26; 6% (18%)	26; 54% (32%)	26; 35% (49%)
Week 13	26; 17% (20%)	26; 10% (27%)	25; 48% (46%)	24; 48% (45%)
Week 14	26; 20% (18%)	25; 5% (13%)	26; 45% (39%)	26; 41% (40%)
Week 15	25; 16% (18%)	25; 6% (19%)	25; 56% (49%)	12; 42% (51%)
Week 16	25; 17% (18%)	25; 10% (26%)	25; 46% (37%)	25; 38% (39%)

Figure 1. Average responsivity across text messages categories received weekly over a 16-week period.

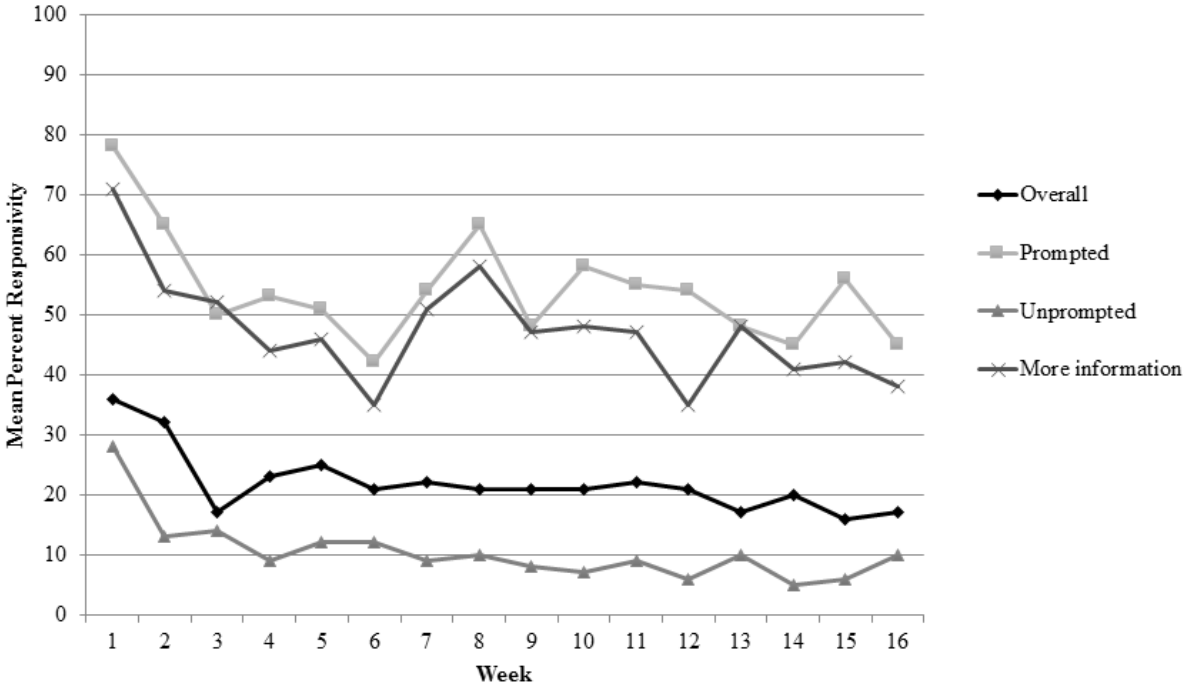


Table 4. Means and SDs for text message responsivity (%) across categories of text messages received intermittently during the intervention.

Week	Medication reminders, n; mean (SD)	Appointment reminders, n; mean (SD)	Goal motivation, n; mean (SD)	Goal progress, n; mean (SD)	Patient experience, n; mean (SD)	Trivia, n; mean (SD)[a]
Week 1	17; 78% (44%)	__[b]	—	—	—	26; 89% (33%)
Week 2	—	—	26; 73% (45%)	—	25; 80% (41%)	—
Week 3	—	—	—	—	—	—
Week 4	17; 67% (50%)	2; 50% (71%)	—	26; 69% (47%)	—	—
Week 5	—	2; 50% (71%)	—	—	26; 65% (49%)	—
Week 6	—	1; 0% (0%)	26; 62% (50%)	—	1; 100% (0%)	—
Week 7	—	2; 50% (71%)	—	—	8; 50% (53%)	—
Week 8	17; 78% (44%)	1; 100% (0%)	—	26; 62% (50%)	—	—
Week 9	—	—	—	—	—	—
Week 10	—	1; 0% (0%)	26; 65% (49%)	—	9; 78% (44%)	—
Week 11	—	5; 60% (55%)	—	—	—	26; 62% (50%)
Week 12	17; 67% (50%)	6; 58% (49%)	—	26; 54% (51%)	12; 92% (29%)	—
Week 13	—	5; 70% (45%)	—	—	1; 0% (0%)	—
Week 14	—	4; 50% (58%)	25; 60% (50%)	—	1; 100% (0%)	—
Week 15	—	4; 50% (58%)	—	—	—	25; 52% (51%)
Week 16	17; 67% (50%)	5; 50% (50%)	—	14; 57% (51%)	—	—

[a]Significant linear decrease across weeks.

[b]Not applicable.

Responsivity to Promoted Content-Specific Messages

We also examined trends in responsivity across certain types of prompted text messages (see Table 1). Participant responsivity to medication reminders (P=.64), appointment reminders (P=.31), goal motivation (P=.12), goal progress (P=.25), and patient experience texts (P=.74) did not significantly change over the course of the 16-week intervention, suggesting relatively stable participant engagement with these types of text messages (see Table 4). In contrast, there was a significant decrease in responsivity to *more information* texts in the first 3 weeks, with a 10% decrease per week (95% CI −17.2 to −1.8; P=.02), which plateaued after week 3 (decrease 0.5% per week; 95% CI −1.4 to 0.5; P=.33; see Figure 1). The slopes before and after week 3 were significantly different (P=.03). Descriptively, the average responsivity to *more information* texts at week 1 was 71%, which decreased to 38% by week 16 (see Table 3). Participants also responded to fewer health trivia texts each week, with a 3% decrease per week (95% CI −4.24 to −0.97; Z=−3.13; P=.002). Mean responsivity for trivia texts was 89% at week 1, which decreased to 52% by the last week of the intervention.

Discussion

Principal Findings

Over the course of a 16-week 2-way SMS intervention for AYA cancer survivors called THRIVE, we found that responsivity to text messages (overall, prompted, unprompted, and *more information* texts) peaked during the first week and rapidly decreased by week 3. However, participants were more responsive to prompted than to unprompted text messages and demonstrated stable responsivity to certain prompted content: medication reminders, appointment reminders, goal motivation, goal progress, and patient experience texts. Our analysis of text message responsivity represents an important contribution to the literature, as the majority of existing mHealth interventions have focused exclusively on efficacy data and neglected objective engagement data [38]. Our approach, involving coding how many times a participant sent a *text back* to different types of text messages, was practical, feasible to do with a small research team, and unveiled important patterns in AYA survivors' text message responsivity.

The decline in text message responsivity is consistent with past research that has demonstrated that engagement with mHealth tools is generally low and rapidly decreases in the first few weeks of use [11,12]. For example, in a mHealth intervention called HeartSteps that delivered contextually tailored activity suggestions to sedentary adults, the effect of activity suggestions diminished over time and largely disappeared after 1 month [39]. Similarly, 74% of consumers report discontinuing the use of commercial health apps after the tenth use [40]. This study adds to this body of literature and contributes novel information about how text message responsivity declined and plateaued in an AYA cancer survivor sample, likely because of habituation. At the same time, across weeks, participants demonstrated the highest percentage responsivity to prompted text messages. Among the prompted text message content categories, participants demonstrated relatively stable engagement (generally >50%) with certain content: medication reminders,

appointment reminders, goal motivation, goal progress, and patient experience texts.

The data generated by this study have furthered our understanding of AYA survivors' objective engagement with specific text message content in a few ways. First, deteriorating responsivity in the first 3 weeks of the intervention highlights the importance of planning for habituation in future trials. To sustain long-term engagement with text messages, researchers should consider the implementation of a text message delivery schedule that minimizes the risk of habituation by favoring times when participants are most likely to be receptive and available to respond [39]. Although each participant in our study selected the timing of their text messages, employing reinforcement learning algorithms [41] would help to further personalize the timing of text message delivery based on a participant's prior responses (eg, their responsivity to messages at certain times, days of the week, and based on certain content). Notably, sending 1 to 2 health-related text messages per day for 16 weeks may have not been ideal for establishing and sustaining the level of responsivity needed to support meaningful behavior change. Indeed, based on learning theories, it may have been beneficial to temporarily suspend text messages when a participant's responsivity decreased (eg, at week 3 of our intervention) to reduce burden and encourage a spontaneous recovery in engagement once text messages were reinitiated [42]. Similarly, it may have been useful to vary text messages by sending engaging (noninterventional) content on some days, such as memes, gifs, and life insights that have been used to increase engagement in daily mobile assessments in other AYA populations [43].

Second, although greater responsivity to prompted text messages is somewhat intuitive as these messages requested a response, relatively few 2-way text message interventions (as opposed to 1-way text messages) have been tested to address AYA health behaviors [10] and even fewer for AYA cancer survivors [26]. Our research provides evidence that bidirectional messaging may be an important component for promoting AYA survivors' engagement with text message interventions.

Third, AYA survivors demonstrated sustained responsivity to prompted texts about their temporal health behaviors (eg, medication adherence), a personally selected health goal, and personal experiences compared with text messages that were education based (ie, seeking more information or responding to health trivia). Such findings are consistent with evidence that tailored, personally relevant text messages are more engaging that generic text messages [11,27]. Although it is well documented that educational interventions alone are insufficient to improve health behaviors such as medical adherence [44,45], our findings show that purely informational content may have also been less engaging for AYAs in our research study. Alternatively, it is possible that participants were more engaged with these content-specific text messages because they were received less often, appeared more novel, and thus were less prone to habituation. Experimental designs such as microrandomized trials will help to disentangle how the frequency, timing, and content of tailored text messages impact AYA survivors' proximal engagement in mHealth interventions [46,47].

Limitations

With limited methodological guidance about how to best evaluate text message user data, we recognize that our approach represents only one of the several possible methods of examining text message responsivity. This approach is not without limitations. Our texting platform was not equipped to measure whether participants read the text messages. As a result, we relied exclusively on a text back from participants to determine their engagement, which neglects the possibility that participants passively engaged with content (especially with unprompted messages that did not request a response). To our knowledge, no research has examined differences in intervention efficacy for participants who actively respond to text messages compared with those who passively read, but did not respond, to text messages. Such research would be beneficial for illustrating dose-response relationships within a texting intervention or the amount of responsivity that is minimally needed to experience certain intervention effects [38,48]. In addition, all text messages were delivered via a secondary study iPhone, which may have resulted in lower responsivity than if messages were delivered to a personal phone.

Another notable limitation was the large variability across participants in their responsivity to text messages, as evidenced by large SDs in most text message categories. Although our focus was assessing responsivity to text messages at the group level, heterogeneity between participants highlights the importance of examining within-subject responsivity in future research. Furthermore, our investigation was underpowered to test whether specific participant characteristics (such as age, sex, or cancer diagnosis) predicted trends in text message responsivity, which could provide further insight into responsivity patterns. More research, with larger samples, is clearly needed to demonstrate whether our method of examining text message responsivity contributes meaningful knowledge about user engagement and explanatory data about the efficacy of the intervention. Despite these limitations, this generalizable approach unveiled practical information about how to enhance our text bank for AYA cancer survivors before a scaled-up randomized controlled trial (ie, increase the number of prompted text messages and expand tailoring of text messages to increase personal relevance).

Conclusions and Future Directions

This study illustrates 1 method for analyzing user engagement with 2-way text messages and contributes knowledge about the types of text messages that AYA survivors responded to the most. The analysis of participant responsivity to text messages, as well as other mHealth log data (eg, such as opening a mobile app, clicking a link, and reading a text message), can help provide critical insight into participant engagement with mHealth technology [15]. Consistent with prior research [11,28,39], our findings showed that AYA survivors' responsivity to text messages rapidly decreased during the first few weeks of the intervention, but they demonstrated higher engagement with prompted and personally relevant text messages. Future mHealth interventions should integrate bidirectional and tailored content to maximize AYA survivors' engagement in text message interventions. For example,

following this pilot intervention, we expanded THRIVE text messages to include educational content that was tailored to each AYA survivor's cancer treatment history and related risks for late effects, for example:

> *[Name], Ever notice that you have trouble hearing the TV or other people at large gatherings? If so, let a member of your medical team know and ask to get your hearing checked. Cisplatin (a chemotherapy you received) can affect your hearing.*

The second wave of this intervention is described in a separate publication [49] and is currently being tested in a randomized controlled trial.

Other intervention features may have increased responsivity, such as delivering text messages directly to an AYA's personal phone and adding gamification elements (eg, earning points for responding to messages) [43]. Future research can systematically test the causal effects of various message delivery schedules, as well as different engagement strategies (eg, gamification and incentives), on increasing an individual's responsivity to text messages over time [43,50]. In addition to analyzing responsivity to specific text message content, analysis of additional text message components, such as who sent the text message, when was it delivered, and how frequently were text messages sent, will help elucidate salient user patterns and preferences [15]. We recommend that future research make use of text message log data when possible and examine both intervention (eg, length of intervention, types of text messages, gamification, and target health behavior) and contextual factors (eg, participant age, sex, race/ethnicity, and health status) that influence engagement [51,52]. Future research should also evaluate the latency of responsivity, such as the elapsed time to respond to text messages, as studying response times has yielded valuable information in other health areas (eg, factors that influence health providers' response times to patient physiological monitors) [53]. Building and sharing knowledge in this area can encourage additional research on engagement with mHealth text messages; inform important modifications to text banks; and answer multiple calls for rigorous methodological research in the development, implementation, and evaluation of user-centered mHealth tools [12,54,55].

Acknowledgments

This work was supported by a Department of Pediatrics Chair's Initiative from The Children's Hospital of Philadelphia to LAS. A postdoctoral fellowship from the American Cancer Society (PF-16-166-01-CPPB) to AMP supported the writing of this manuscript. The authors are grateful to the childhood cancer survivors who participated. The authors would also like to thank their text messaging vendor Reify Health.

Conflicts of Interest

None declared.

References

1. Lenhart A, Purcell K, Smith A, Zickuhr K. Pew Research Center. 2010. Social media & mobile internet use among teens and young adults URL: http://www.pewinternet.org/2010/02/03/social-media-and-young-adults/ [accessed 2019-03-11] [WebCite Cache ID 76nTkEbQq]
2. Anderson M, Jingjing J. Pew Research Center. 2018. Teens, social media & technology overview 2018 URL: http://www.pewinternet.org/wp-content/uploads/sites/9/2018/05/PI_2018.05.31_TeensTech_FINAL.pdf [accessed 2019-03-11] [WebCite Cache ID 76nTmaeLM]
3. Rapoff, MA. Adherence To Pediatric Medical Regimens (Issues In Clinical Child Psychology). New York, NY: Springer; 2010.
4. Gold J, Lim MS, Hellard ME, Hocking JS, Keogh L. What's in a message? Delivering sexual health promotion to young people in Australia via text messaging. BMC Public Health 2010;10:792 [FREE Full text] [doi: 10.1186/1471-2458-10-792] [Medline: 21190584]
5. Rabin C, Simpson N, Morrow K, Pinto B. Intervention format and delivery preferences among young adult cancer survivors. Int J Behav Med 2013 Jun;20(2):304-310. [doi: 10.1007/s12529-012-9227-4] [Medline: 22328444]
6. Hanauer DA, Wentzell K, Laffel N, Laffel LM. Computerized Automated Reminder Diabetes System (CARDS): e-mail and SMS cell phone text messaging reminders to support diabetes management. Diabetes Technol Ther 2009 Feb;11(2):99-106. [doi: 10.1089/dia.2008.0022] [Medline: 19848576]
7. Militello LK, Kelly SA, Melnyk BM. Systematic review of text-messaging interventions to promote healthy behaviors in pediatric and adolescent populations: implications for clinical practice and research. Worldviews Evid Based Nurs 2012 Apr;9(2):66-77. [doi: 10.1111/j.1741-6787.2011.00239.x] [Medline: 22268959]
8. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. JAMA Pediatr 2017 May 1;171(5):461-469. [doi: 10.1001/jamapediatrics.2017.0042] [Medline: 28319239]
9. Eapen ZJ, Peterson ED. Can mobile health applications facilitate meaningful behavior change?: time for answers. JAMA 2015;314(12):1236-1237. [doi: 10.1001/jama.2015.11067] [Medline: 26393844]

10. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text messaging and mobile phone apps as interventions to improve adherence in adolescents with chronic health conditions: a systematic review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]
11. Franklin VL, Greene A, Waller A, Greene SA, Pagliari C. Patients' engagement with "Sweet Talk" - a text messaging support system for young people with diabetes. J Med Internet Res 2008 Jun 30;10(2):e20 [FREE Full text] [doi: 10.2196/jmir.962] [Medline: 18653444]
12. Wu YP, Steele RG, Connelly MA, Palermo TM, Ritterband LM. Commentary: pediatric eHealth interventions: common challenges during development, implementation, and dissemination. J Pediatr Psychol 2014 Jul;39(6):612-623 [FREE Full text] [doi: 10.1093/jpepsy/jsu022] [Medline: 24816766]
13. Buller DB, Berwick M, Lantz K, Buller MK, Shane J, Kane I, et al. Smartphone mobile application delivering personalized, real-time sun protection advice: a randomized clinical trial. JAMA Dermatol 2015 May;151(5):497-504 [FREE Full text] [doi: 10.1001/jamadermatol.2014.3889] [Medline: 25629710]
14. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
15. Muench F, Baumel A. More than a text message: dismantling digital triggers to curate behavior change in patient-centered health interventions. J Med Internet Res 2017 May 26;19(5):e147 [FREE Full text] [doi: 10.2196/jmir.7463] [Medline: 28550001]
16. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
17. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014;9(2):e88166 [FREE Full text] [doi: 10.1371/journal.pone.0088166] [Medline: 24505411]
18. Tao D, Xie L, Wang T, Wang T. A meta-analysis of the use of electronic reminders for patient adherence to medication in chronic disease care. J Telemed Telecare 2015 Jan;21(1):3-13. [doi: 10.1177/1357633X14541041] [Medline: 25147178]
19. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
20. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
21. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
22. Enwald HP, Huotari MA. Preventing the obesity epidemic by second generation tailored health communication: an interdisciplinary review. J Med Internet Res 2010;12(2):e24 [FREE Full text] [doi: 10.2196/jmir.1409] [Medline: 20584698]
23. Kopp LM, Gastelum Z, Guerrero CH, Howe CL, Hingorani P, Hingle M. Lifestyle behavior interventions delivered using technology in childhood, adolescent, and young adult cancer survivors: A systematic review. Pediatr Blood Cancer 2017 Jan;64(1):13-17. [doi: 10.1002/pbc.26166] [Medline: 27468131]
24. Davis SW, Oakley-Girvan I. Achieving value in mobile health applications for cancer survivors. J Cancer Surviv 2017 Aug;11(4):498-504. [doi: 10.1007/s11764-017-0608-1] [Medline: 28342093]
25. Schwartz L, Szalda D, Hobbie W, Kazak A. Predictors of engagement in follow-up care of young adult survivors of childhood cancer. Pediatr Blood Cancer 2010;55(5):847.
26. Devine K, Viola A, Coups E, Wu Y. Digital health interventions for adolescent and young adult cancer survivors. JCO Clin Cancer Inform 2018 Dec;2:1-15 [FREE Full text] [doi: 10.1200/CCI.17.00138] [Medline: 30652583]
27. Mendoza JA, Baker KS, Moreno MA, Whitlock K, Abbey-Lambertz M, Waite A, et al. A Fitbit and Facebook mHealth intervention for promoting physical activity among adolescent and young adult childhood cancer survivors: a pilot study. Pediatr Blood Cancer 2017 Dec;64(12). [doi: 10.1002/pbc.26660] [Medline: 28618158]
28. Casillas J, Goyal A, Bryman J, Alquaddoomi F, Ganz PA, Lidington E, et al. Development of a text messaging system to improve receipt of survivorship care in adolescent and young adult survivors of childhood cancer. J Cancer Surviv 2017 Aug;11(4):505-516 [FREE Full text] [doi: 10.1007/s11764-017-0609-0] [Medline: 28364263]
29. Rodgers CC, Krance R, Street RL, Hockenberry MJ. Feasibility of a symptom management intervention for adolescents recovering from a hematopoietic stem cell transplant. Cancer Nurs 2013;36(5):394-399 [FREE Full text] [doi: 10.1097/NCC.0b013e31829629b5] [Medline: 23842522]
30. Rabin C, Horowitz S, Marcus B. Recruiting young adult cancer survivors for behavioral research. J Clin Psychol Med Settings 2013 Mar;20(1):33-36 [FREE Full text] [doi: 10.1007/s10880-012-9317-0] [Medline: 22810954]
31. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]

32. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
33. Schwartz LA, Brumley LD, Tuchman LK, Barakat LP, Hobbie WL, Ginsberg JP, et al. Stakeholder validation of a model of readiness for transition to adult care. JAMA Pediatr 2013 Oct;167(10):939-946 [FREE Full text] [doi: 10.1001/jamapediatrics.2013.2223] [Medline: 23959392]
34. Strecher V, Rosenstock I. The health belief model. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research and Practice. San Francisco: Jossey-Bass; 1997:4-117.
35. Bandura A. Health promotion from the perspective of social cognitive theory. Psychol Health 1998 Jul;13(4):623-649 [FREE Full text] [doi: 10.1080/08870449808407422]
36. Ostojic V, Cvoriscec B, Ostojic SB, Reznikoff D, Stipic-Markovic A, Tudjman Z. Improving asthma control through telemedicine: a study of short-message service. Telemed J E Health 2005 Feb;11(1):28-35. [doi: 10.1089/tmj.2005.11.28] [Medline: 15785218]
37. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]
38. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR Res Protoc 2017 Aug 7;6(8):e156 [FREE Full text] [doi: 10.2196/resprot.6452] [Medline: 28784592]
39. Klasnja P, Smith S, Seewald N, Lee A, Hall K, Luers B, et al. Effects of contextually-tailored suggestions for physical activity: the HeartSteps micro-randomized trial. Ann Behav Med 2018 Sep 5. [doi: 10.1093/abm/kay067] [Medline: 30192907]
40. Consumer Health Information Corporation. 2012. Motivating patients to use smartphone health apps URL: http://www.consumer-health.com/motivating-patients-to-use-smartphone-health-apps/ [accessed 2019-03-11] [WebCite Cache ID 76nUDq8Qe]
41. Shortreed S, Laber E, Lizotte D, Stroup T, Pineau J, Murphy S. Informing sequential clinical decision-making through reinforcement learning: an empirical study. Mach Learn 2011 Jul 1;84(1-2):109-136 [FREE Full text] [doi: 10.1007/s10994-010-5229-0] [Medline: 21799585]
42. Rankin C, Abrams T, Barry R, Bhatnagar S, Clayton D, Colombo J, et al. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. Neurobiol Learn Mem 2009 Sep;92(2):135-138 [FREE Full text] [doi: 10.1016/j.nlm.2008.09.012] [Medline: 18854219]
43. Rabbi M, Philyaw Kotov M, Cunningham R, Bonar EE, Nahum-Shani I, Klasnja P, et al. Toward increasing engagement in substance use data collection: development of the Substance Abuse Research Assistant app and protocol for a microrandomized trial using adolescents and emerging adults. JMIR Res Protoc 2018 Jul 18;7(7):e166. [doi: 10.2196/resprot.9850] [Medline: 30021714]
44. Duncan C, Mentrikoski J, Wu Y, Fredericks E. Practice-based approach to assessing and treating non-adherence in pediatric regimens. Clin Pract Pediatr Psychol 2014 Sep;2(3):322-336. [doi: 10.1037/cpp0000066] [Medline: 25506046]
45. Kahana S, Drotar D, Frazier T. Meta-analysis of psychological interventions to promote adherence to treatment in pediatric chronic health conditions. J Pediatr Psychol 2008 Jul;33(6):590-611 [FREE Full text] [doi: 10.1093/jpepsy/jsm128] [Medline: 18192300]
46. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. JMIR Mhealth Uhealth 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: 10.2196/10123] [Medline: 30497999]
47. Smith SN, Lee AJ, Hall K, Seewald NJ, Boruvka A, Murphy SA, et al. Design lessons from a micro-randomized pilot study in mobile health. In: Rehg JM, Murphy SA, Kumar S, editors. Mobile Health: Sensors, Analytic Methods, And Applications. Cham, Switzerland: Springer; 2019:59-82.
48. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
49. Schwartz L, Psihogios A, Henry-Moss D, Daniel L, Ver Hoeve E, Velazquez-Martin B, et al. Iterative development of a mHealth intervention for adolescent and young adult survivors of childhood cancer. Clinical Practice in Pediatric Psychology 2019;7(1):31-43 [FREE Full text] [doi: 10.1037/cpp0000272]
50. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. Health Psychol 2015 Dec;34(Suppl):1220-1228. [doi: 10.1037/hea0000305] [Medline: 26651463]
51. 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 2016 Dec 13. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]

52. Paz CR, Haug S, Filler A, Kowatsch T, Schaub MP. Engagement within a mobile phone-based smoking cessation intervention for adolescents and its association with participant characteristics and outcomes. J Med Internet Res 2017 Nov 1;19(11):e356 [FREE Full text] [doi: 10.2196/jmir.7928] [Medline: 29092811]
53. Bonafide CP, Localio AR, Holmes JH, Nadkarni VM, Stemler S, MacMurchy M, et al. Video analysis of factors associated with response time to physiologic monitor alarms in a children's hospital. JAMA Pediatr 2017 Jun 1;171(6):524-531 [FREE Full text] [doi: 10.1001/jamapediatrics.2016.5123] [Medline: 28394995]
54. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
55. Patel S, Arya M. The BUS Framework: a comprehensive tool in creating an mHealth app utilizing behavior change theories, user-centered design, and social marketing. J Mob Technol Med 2017 Apr;6(1):39-45 [FREE Full text] [doi: 10.7309/jmtm.6.1.6] [Medline: 29104674]

Abbreviations

AYAs: adolescents and young adults
mHealth: mobile health
SMART: Social-ecological Model of AYA Readiness to Transition to Adult Care
SMS: short message service
THRIVE: Texting Health Resources to Inform, motiVate, and Engage

Edited by L Buis; submitted 18.10.18; peer-reviewed by C Cushing, B Holtz; comments to author 20.11.18; revised version received 15.01.19; accepted 05.02.19; published 18.04.19.

Please cite as:
Psihogios AM, Li Y, Butler E, Hamilton J, Daniel LC, Barakat LP, Bonafide CP, Schwartz LA
Text Message Responsivity in a 2-Way Short Message Service Pilot Intervention With Adolescent and Young Adult Survivors of Cancer
JMIR Mhealth Uhealth 2019;7(4):e12547
URL: http://mhealth.jmir.org/2019/4/e12547/
doi:10.2196/12547
PMID:30998225

Original Paper

Translating/Creating a Culturally Responsive Spanish-Language Mobile App for Visit Preparation: Case Study of "Trans-Creation"

Denise Ruvalcaba[1], BA; Hidemi Nagao Peck[2], MA, MPH; Courtney Lyles[3,4], PhD; Connie S Uratsu[1], MS; Patricia R Escobar[2], MPH; Richard W Grant[1], MPH, MD

[1]Division of Research, Kaiser Permanente Northern California, Oakland, CA, United States

[2]Regional Health Education, The Permanente Medical Group, Oakland, CA, United States

[3]Division of General Internal Medicine, Zuckerberg San Francisco General Hospital and Trauma Center, University of California, San Francisco, CA, United States

[4]Center for Vulnerable Populations, University of California, San Francisco, CA, United States

Corresponding Author:
Richard W Grant, MPH, MD
Division of Research
Kaiser Permanente Northern California
2000 Broadway
Oakland, CA, 94612
United States
Phone: 1 510 891 5960
Email: Richard.W.Grant@kp.org

Abstract

Background: Health information technology (IT) tools are increasingly used to improve patient care. However, implementation of English-only health IT tools could potentially worsen health disparities for non-English speakers.

Objective: We aim to describe the "trans-creation" process of developing linguistically and culturally appropriate health IT tools through a detailed case analysis of a waiting room health mobile app designed to help Spanish-speaking Latino people prepare for primary care visits.

Methods: We adapted the English-language Visit Planner mobile app for Spanish-speaking Latino patients. We applied culturally defined themes derived from prior published research and input by both skilled linguists and potential end users. Initial changes were iteratively reviewed and edited by a team of writers, health care educators, subject matter experts, patients, and providers.

Results: The trans-creation process resulted in the following key culturally mediated changes to the tool: replacing the "provider" actors with "patient" actors; changing the choice of "Stress at Home or Work" (represented by an icon of a house) to "Mi Familia" (translation: my family; icon is an outline of family members holding hands); replacing the English terms "anxiety" and "depression" with "Me siento desanimado"(translation: I am feeling down) to avoid mental health stigma; and using more concise text translation to ensure the wording fit the available on-screen space.

Conclusions: The trans-creation process of cultural and linguistic adaptation led to several design changes that would not have been implemented if we had simply translated the words from English to Spanish.

(JMIR Mhealth Uhealth 2019;7(4):e12457) doi:10.2196/12457

KEYWORDS

IT tool development; transcreation; doctor-patient communication; primary care

Introduction

Latino people are the fastest and largest growing minority group in the United States and are projected to account for one-third of the US population by 2060 [1]. Latino people face high rates of chronic conditions like heart disease, diabetes, and cancer [2]. In 2010, the Centers for Disease Control and Prevention reported that 14.2% of Latinos had two or more concurrent chronic conditions [3]. Given this high prevalence of chronic illness and the projected growth of the Latino population, increased efforts are needed to address how to improve health outcomes for this community.

Effective patient-provider communication is the cornerstone of good health care [4,5]. Patients with diabetes who report good

provider-patient communication, for example, have been shown to have better self-care and glycemic control [6,7]. Language barriers represent an important source of health disparities among Latino people [8]. Indeed, communication between English-speaking doctors and Latino patients may often be hampered because 30%-40% of the Latino population speaks English less than "very well" (often defined as limited English proficiency) [9]. Beyond simple language discrepancy, culture also plays a key role in the health of Latino people, because there are culture-specific understandings of illness including stigma associated with certain conditions such as mental health [10]. Research has identified several core cultural constructs such as *familismo* (the central role of family) and *confianza* (importance placed on trusting health care providers) that are particularly relevant to health-related interventions [11,12]. Effective communication with Latino patients must therefore also consider the cultural framework within which Latino people experience their health care.

Health information technology (IT) tools are increasingly being used within health care settings and have been proven to enhance patient-provider communication and lead to better health outcomes [13,14]. However, English-only health IT tools could potentially increase health disparities if they only benefit English speakers. Therefore, developing culturally responsive Spanish-language health IT tools is imperative. As part of a funded project (PCORI/CDR-1403-11992), we previously developed an English-language health mobile app (Visit Planner) to help patients prepare for their primary care visits while in the waiting room [15]. To adapt our tool for Spanish-speaking Latino people, we applied the process of "trans-creation" (derived from combining "translation" and "creation") to culturally and linguistically adapt materials for the target population [16]. Here, we describe the process, changes, and impact of this trans-creation process.

Methods

Setting

This study was conducted within Kaiser Permanente Northern California (KPNC), a nonprofit integrated-care delivery system providing care for over 4.1 million members throughout Northern California. KPNC serves individuals from diverse demographic and socioeconomic backgrounds, including over 700,000 members who self-identify as Latino. The Kaiser Foundation Research Institute Institutional Review Board approved the study.

Visit Planner - English Version

The Visit Planner was first created in English for the "Aligning Patients and Providers" study (trial registration: ClinicalTrials.gov NCT02707146), which is described in detail elsewhere [15]. Briefly, we applied user-centered design principles such as stakeholder involvement and iterative modifications using wire diagrams and paper versions to incorporate end-user feedback in order to develop an iPad waiting room tool (Visit Planner). This tool supported patients in identifying and discussing their top priorities for their primary care visit. Using the Flesch-Kincaid Grade-Level tests, the reading level of the tool was set between a 4th to 6th grade level to ensure accessibility for a broad audience.

The Visit Planner begins with a 30-second video that emphasizes the importance of identifying and discussing top health concerns early in the visit. The next screen prompts the patient to select his/her top two areas of concern out of the six available options (New Problem, Old Problem, Medicines, Stress at Home or at Work, A Personal Concern or Other, Need Something From my Doctor). Once the patient selects one or two option(s), three to four suboptions are offered under each selection. For example, if the patient selects "Medicines" as one of the concerns, three options are presented: (1) Problems With Side Effects, (2) Medicines Cost Too Much, and (3) Stopped Taking. After the patient decides on his/her top priorities for the visit, a second brief video advises the patient to take notes and ask questions during the visit. The actress in both English-language videos portrays a health care provider. The final screen of the Visit Planner is a summary of the patient's selected visit priorities and strategies, which is printed out and given to the patient to help with the ensuing visit.

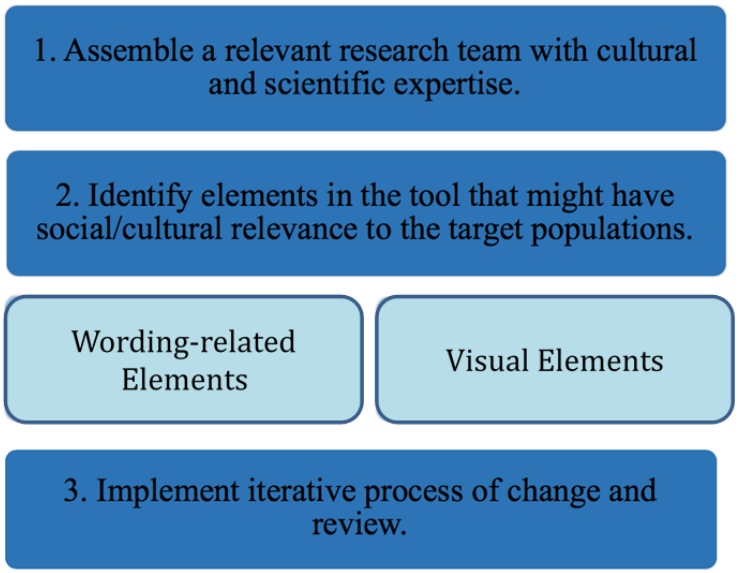
Trans-Creation: A Step-by-Step Process

To better serve the growing Latino population, our research team partnered with health educators and Latino cultural brokers to adapt the Visit Planner for Spanish-speaking Latino people by using the following three-step trans-creation model (Figure 1).

Step 1: Assemble the Team

Team members were chosen based on the concept of "Constituent-Involving," which is the process of hiring and training staff members from within the target population [16]. These cultural brokers provided important insight into nuanced cultural knowledge gained from personal experience. Key team members were hired after they passed a Spanish-language fluency test, a step not usually practiced in other trans-creation/translation projects, but which we found vital for creating an authentic product. Five individuals with relevant skills for this project (three Latino Spanish-language editors, one Latina bilingual health education specialist, and one medical anthropologist) were added to the original research team. The Spanish-language editors ensured that the intent, tone, content strategy, and linguistic fidelity of the Spanish version had as high quality as the original version; the education specialist provided important input on effective communication strategies to reach the target audience; and the medical anthropologist provided cross-discipline expertise in cultural responsiveness and public health interventions.

Figure 1. The three-step transcreation model.



Step 2: Identify Relevant Language and Visual Elements

Our team of cultural experts analyzed the original "Pre-Visit Tool" with a focus on wording, visual elements, and culturally sensitive content area. This process included using the tool itself as well as reviewing transcripts of the video segments within the tool. The editorial staff was specifically trained to analyze the health literacy levels of the general population and to tailor the text to ensure its accessibility to most patients. Analysis of visual elements included assessment of icons, color, and video presentation. Finally, the team identified conceptual areas within the tool where Latino cultural values might influence how the Spanish version of the tool should be created.

Step 3: Testing and Iteration

After the initial process of translation and the changes were made based on visual and cultural elements, the Spanish content was back translated into English to ensure that none of the original intended meaning was lost. Scripts for the video clips were also revised in real-time during filming to ensure optimal comprehension.

Assessing Use of the Mobile App

The Spanish version of the Visit Planner was implemented at two practices within KPNC as part of the larger, externally funded "Aligning Patients and Providers" clinical trial. Time patients spent using the tool was recorded, and a postvisit survey about the tool was administered that asked patients about their visit experiences after using the Pre-Visit Planner. Survey questions reported here were adapted from the Stanford Communication with Physicians Scale [17]. Comparison of results between English- and Spanish-language tool users was made using *t* test or Chi-square tests, as appropriate.

Results

Our trans-creation process resulted in several key changes related to both the wording of the text and the visual representation on the screen.

Wording Modifications

Family

In Latino culture, family (both immediate and extended family members) is often highly valued and can play a key role in medical care [18]. The importance of family as a cultural influence led our research team to embed concepts of family throughout the Visit Planner tool. For example, the team changed the visit priority option "Stress at Home or Work" to "Mi Familia" (translation: my family) to acknowledge the emphasis Latino culture places on family and multigenerational homes. Another change to emphasize the role of family was to replace "caregiving issues" as a concern to the more specific "No sé cómo atender a un ser querido" (translation: I don't know how to care for a loved one).

Literacy

The English version of the Visit Planner was designed to have a 4th to 6th grade reading level using the Flesch-Kincaid grade-level assessment tool. Because there is currently no validated tool to assess Spanish reading grade level, we formally assessed the literacy level by translating our final Spanish text back into English and then applying the Flesch-Kincaid grade-level assessment. Health-specific literacy was also an important consideration. For example, we changed the phrase "Referral to specialist" to "Que me diga si necesito un especialista" (translation: Tell me if I need a specialist) to remove the word "referral" due to the concern that patients with less medical literacy may not immediately understand the term.

Stigma

During the trans-creation process, the team is trained to consider best practices for a given cultural community. This knowledge included identifying that mental health can be a particularly difficult subject to discuss. Indeed, prior research has shown that Latino people are less likely to inquire about and receive mental health care due to the stigma associated with mental health [19,20]. Accordingly, under the topic of "Personal Concerns," the choices of "anxiety" and "depression" in the

English version were changed to the less stigma-associated "Me siento desanimado" (translation: I am feeling down). Use of this euphemism was designed to encourage Latino patients to talk about emotional health and help those who may not recognize that certain symptoms they experience may be related to anxiety or depression.

Concise Text

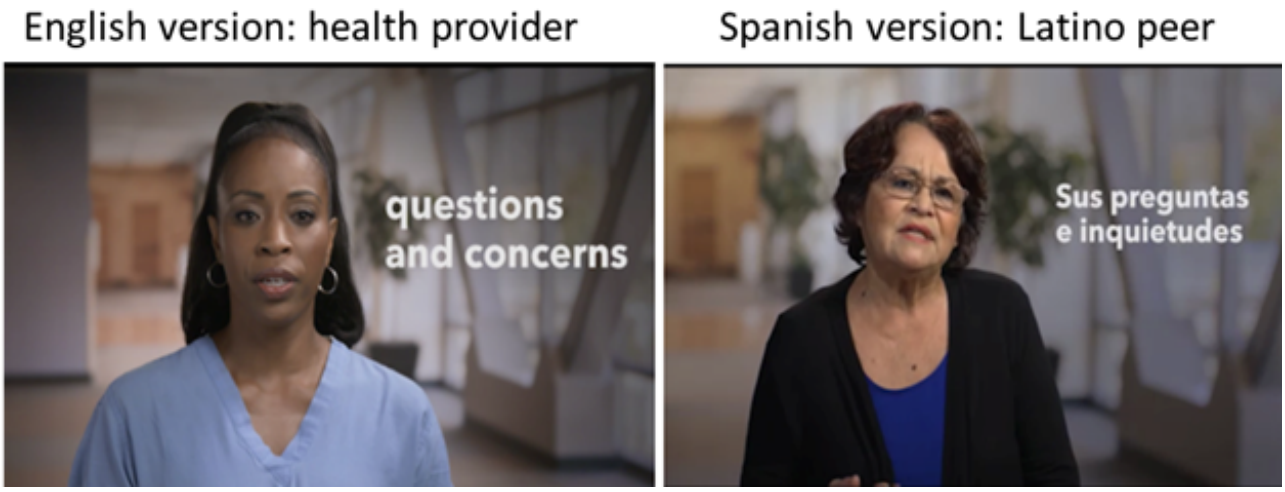
Spanish tends to require more syllables to convey the same direct wording as English [21]. Thus, in several cases, we needed to change the wording of the original English text before translating it to Spanish. For example, one of the choices in the "Need something more from my doctor" category was "Blood test, x-ray, or other test" (32 spaces). Direct translation of these words into Spanish yields "Prueba de sangre, rayos X u otra prueba" (39 spaces, 18% longer). In this case, we adapted the English version to "Ask if I need more tests," which translates to Spanish as "Confirmar si necesito mas pruebas" (33 spaces).

Visual Modifications

Actors

In the original Visit Planner, the actors in the video clips portrayed a "health provider," whereas in the videos for the Spanish-language version, the actors portrayed fellow patients (Figure 2). In conjunction with this patient-level portrayal, the script was modified, so that the coaching advice was framed as "en mi experiencia" (translation: In my experience), because prior research has shown that peer-to-peer advising is more effective within the Latino population [22]. The change in phrasing was also intended to increase *confianza* (trust). The actors were selected to resemble the typical age and appearance of our Latino patients, and the actors were instructed to dress casually and choose everyday wear for the video clips.

Figure 2. Actors in the Spanish version were switched with a peer.



Icons

All the icons from the English Visit Planner were used in the Spanish version, with the exception of the icon originally associated with "Stress at home or work," which in the English version, was an outline of a house. To better align with the changed text (the English equivalent of "my family"), we also changed the icon to an outline of a family (Figure 3).

Comparison of the Spanish and English Visit Planner Versions

Spanish-speaking patients (n=41) spent a similar amount of time going through the Spanish version of the Visit Planner as users of the original English tool (n=273; 5.26 minutes vs 5.05 minutes; *P*=.48). In the postvisit survey, responses were similar between groups (Table 1), with the exception that Spanish speakers were more likely to discuss their visit experience with friends and loved ones than English speakers (35/41 [86%] vs 188/273 [69%], *P*=.02).

Figure 3. Screenshots of the English and Spanish versions of the primary prioritization page.

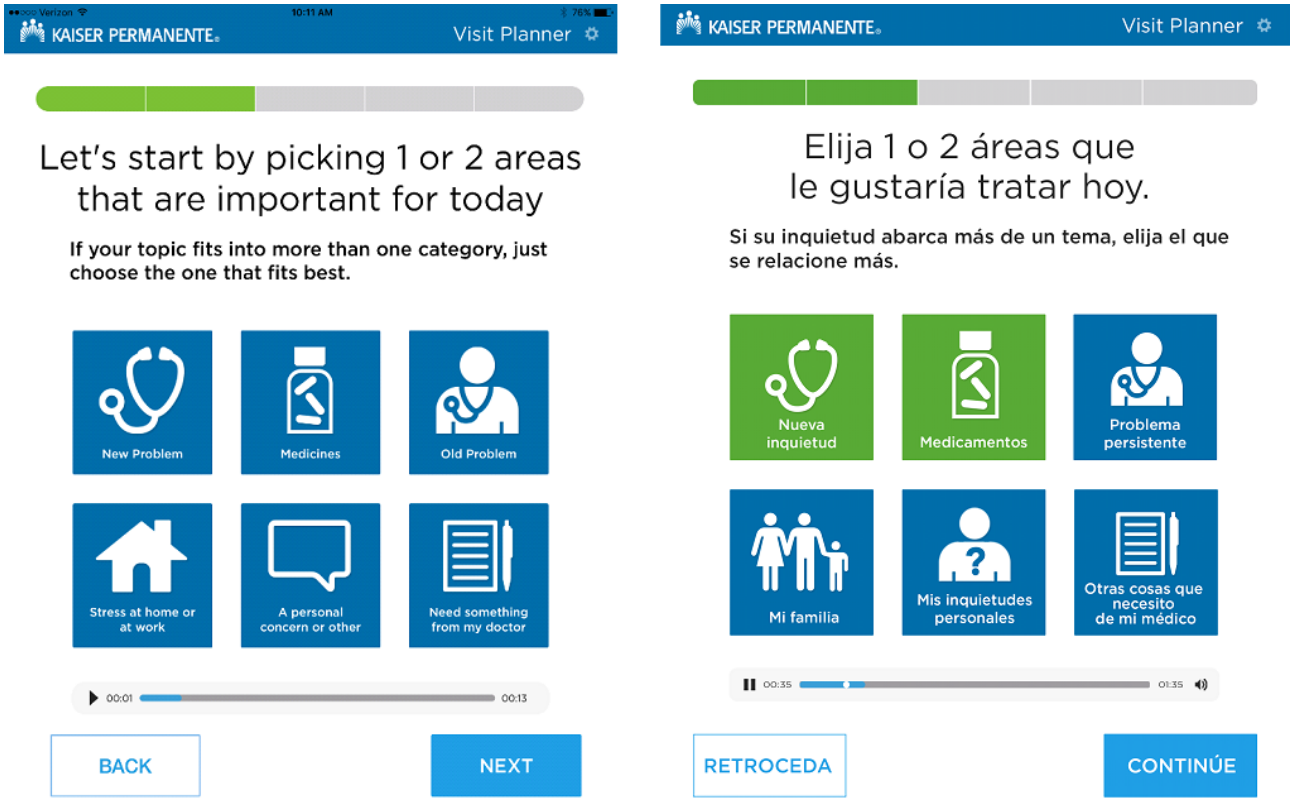


Table 1. Postvisit survey results comparing patients using the English version and those using the Spanish version of the Visit Planner Tool.

Questions	Proportion of respondents answering "Yes"		*P* value
	English version (N=273), n (%)	Spanish version (N=41), n (%)	
Did you ask questions about the things you want to know and the things you don't understand about your treatment?	259 (95)	36 (88)	.09
Did you tell your doctor about your top concerns at the beginning of the visit?	257 (94)	37 (91)	.39
Did you take notes or ask questions during your visit?	246 (90)	33 (81)	.11
After a visit, did you talk to friends or loved ones about the visit?	188 (69)	35 (86)	.02
Did you discuss any personal problems that may be related to your illness?	175 (64)	26 (63)	.89

Discussion

In this case study, we describe the process of adapting an English-language health mobile app for Spanish-speaking Latino people. Because merely translating health education materials from one language to another does not address cultural differences in how medical care is perceived by patients from different backgrounds, we implemented a three-step trans-creation process to identify, reinforce, and build upon the Latino cultural values and concepts [23]. Trans-creation steps included (1) adding members to the research team with appropriate linguistic and cultural expertise, (2) carefully assessing every element of the tool for its potential cultural significance, and (3) iteratively testing and assessing the resulting changes. This iterative process required five distinct rounds of testing and associated modifications.

Efforts to design effective health IT tools are highly relevant to addressing the problem of health disparities within the large and growing US Latino population. More than 50 million Latinos in the United States have limited English proficiency and face the possibility of being left further behind as new health IT tools are developed to improve doctor-patient communication and other aspects of health care [9]. Our Visit Planner study represents one of the first published examples of applying the trans-creation process to a waiting room, iPad-based health mobile app designed to help patients more effectively communicate with their providers during visits. The lessons learned in this process, while specific to our Visit Planner tool, provide illustrative examples of the broad issues that may be

faced by others planning to create educational tools for Latino people.

Many of the changes we made were reflective of several core cultural constructs that have been previously described in the US Latino population. These constructs, applied to the Visit Planner trans-creation process, included *familismo* (family), *confianza* (trust), and the role of mental health stigma. The key changes that we made in our Visit Planner (eg, altering language and images to emphasize family, replacing "medical" actors with "peer" actors to increase trust and softening the language used to describe anxiety and depression to reduce stigma), all exemplify how the trans-creation process adds crucial enhancements beyond mere translation. Lessons we learned during this process are all potentially transferable to other trans-creation projects.

Initial evaluation of how patients used the Visit Planner tool revealed that Latino users of the Spanish version spent a similar amount of time as they did on the English version, to work through the sequential screens, indicating that we were successful in keeping the time and effort required to complete the Visit Planner similar in the two versions. Moreover, the survey responses completed after the visit indicated that both English- and Spanish-speaking users had similar experiences during the visit. The only difference between groups was the greater discussion with family among Latino people, which provides further empiric evidence to support our emphasis on *familismo*.

In the English-language version, our team asked participants about mental health symptoms directly; however, this method was not appropriate for the Spanish-language version. In Latino culture, mental health problems are not necessarily validated, are often seen as a sign of weakness, and may carry stigma [20]. In fact, although the rates of mental illness are similar among white and Latino people, the latter are much less likely to seek health care for mental health concerns [20]. Association with a mental illness may even reflect poorly on the family and place stigma not only on the individual but also their relatives [22]. Concern about this stigma influenced our trans-creation process in a way that would not have occurred with simple translation.

A recent review of health IT implementation for Latino people demonstrated several important gaps in the literature, which our study helps to fill [24]. This review found that only 60% of studies targeted toward Spanish-speaking Latino people were culturally tailored. Increasing culturally tailored health IT is necessary because a culture-specific platform can better engage patients and improve health outcomes. This narrative synthesis also highlighted the crucial need for future studies to explicitly detail their cultural tailoring process and support it with evidence-based literature. Our study does this by providing a model to follow and a description of our process; future studies can more easily replicate cultural tailoring with our trans-creation model and description.

Two limitations of our study deserve mention. First, because of the funding sequence, the English version was created first and used as a template for the Spanish version. Many experts recommend that culturally tailored materials be created *de novo* or in tandem, rather than as variations on existing materials [25,26]. Second, the results of our randomized clinical trial are not yet available; therefore, we cannot report on the clinical impact of our tool.

Beyond simple translation, trans-creation provides a powerful framework for adapting English health IT tools for patients from different cultures who speak different languages. Our trans-creation process, while specific to the Visit Planner tool, demonstrated several key principles that can be applied to further the reach of health education tools for patients from non-English speaking cultures.

Acknowledgments

Funding for this study was obtained from the Patient-Centered Outcomes Research Institute (CDR-1403-11992), National Institute of Diabetes and Kidney Diseases (K24DK109114), and The Permanente Medical Group. Sponsors had no role in review or approval of the manuscript. The authors wish to thank Margaret Lapiz for her support and the actors and patients who provided their time and advice.

Conflicts of Interest

None declared.

References

1. US Census Bureau. Washington, DC: US Dept of Commerce; 2012. U.S. Census Bureau Projections Show a Slower Growing, Older, More Diverse Nation a Half Century from Now URL: https://www.census.gov/newsroom/releases/archives/population/cb12-243.html [accessed 2019-02-20] [WebCite Cache ID 76KuxKX3w]
2. McWilliams JM, Meara E, Zaslavsky AM, Ayanian JZ. Differences in control of cardiovascular disease and diabetes by race, ethnicity, and education: U.S. trends from 1999 to 2006 and effects of medicare coverage. Ann Intern Med 2009 Apr 21;150(8):505-515. [Medline: 19380852]
3. Ward BW, Schiller JS. Prevalence of multiple chronic conditions among US adults: estimates from the National Health Interview Survey, 2010. Prev Chronic Dis 2013 Apr 25;10:E65 [FREE Full text] [doi: 10.5888/pcd10.120203] [Medline: 23618545]

4. Stewart MA. Effective physician-patient communication and health outcomes: a review. CMAJ 1995 May 1;152(9):1423-1433 [FREE Full text] [Medline: 7728691]
5. Ha JF, Longnecker N. Doctor-patient communication: a review. Ochsner J 2010;10(1):38-43 [FREE Full text] [Medline: 21603354]
6. Heisler M, Bouknight RR, Hayward RA, Smith DM, Kerr EA. The relative importance of physician communication, participatory decision making, and patient understanding in diabetes self-management. J Gen Intern Med 2002 Apr;17(4):243-252. [doi: 10.1046/j.1525-1497.2002.10905.x]
7. Aikens JE, Bingham R, Piette JD. Patient-provider communication and self-care behavior among type 2 diabetes patients. Diabetes Educ 2005 Dec;31(5):681-690. [doi: 10.1177/0145721705280829] [Medline: 16203852]
8. Sudore R, Landefeld C, Pérez-Stable EJ, Bibbins-Domingo K, Williams B, Schillinger D. Unraveling the relationship between literacy, language proficiency, and patient-physician communication. Patient Educ Couns 2009 Jun;75(3):398-402 [FREE Full text] [doi: 10.1016/j.pec.2009.02.019] [Medline: 19442478]
9. Parker MM, Fernández A, Moffet HH, Grant RW, Torreblanca A, Karter AJ. Association of Patient-Physician Language Concordance and Glycemic Control for Limited–English Proficiency Latinos With Type 2 Diabetes. JAMA Intern Med 2017 Mar 01;177(3):380. [doi: 10.1001/jamainternmed.2016.8648] [Medline: 28114680]
10. Vega WA, Rodriguez MA, Ang A. Addressing stigma of depression in Latino primary care patients. Gen Hosp Psychiatry 2010 Mar;32(2):182-191. [doi: 10.1016/j.genhosppsych.2009.10.008] [Medline: 20302993]
11. Zamudio CD, Sanchez G, Altschuler A, Grant RW. Influence of Language and Culture in the Primary Care of Spanish-Speaking Latino Adults with Poorly Controlled Diabetes: A Qualitative Study. Ethn Dis 2017 Dec;27(4):379-386 [FREE Full text] [doi: 10.18865/ed.27.4.379] [Medline: 29225438]
12. Chong N. The Latino Patient: A Cultural Guide for Health Care Providers. Boston: Nicholas Brealey; 2002:A.
13. Costa BM, Fitzgerald KJ, Jones KM, Dunning AT. Effectiveness of IT-based diabetes management interventions: a review of the literature. BMC Fam Pract 2009 Nov 17;10:72 [FREE Full text] [doi: 10.1186/1471-2296-10-72] [Medline: 19917136]
14. López L, Grant RW. Closing the Gap: Eliminating Health Care Disparities among Latinos with Diabetes Using Health Information Technology Tools and Patient Navigators. J Diabetes Sci Technol 2012 Jan;6(1):169-176. [doi: 10.1177/193229681200600121]
15. Lyles CR, Altschuler A, Chawla N, Kowalski C, McQuillan D, Bayliss E, et al. User-Centered Design of a Tablet Waiting Room Tool for Complex Patients to Prioritize Discussion Topics for Primary Care Visits. JMIR Mhealth Uhealth 2016 Sep 14;4(3):e108 [FREE Full text] [doi: 10.2196/mhealth.6187] [Medline: 27627965]
16. Kreuter MW, Lukwago SN, Bucholtz RDDC, Clark EM, Sanders-Thompson V. Achieving cultural appropriateness in health promotion programs: targeted and tailored approaches. Health Educ Behav 2003 Apr;30(2):133-146. [doi: 10.1177/1090198102251021] [Medline: 12693519]
17. Lorig A, Ritter P, González V, Laurent D, Lynch J. Outcome Measures for Health Education and other Health Care Interventions. Thousand Oaks, CA: Sage Publications; 1996.
18. Larkey L, Hecht M, Miller K, Alatorre C. Hispanic cultural norms for health-seeking behaviors in the face of symptoms. Health Educ Behav 2001 Feb;28(1):65-80. [doi: 10.1177/109019810102800107] [Medline: 11213143]
19. NIH US National Library of Medicine. 2001. Mental Health: A Report of the Surgeon General URL: https://profiles.nlm.nih.gov/ps/retrieve/ResourceMetadata/NNBBHS [accessed 2019-02-20] [WebCite Cache ID 76Kw3R4ae]
20. Breslau J, Cefalu M, Wong EC, Burnam MA, Hunter GP, Florez KR, et al. Racial/ethnic differences in perception of need for mental health treatment in a US national sample. Soc Psychiatry Psychiatr Epidemiol 2017 Dec;52(8):929-937 [FREE Full text] [doi: 10.1007/s00127-017-1400-2] [Medline: 28550518]
21. Pellegrino F, Coup. A Cross-Language Perspective on Speech Information Rate. Language 2011;87(3):539-558.
22. Kouyoumdjian H. Barriers to Community Mental Health Services for Latinos: Treatment Considerations. Clinical Psychology: Science and Practice 2003 Nov 01;10(4):394-422. [doi: 10.1093/clipsy/bpg041]
23. Simmons VN, Quinn G, Litvin EB, Rojas A, Jimenez J, Castro E, et al. Transcreation of validated smoking relapse-prevention booklets for use with Hispanic populations. J Health Care Poor Underserved 2011 Aug;22(3):886-893 [FREE Full text] [doi: 10.1353/hpu.2011.0091] [Medline: 21841285]
24. Chaet AV, Morshedi B, Wells KJ, Barnes LE, Valdez R. Spanish-Language Consumer Health Information Technology Interventions: A Systematic Review. J Med Internet Res 2016 Aug 10;18(8):e214. [doi: 10.2196/jmir.5794]
25. Montague E, Perchonok J. Health and Wellness Technology Use by Historically Underserved Health Consumers: Systematic Review. J Med Internet Res 2012 May 31;14(3):e78. [doi: 10.2196/jmir.2095] [Medline: 22652979]
26. Kreuter MW, Sugg-Skinner C, Holt CL, Clark EM, Haire-Joshu D, Fu Q, et al. Cultural tailoring for mammography and fruit and vegetable intake among low-income African-American women in urban public health centers. Prev Med 2005 Jul;41(1):53-62. [doi: 10.1016/j.ypmed.2004.10.013] [Medline: 15916993]

Abbreviations

IT: information technology
KPNC: Kaiser Permanente Northern California

Edited by G Eysenbach; submitted 09.10.18; peer-reviewed by S Khan, S Richardson; comments to author 19.12.18; revised version received 19.12.18; accepted 03.01.19; published 05.04.19.

Please cite as:
Ruvalcaba D, Nagao Peck H, Lyles C, Uratsu CS, Escobar PR, Grant RW
Translating/Creating a Culturally Responsive Spanish-Language Mobile App for Visit Preparation: Case Study of "Trans-Creation"
JMIR Mhealth Uhealth 2019;7(4):e12457
URL: https://mhealth.jmir.org/2019/4/e12457/
doi:10.2196/12457
PMID:30950803

Original Paper

A Mobile Phone App for Bedside Nursing Care: Design and Development Using an Adapted Software Development Life Cycle Model

Frederic Ehrler[1], PhD; Christian Lovis[1,2], MD, MPH, FACMI; Katherine Blondon[3], MD, PhD

[1]Division of Medical Information Sciences, University Hospitals of Geneva, Geneva, Switzerland

[2]Faculty of Medicine, University of Geneva, Geneva, Switzerland

[3]Medical Directorate, University Hospitals of Geneva, Geneva, Switzerland

Corresponding Author:
Frederic Ehrler, PhD
Division of Medical Information Sciences
University Hospitals of Geneva
4, Gabrielle-Perret-Gentil
Geneva, 1205
Switzerland
Phone: 41 223728697
Email: frederic.ehrler@hcuge.ch

Abstract

Background: Nurses are increasingly spending time on computers, and providing them with a tailored tool to access clinical information and perform documentation at the bedside could help to improve their efficiency. Designing an app to support nurses' work at the bedside is a challenging task, given the complexity of the care process.

Objective: This study aimed to present the design, development, and testing of a smartphone app for nurses guided by an adapted software development life cycle model that takes into consideration the complexity and constraints of a health care setting.

Methods: The model drives us through an iterative development process intersected by 3 stages of formative evaluation of growing ecological validity.

Results: The initial requirements identification stage included 11 participants who helped us select the most important functionalities to integrate into the tool. Starting with a usability evaluation allowed for the identification of design issues that could have caused misuse. Then, making on-site evaluations under the supervision of an investigator helped to understand the adequacy of the tool with limited risks. Finally, the on-site evaluation allowed us to validate the acceptance of the app by caregivers.

Conclusions: The interpretation of the collected evaluation confirms the necessary involvement of end users early in the process to help address the heterogeneity of the nursing workflow processes in the different wards. We also highlight the delicate balance between high-security measures to protect access to patient data and maintaining ease of access for efficiency and usability. Although a close collaboration with clinicians throughout the entire project facilitated the development of a tailored solution, it was also important to involve all stakeholders, in particular, the information technology (IT) security officers.

(JMIR Mhealth Uhealth 2019;7(4):e12551) doi:10.2196/12551

KEYWORDS

mHealth; nursing; hospital information systems

Introduction

Background

Studies on workplace efficiency and patient care have brought forward the need to improve workflow processes [1], particularly with the implementation of electronic health records (EHRs) and documentation requirements. Nurses and nursing assistants provide a range of interventions for each patient, which are determined through planned nursing care or doctor prescriptions. For example, they administer medications, check vital signs, change bandages, handle meals, and provide patient support and education. All these bedside activities vary for each patient. In our hospital, nurses print out the intervention list for each patient at the beginning of each shift to guide the provision of care. This list is also used as support for patient handoffs at the

beginning of the shift and helps nurses prepare the necessary equipment for patient care before entering the rooms. Throughout the shift, nurses document assessments such as vital signs, observations, as well as the planned interventions.

Although the printed task lists guide the bedside activities, they are also potential sources of errors. Printouts are snapshots of patient interventions at a point of time, and updates in the EHR need to be scribbled in manually and may even be missed. Furthermore, these printouts do not replace the need for nurses to log into the patients' EHRs to document their actions and observations [2]. These paper-based supports are also used as temporary repositories for clinical data (ie, vital signs) before they are entered into the EHR; besides delaying the availability of these data in the EHR, this transcription process increases the risk of errors [3]. Finally, the institution's move toward going paperless is another reason to look for alternative solutions to improve the bedside patient care process.

Mobile devices provide many new opportunities and tools for patient care [4-10]. In fact, a smartphone app can help address some of the concerns raised above [11]. An EHR-connected app can provide the needed intervention list at timely moments and allows real-time documentation [12,13]. For example, entering data directly on the device at the bedside could help avoid potential transcription errors, with data immediately available in the EHR. In addition, access to elements in the EHR can help nurses respond to patient questions more readily, thus encouraging patient empowerment. Most importantly, however, we hypothesize that a smartphone app can help decrease the amount of time spent on clinical documentation and updates in the medical charts by supporting the nurses' workflow process. Previous studies report variable time spent on documentation (20%-35% of a shift), which is often completed away from the patient [14-17].

Although several projects have already explored the use of mobile health (mHealth) for physicians [18,19], providing mobile access to EHRs is still in its budding stage. However, it is recognized that designing and developing a mobile app in the complexity of a health care environment includes many challenges linked to specific constraints of the medical domain as well as linked to the form factor of mobile apps [20,21].

Despite an accumulation of best practices research that identifies success factors, IT implementation projects are often not successful. Across industry sectors, at least 40% of such generic IT projects are either abandoned or fail to meet business requirements, whereas fewer than 40% of large systems purchased from vendors meet their goals [22]. Some sources even report failure rates of up to 70% [23].

There are serious issues with the implementation of the health information system (HIS), and reports of HIS implementation failure are not hard to find in the literature. The reasons for failure include inadequate funding; lack of IT infrastructure; poor leadership; inadequate end-user engagement and unrealistic timelines; and lack of compatibility of HIS with current work processes, the organizational culture, and vision [24-27].

Previous Works

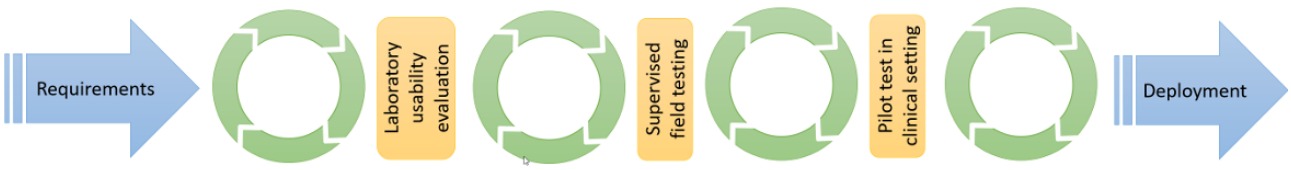
We experienced these difficulties during a 2014 pilot study conducted in 3 wards of our institution to explore the use of tablets with EHR access for nursing teams. The EHR version was the same as the desktop one and was used during a week-long test with workstations on wheels (WOWs). These tablets not only provided support for documentation at the bedside but also helped avoid iterative trips to the nursing station to collect the various supplies for bedside care. WOWs with laptops were also available in the wards during this period. Nurses printed less task lists, reported higher efficiency in validating tasks, and spent more time at the bedside. However, the tablet EHR did not adequately support the nurses' workflow, for instance, it did not provide an overview of the assigned tasks, so the nurses could not plan the tasks for the day. In fact, some tasks were even forgotten or performed with a delay. Usability of EHR on the tablet was low, with difficulty reading the data due to the screen size, fonts, and the lack of adaptation of the EHR to the mobile device. Suggested improvements were the grouping of similar tasks together and providing multipatient views of the assigned tasks to organize their work. Furthermore, rethinking the design of the EHR for tablets could help address the usability issues and could help provide an improved support tool throughout the nursing workflow. This was confirmed in a study in another setting [28].

Developing mobile apps in a health care setting is not a simple task. Several approaches have been proposed that can be followed to drive the development of apps, but there is often a lack of feedback on the impact of the chosen method on the quality and acceptance of the produced app [29].

As an attempt to deal with the strong constraint associated with the development of health information technology (HIT) in health care, we proposed, in a previous paper, a tailored software development life cycle (SDLC) model that takes into consideration the constraints of mobile app in a health care environment and integrates both development and evaluation frameworks (Figure 1) [30].

There is no clear consensus on existing SDLC models, but we can confidently divide them into traditional and agile models. Traditional models are mainly sequential and include models such as the waterfall, the spiral, or the V-shape. Agile models, on the other hand, are a sequence of short cycles that allow better responsiveness [31-33]. Our adapted SDLC model differs from the existing ones as it adds more detail in the development stage; several substages are separated by evaluations of increasing ecological validity.

Figure 1. Overall life cycle process with the different evaluation stages.



This model begins with a requirements identification stage, followed by a series of development cycles, including specification, prototyping, development, and functional validation, interspersed by formative evaluations of growing ecological validity. Starting with low constraints in the early testing stage allowed us to validate different aspects of the tool without being limited by unnecessary complexity.

Objectives

In this paper, we aim to demonstrate how our tailored SDLC model based on evidence and user feedback can drive the design and development process of an app in a health care setting, from the initial brainstorming to the testing of a tool connected to our local EHR system. We also describe the conceptual framework adopted to address the hospital security requirements for implementation.

Methods

Study Design

We recruited nurses, nursing aids, and head nurses from the departments of medicine and surgery as cocreators throughout our project. These departments have the highest number of beds for acute care in our hospital. Choosing one unit from each department aimed at being the most representative of ward work in acute care settings in our hospital. Therefore, by creating a tool that addresses their combined needs, we hoped to better anticipate a future all-hospital deployment. The objective of the app was to provide access to certain parts of the EHR, targeting the elements in the bedside workflow, which can benefit from timelier data entry into the EHR. We proceeded according to the adapted SDLC model (Table 1), starting with requirements identification, followed by subsequent iterative development and evaluation stages.

Requirements Identification

The overall process started with the identification of the needs through a couple of focus groups with the end users. The first focus group focused on brainstorming and determining which target population (nurses, physicians, and patients) and which needs could be addressed with mobile access to the EHR. The results were subsequently analyzed in depth, using a thematic analysis approach and included the use of card-sorting, to define institutional priorities and feasibility of the various propositions, taking into consideration the available resources (ie, available data, estimated complexity, and potential gain). Card-sorting is a reliable approach to find patterns in the way participants organize content and was used to identify common elements in design from the various brainstorming ideas [34]. The subsequent analysis addressed the feasibility of each idea and led to the choice of an app for nurses at the bedside. The aim of the following focus group session was to understand the workflow of the nurses and to identify the main interactions that they have with the clinical information system. Discussions with end users were transcribed with a subsequent thematic analysis to identify the main interactions between the users and the system.

Development Cycles

The initial requirement fueled the start of iterative user-centered development cycles. The cycle started with the creation of mockups integrating the specifications identified from the previous cycle. The mockups were discussed with the project team for validation. The validated evolutions were then implemented into the functional prototype and tested using functional evaluations. In these evaluations, all app functions were executed in a test case procedure. Then, the actual and expected outputs were compared to check whether the app addressed the specified end user needs. At each stage, the prototype of increasing complexity was presented to a group of end users to gather feedback about the proposed concepts. This feedback was integrated into the next cycle to refine the functionalities as well as the interaction through the user interface.

Formative Evaluations

We conducted formative evaluations at key milestones of the project. These formative evaluations were more in-depth than those performed during the iterative development cycles and often raised issues that had not been noticed during the development cycles. They, thus, contributed to the evolution of the tool specifications.

Table 1. Summary of the software development life cycle (SDLC) stages and goal description.

Number	Stage	Methodology	Stage purpose
1	Brainstorming	Focus group	Identification of the most promising areas to develop and intervene
2	Requirements	Focus group	Identification of the basic functional requirements
3	Development cycle 1	Development cycles	Development of the Alpha prototype
4	Usability test	Laboratory usability test	Evaluation of the tool usability
5	Development cycle 2	Development cycles	Development of the Beta prototype
6	User test	Supervised on-site test	Test of the tool on site to validate workflow adequacy
7	Development cycle 3	Development cycles	Development of the final prototype
8	Pilot test	Unsupervised on-site test	Summative evaluation to test app acceptance

Usability Test

The first formative evaluation was a laboratory usability test [35-38]. During this evaluation, nurses performed one of the two predefined scenarios in a controlled but realistic environment. The scenarios were created according to the results of previous field observations. Overall, one scenario was designed for the medical ward nurses and one for the surgical ward nurses. The scenarios guide the participants through a sequence of actions that the nurses are likely to perform in real life and that are supported by our app.

The outline of the scenario is as follows:

1. Identify the patient by scanning the Quick Response (QR) code on his bracelet
2. Review the interventions performed during the night
3. State the necessary interventions for the medication rounds, validate the administration of drugs, and cancel the validated breakfast task
4. Postpone and duplicate an intervention
5. Validate the start of an intravenous (IV) drug; check the pro re nata (PRN) painkillers, administer a dose of painkiller, and validate this action in the app
6. Indicate the elapsed volume for the IV drug and document the patient's pain level
7. Document that the patient refused to eat his dinner
8. Complete a Braden scale assessment and take a photo of the lesion
9. List the remaining interventions to be completed before the end of the shift
10. Log out

After a brief presentation of the tool, participants were asked to follow the scenario and to perform the sequence of tasks. All the tasks were video-recorded for subsequent analysis. We measured the success rate for the tasks and reported the errors with the proposed improvements. The participants completed the System Usability Scale (SUS) questionnaire [39] and discussed their satisfaction and impressions at the end of the session. The investigators then conducted an analysis of the usability of the tool.

Supervised Field Tests

Supervised field tests took place in 2 wards of the University Hospitals of Geneva (HUG). After a short personalized training, the care providers were asked to use the app to perform their daily tasks and were followed by an observer. The observer closely monitored their use of the app and identified usability issues and bugs. After each session, bugs were systematically reported to the development team for corrections. This close supervision was vital to ensure patient safety and included the verification that all patient information handled with the app was correctly recorded in the EHR at the end of each test session.

Pilot Test in the Wards

The pilot study took place in a surgical ward and a medical ward of 18 beds, each at the HUG [40]. The app was provided to the participants on institutional smartphones. The app usage was restricted to weekdays from 7 am to 6 pm to ensure technical support in case of problems.

During the week before the start of the study, ward nurses and nursing assistants received a short training about the app. As participation in the study was on a voluntary basis, the participants were informed that the use of the app was encouraged but not compulsory during the study period. During the whole study, the investigators conducted frequent visits to the wards to provide support and to collect feedback from the users, in particular, for bugs and suggestions for improvement. These findings were forwarded to the IT team, who corrected the bugs and added minor improvements when possible.

To evaluate the usage of the tool by the participants, the app logged all of the participants' interactions with the app automatically, for subsequent analysis. At the end of the study, participants were asked to complete a tailored technology acceptance questionnaire. The questionnaire was derived from the unified theory of acceptance and use of technology (UTAUT) [41] to fit our particular setting. It contained 21 questions with a 7-point Likert scale, which were divided into 5 dimensions: performance efficiency (4 questions), effort expectancy (4 questions), social influence (4 questions), facilitating conditions (4 questions), and behavioral intention to use the system in the future (3 questions). Perceived usage during the study and general comments were also collected.

Results

Setting

Our research took place at a tertiary teaching hospital of 1800 beds, with 60,000 admissions per year and more than 4000 nurses. Table 2 presents a summary of the methods and number of participants at each stage of the SDLC model.

Requirements Identification

Ideation

The initial brainstorming session with 24 nurses and nursing assistants and 8 physicians from medical and surgical wards produced 30 different ideas about how mobile devices could help address needs of nurses, physicians, and patients in our institution. Among these propositions, many required data that were not available electronically or that were not available within the EHR (eg, menu choices for patients) or had low feasibility or low priority (scan of ward pharmacy to automatically generate stock refills). After considering estimated cost and complexity as well as potential gain for the institution, the project team decided to focus on a bedside support tool for nursing teams. This was the beginning of Project *Bedside Mobility*.

Requirements Collection

In the subsequent focus group with 5 nurses, 4 nursing assistants, and 2 head nurse investigators, we explored in more detail the functionalities that nurses envisioned to support their workflow. The thematic analysis of the focus group transcript revealed the need to improve current processes: scattered information retrieval, photo and vital sign documentation needs at the bedside, and lack of a continuously updated task list (Textbox

1). We found that nurses spent considerable time retrieving information and entering data in the EHR because it is scattered in several sections of the EHR. Nurses wanted to be able to easily document wound progress with pictures at the bedside. Picture uploads in the EHR are a multistep process, which involved using and connecting an additional device (ie, camera) and manually entering data (ie, consent, date [current and date of photo], type of photo, etc).

We also collected suggestions about new possibilities that do not exist in the EHR (Textbox 2). First, by identifying one's assigned patients for a given shift, one should be able to visualize the required tasks in both an individual or multipatient view. In addition, caregivers requested to simplify patient identification at the bedside. Second, users wished to simplify the process between task validation and data entry, such as for blood glucose levels or fluids (ins, and outs), rather than visualizing and validating the task list in one page and then going to another page to enter the data. They also wished to be able to receive notifications (eg, at the end of an IV drip). Finally, the nurses also needed assistance in calculations such as for fluid balance; currently, data for the fluid intake and output are collected on pieces of paper at the bedside or in the toilet facilities, and the total is calculated with a calculator before data entry in the EHR.

Selected Functionalities

Functionalities were selected from the initial needs assessment and were refined with iterative discussions with the nursing staff. We provide a brief overview of the app's main functionalities in Figure 2.

In accordance with our hospital's patient safety strategy, the app provides access to patient charts by scanning a QR code on the patients' hospital bracelets with the smartphone camera.

Each patient chart includes general patient information data (identity, age, and length of hospital stay) with clinical data on the current hospitalization, comorbidities, and daily nursing objectives. These components were included to provide support for the handoff process.

In each patient's chart, nursing interventions are sorted in a chronological order, starting from the current time of app use. Interventions of similar nature (eg, medication) are grouped together for easier readability (Figure 3). Interventions can be validated when completed with a rapid swipe motion, but can also be modified, delayed, repeated, or validated as incomplete (eg, in the case of patient refusal). These functions are often used by nurses during charting in the EHR.

Table 2. Summary of the software development life cycle (SDLC) stages and user involvement.

Number	Stage	Methodology	Number of participants	Iterations	Duration
1	Brainstorming	Focus group	32	1	1 month
2	Requirements	Focus group	11	2	2 months
3	Development 1	Development cycles	5	8	12 months
4	Usability test	Laboratory usability test	10	1	1 month
5	Development 2	Development cycles	5	2	4 months
6	User test	Supervised on-site test	20	1	2 weeks per ward
7	Development 3	Development cycles	5	2	1 month
8	Pilot test	Unsupervised on-site test	30	2	2 months

Textbox 1. Weaknesses identified in the electronic health records.

- Scattered information
- Complicated data entry: vital signs and photos
- Lack of real-time task updates

Textbox 2. New desired functionalities.

- Identifying assigned patients during a shift
- Simplified patient identification at the bedside
- Direct data entry: ins and outs and blood glucose levels
- Assistance for calculations (eg, total volumes of fluid)
- Notifications (eg, end of intravenous drip)

Figure 2. Overview of app functionalities (PRN: pro re nata).

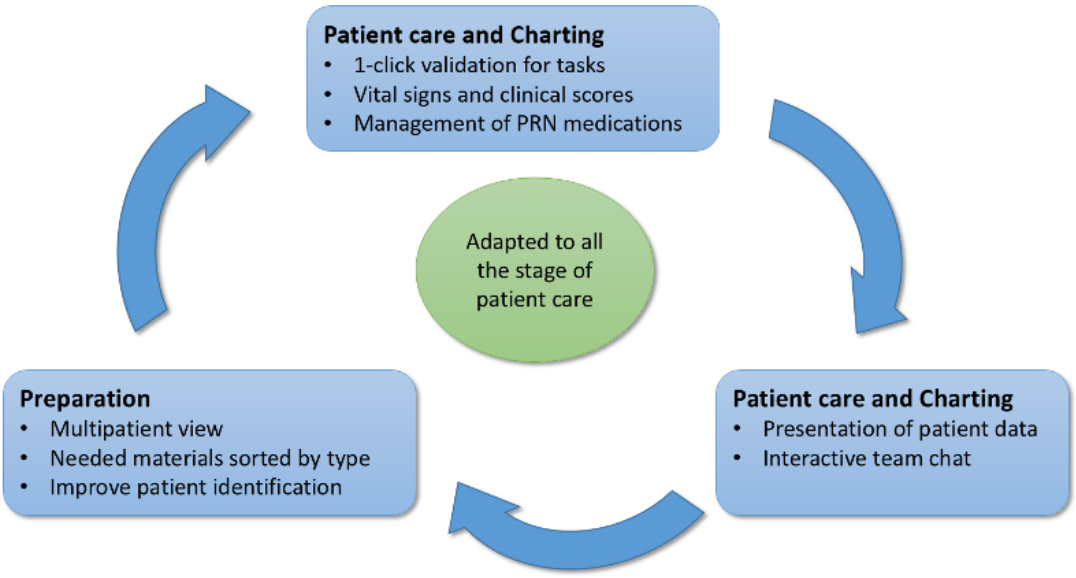


Figure 3. Screenshot of the Bedside Mobility app, view of the daily intervention.

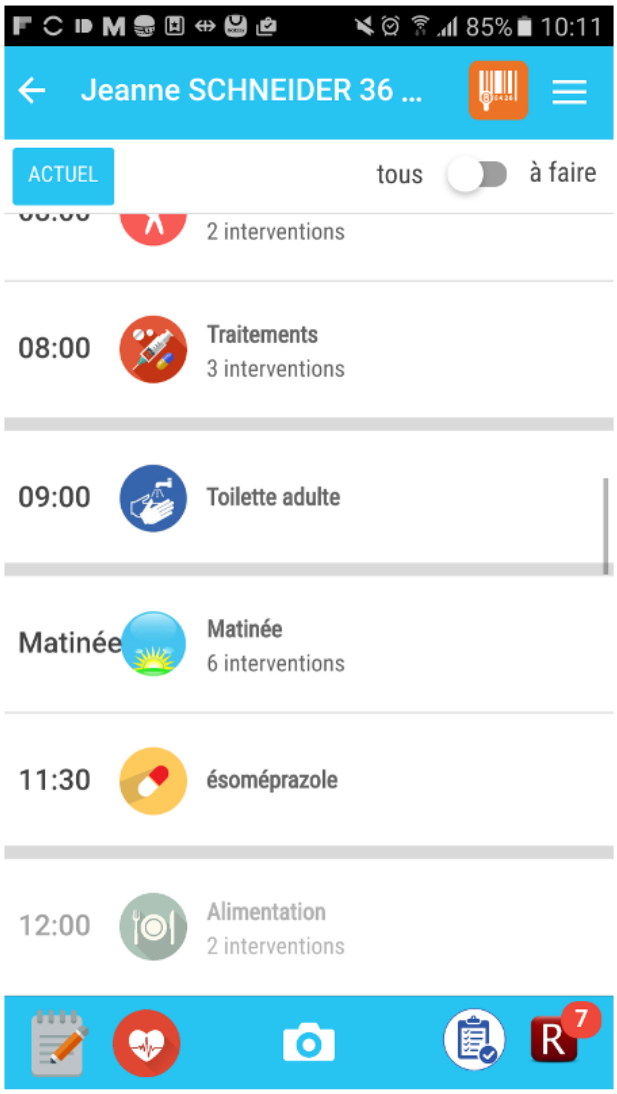


Figure 4. Screenshots of the Bedside Mobility app, input screen of clinical score.

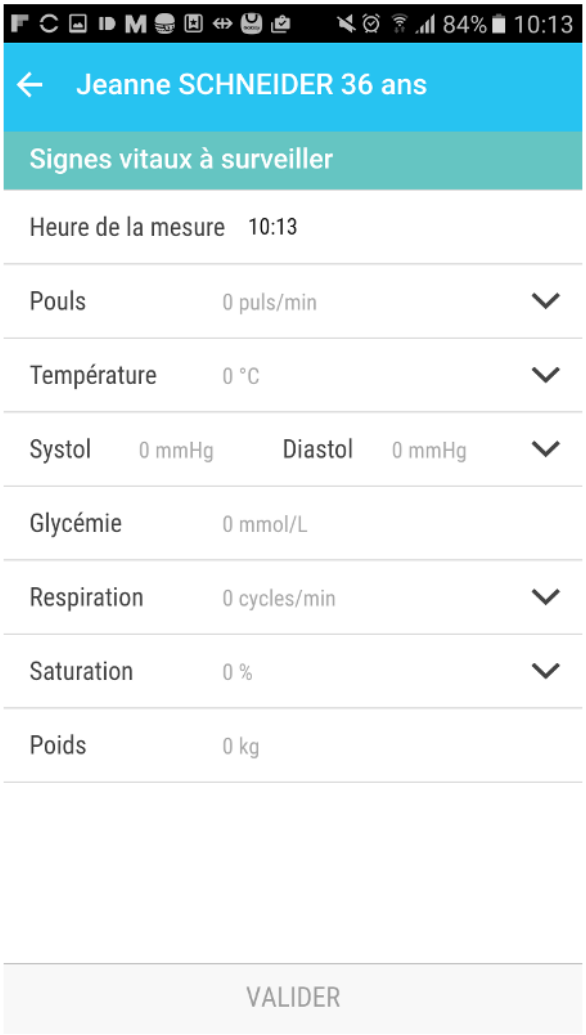
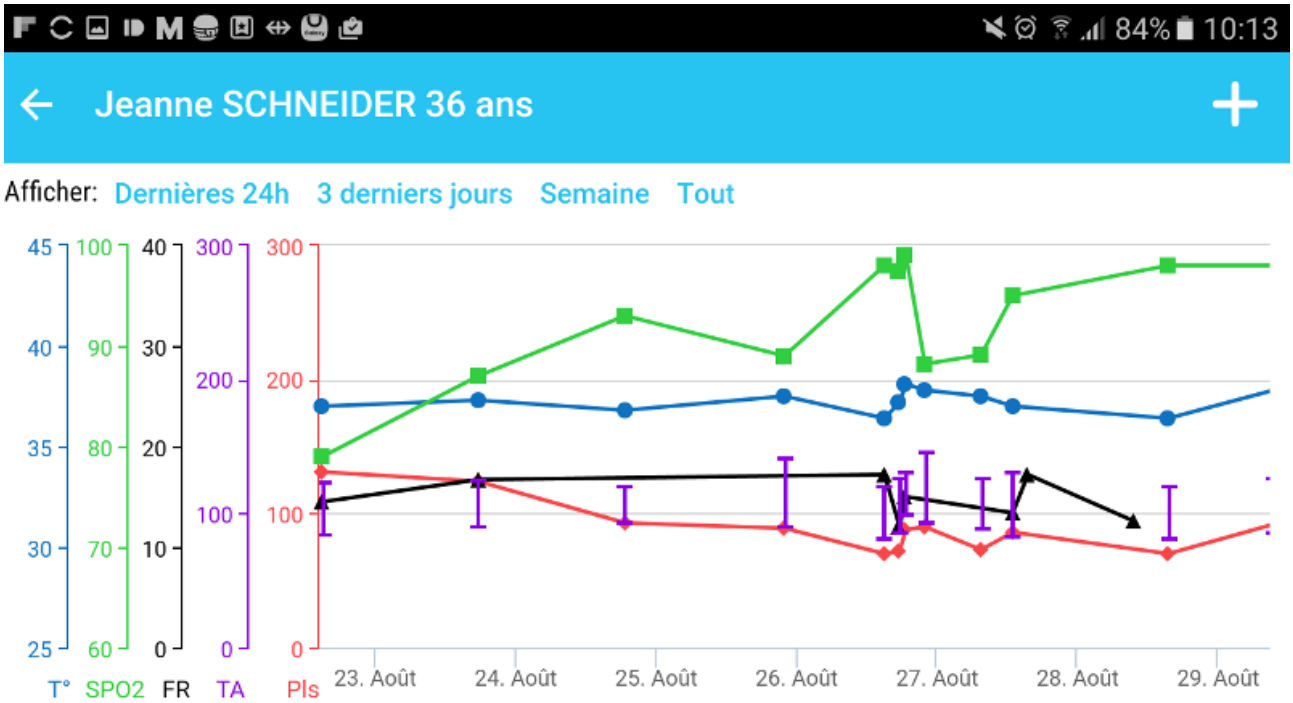


Figure 5. Screenshots of the Bedside Mobility app, vital signs graph.



Vital signs and clinical scores can be entered and visualized easily in the app; data can be entered directly from the task list to facilitate usability (Figure 4) when these assessments are planned. The app also allows users to enter these assessments in the data visualization screen when they are not scheduled (Figure 5). We chose to use vital sign graphs that are similar to the EHR graphs for easier readability.

PRN use medication (or medication "as needed") is available in another part of the app, with indications of prescribed doses and frequencies. It also records when the doses are administered during the past 24 hours, as requested by the nursing teams.

A "notes" function was implemented to allow easy note taking during the day for subsequent production of progress notes, charting, or handoffs. A note can also be created from the to-do list section.

Careful consideration was given to the design of the app to provide efficiency and usability, while offering the range of frequent tasks used in clinical documentation from a computer-based EHR.

Development Cycles

Technical Considerations

From an architectural point of view, we set up a client-server architecture. Our server, hosted on a Java Beans Open Source Server of the hospital infrastructure, is programmed in JAVA and is responsible to ensure communication between the app and the exposed services of our local EHR. The communication is proprietary using Representational State Transfer (REST) or XML messages and allows access to all the necessary information in a standardized way. Between our server and the client, a proprietary REST or JavaScript Object Notation exchange protocol was set up to simplify the interpretation of the data at the client side. The client is programmed in HTML5 or JavaScript using the Angular or Ionic framework. This framework ensures its adaptation at minor costs on different operating systems and allows access to the core functionalities of the devices.

Security Issues

The transition from laboratory to field test confronted us to questions regarding the access to the clinical data. As mentioned above, one of the priorities in developing our app was to provide efficiency and high usability. This implied that we needed to find a way to simplify the user authentication procedure without compromising overall security. Another constraint was the sharing of devices among the members of a nursing team (over 20 nursing staff per team), which therefore excluded the possibility of using biometric authentication methods. The most sensitive question was how to restrict access to authorized information in case of theft of the device.

Our current solution is the result of a Delphi process with the IT security experts of our institution. We developed a solution with a combination of Personal Identification Number (PIN) code, institutional password, and location beacons (Figure 6).

Figure 6. State diagram of the authentication process. Each state is composed of 3 substates: the phone (top line), and app (middle line) locking, as well as the user login (bottom line). PIN: Personal Identification Number.

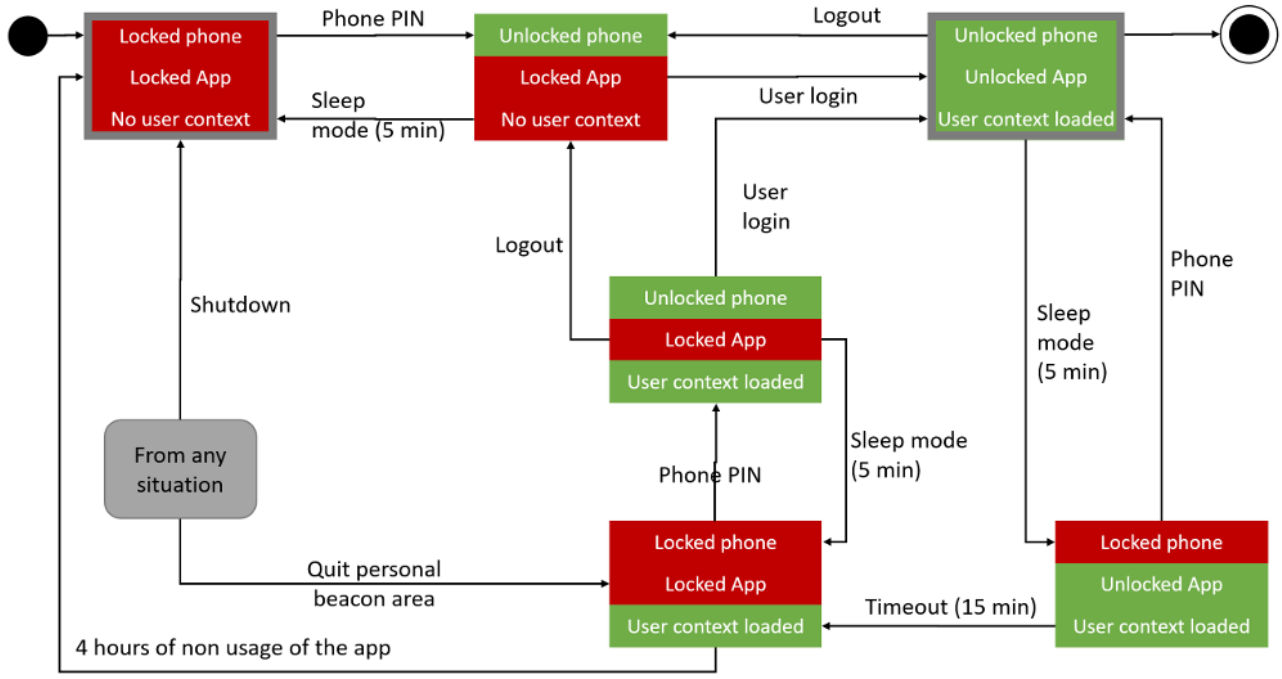


Table 3. Participant demographics (N=10).

Characteristics	Statistics, n (%)
Age (years)	
21-30	3 (30)
31-40	5 (50)
41-50	2 (20)
Gender	
Female	6 (60)
Male	4 (40)

At the initial state, the phone is locked, and no one is logged into the app. The user has to first unlock the phone using a PIN that is common to all devices for a ward. This PIN prevents access to the device in case of theft, as it is usually the case for every smartphone. As the PIN is shared between all ward users, it is not sufficient to ensure a personal access to the app. To ensure this personal access, each user has to log into the app using their institutional password, which associates a user identity to its authorization. Once logged in, the user can access clinical information. In the case of sleep mode, the phone gets locked, but the app does not lose its context. The user must then unlock it using the device PIN to get back to the previous context. If the phone remains unused for a longer period (>15 min), the app is locked; this means that the user has to enter both the PIN and their institutional password to retrieve their session. The user context remains loaded until the user is logged off by choice or if the app remains unused for more than 4 hours. In all cases, if the phone leaves the ward area covered by beacons, the app logs off and the phone gets locked. The user must then restart the whole authorization process and reset its context.

Functional Evaluation

Functional evaluation consisted of setting a list of likely use cases, which allow the tester to assess the proper behavior of the apps. Each functional test was composed of an objective, a precondition, a list of execution stages, input data, expected results, actual results, test report, and name of the testers. We ended up with 15 different tests going from user identification, patient identification, interactions with interventions, and so on.

Evaluation

Usability Test

A total of 10 volunteer nurses from medicine and surgery wards took part in the usability test (Table 3). Overall, user satisfaction with the app was high, with a SUS of 75 (SD 16.5), as presented in Figure 7.

Overall, 9 out of 10 participants managed to accomplish the 10 required tasks, even though some tasks took more time for some users. Table 4 provides an overview of the difficulties encountered.

The most difficult tasks for the participants were the review of interventions, canceling an intervention, and the validation of PRN meds.

Overall, 1 user encountered several difficulties. The user completed 3 of the tasks without help, but had navigation issues and had trouble understanding the icons in the app. The interpretation of certain icons was also reported as a source of error or delay among other participants. Moreover, two of the other participants made errors during the test by clicking on the wrong buttons, but spontaneously corrected them. All potential sources of error or delay were revised after completing the testing, and the project team tested the revised version of the app.

Figure 7. The system usability scale score of the Bedside Mobility app.

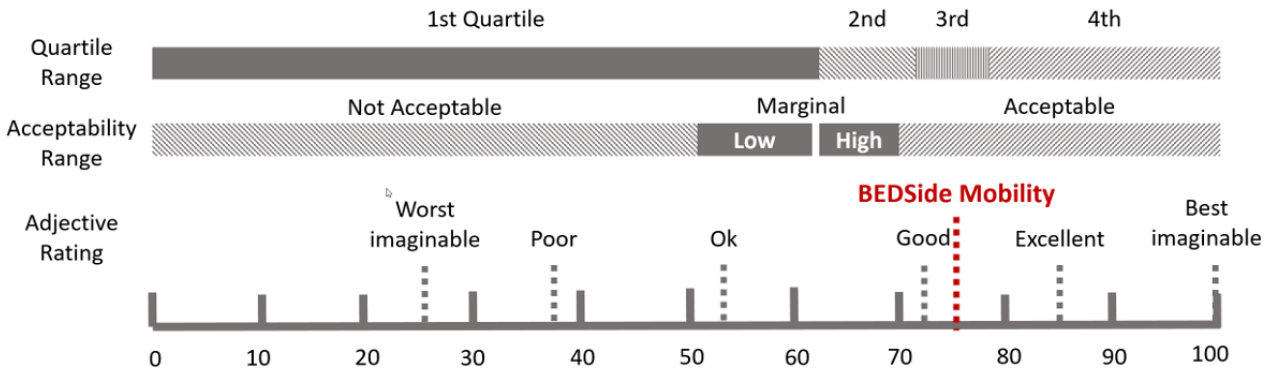


Table 4. Identified shortcomings and their correction measures.

Identified shortcomings	Correction measures
Miscomprehension of the clinical scale icon	Identification with the users of a more appropriate icon to represent clinical scale
Unexpected navigation of the back button when an intervention is open	Modification of the navigation mechanism by closing the intervention when opened rather than returning to the previous page
Canceling the validation of an intervention	Improved explanations before app use can help avoid this confusion
Inconsistent implementation of the functional design validating the administration of a pro re nata drug	Integration of similar validation mechanism to administer PRN drug using consistent icons

Supervised Field Testing

Supervised field testing took place after the second cycle of development. The local ethical committee exempted this study of authorization due to its nature (quality improvement project) and the absence of nominative data collection. Overall, 1 investigator conducted daily observations over 2 weeks of about 5 different nurses per ward to see how they interacted with the tool. This test underlined the necessity to train the users adequately and to explain the concepts underlying the app (ie, it was not intended to replace the laptops on the WOWs) and helped to identify various bugs and inconsistencies, thus driving small improvements. After each field test, the bugs were fixed and the evolutions were integrated into preparation of the next field test.

Pilot Test in the Wards

The study in the surgical unit took place for 5 weeks in July 2017 in the surgical ward and for 5 weeks in August 2017 in the medical ward. The first week was used to collect baseline data of workflow and EHR documentation. The fifth-week observations also included app use. During the study period, 27 nurses and nursing assistants used the app 427 times in the surgical ward and 23 participants used the app 239 times in the medical ward (Figure 8).

Participants were asked to verify their data entry before ending their shift to ensure the quality of clinical documentation. We explained that this was an initial pilot test with uncertain future, as the results would allow the hospital policy makers to decide whether to pursue the deployment of mobile devices.

Technology Acceptance

Our tailored UTAUT questionnaires were completed on a voluntary basis. In total, in the 2 wards, 16 care providers (8 in each ward) responded to the questionnaire. Responders were aged between 25 and 58 years (mean age 37 years). Table 5 provides a summary of the UTAUT results.

The questionnaire results revealed that the users considered the app easy to use, with a mean score of 5.6 for the effort dimension. The promotion of the app usage by the institution was evaluated as satisfactory, with a mean of 5.4. In terms of motivators of app use, the influence of the hierarchy for app use was clearly present (average of 5.9), whereas the influence of coworkers was much lower (average of 3.9).

Figure 8. Daily usage per ward.

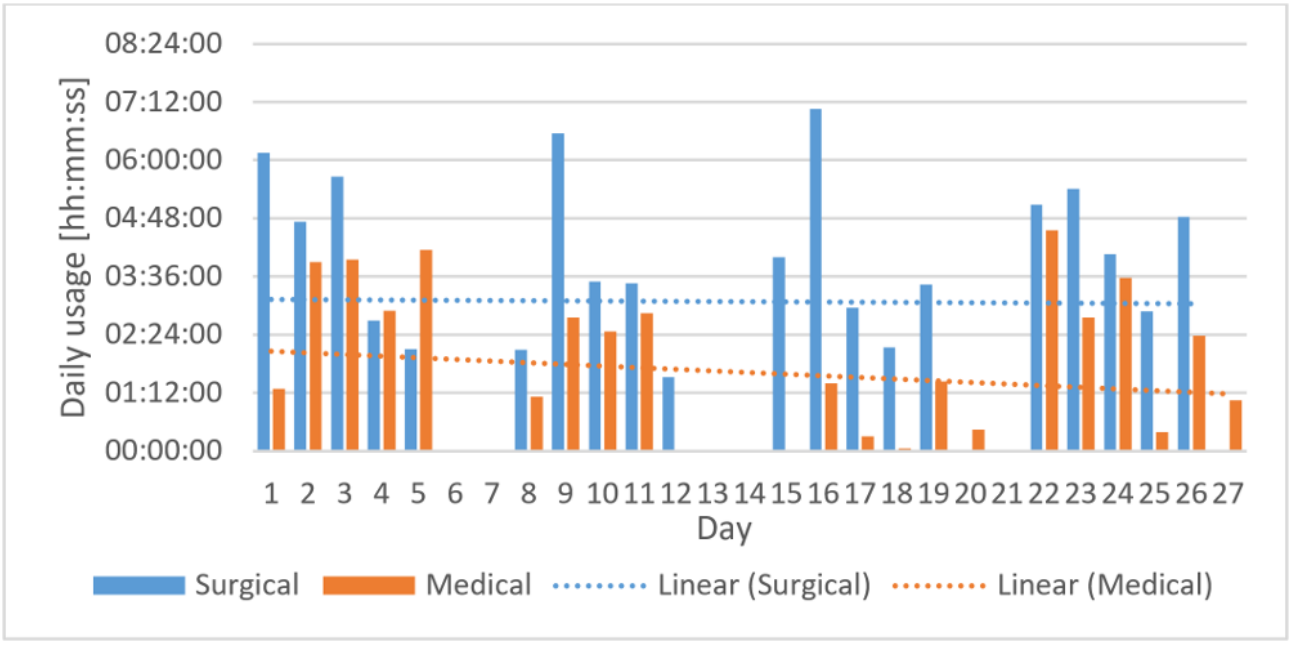


Table 5. Mean and SD in the 5 dimensions of our tailored unified theory of acceptance and use of technology questionnaire (7-point Likert scale).

Measure	Statistics, mean (SD)
Performance	3.7 (1.8)
Effort	5.6 (1.4)
Influence	5.4 (1.3)
Condition	4.7 (1.7)
Intention	4.3 (2.3)

The conditions facilitating the usage were also considered adequate with an average score of 4.7. The weakest score was for the performance dimension, with a mean score of 3.7. The question "Do you think the BEDside app increased your productivity?" received the lowest score, with an average of 3. Intention to use the app in the future obtained an average score of 4.5.

Discussion

Use of Software Development Life Cycle Model

This study presents the development lifecycle of an app to support nursing bedside care, guided by an SDLC model that is adapted to the complexity of a health care setting. The strengths of this model are the strong involvement of end users throughout the process as well as the increasing ecology of subsequent evaluation stages. Gathering information from the end users helped to identify weaknesses in the workflow process and opportunities for supportive tools. The feasibility assessment not only involved technical issues within the EHR but also human resources, potential sources of funding, and acceptability from the users. Involving clinicians from the onset of the project helped engage them throughout the process, allowing them to make suggestions and understand feasibility and funding issues. Ultimately, we also needed to consider the potential benefits and deployment at an institutional level.

Iterative Interactions With End Users

As already reported in the literature, involving users throughout the design and development process allowed them to better understand the possibilities and constraints for the app and ensure a good adoption [42-45].

The technical team also benefited from the proximity with end users, as it allowed them to suggest and develop functionalities that corresponded better to the workflow or to usual practice. This was in contrast to the interactions with the technical team for the EHR system. End users described frustrations with long-awaited "big" EHR improvements and subsequently did not always take the time to report the "small problems" to the technical support team. Communication between the end users and the developers was facilitated through frequent planned sessions as well as through the clinical team members who spent considerable time in the wards collecting information or testing functionalities with the users. Having this close and facilitated proximity between the users and the developers was therefore greatly beneficial to both parties.

Heterogeneity of the Workflow

There is a strong variability in the nursing workflow between medical and surgical wards and even between the different surgical specialty teams (orthopedics, plastic surgery, etc). For example, differences lay in the patient monitoring and wound care, which involved more equipment in the surgical wards than in the medical wards. Integrating nursing representatives (directorate and head nurses) in the project team helped ensure that the specifications and functionalities designed were adapted to the nursing workflow, such as for clinical documentation, bedside tasks, and nursing handoffs.

Use of an Institutional Device

In our setting, the institution provided devices to the nursing teams. The advantage of this solution is a more controlled and homogeneous environment, whereas disadvantages include higher costs due to the acquisition and maintenance of a large number of devices. Furthermore, having institutional devices provides an equitable access among health care providers, without discriminating against those who are not smartphone owners. Providing a device for each caregiver can be very costly; sharing the devices within the team may be more cost-effective. However, it requires a specific authentication policy that supports several users on a single device. This is complicated as smartphones are usually considered as personal devices and therefore have poor multi-identity management. Several solutions can be considered, but the final choice should carefully balance security and usability constraints. Indeed, authentication should not be too time-consuming as smartphone usage pattern is associated with a high number of authentications during daily usage [46]. A simple authentication is likely to improve user experience but may be insufficient for professional use in a health care environment with sensitive data [47-49].

Appropriateness and Strengths of our Software Development Life Cycle Model

In health care environments, challenges to the success of HIT are largely due to nontechnical issues such as poor usability that impact communication and workflow. Therefore, involving end users all along the process is of vital importance. As most of the existing SDLC models are more focused on technical issues, we have defined our SDLC model to ensure end users' participation in the early stage of the evaluation to maximize acceptance. The different stages of evaluation are designed to maximize the usefulness of the users' involvement. Indeed, it is critical to make the best use of the care providers' time in the implementation process because including participants is often difficult due to limited time. In addition, performing evaluations of growing ecological validity allowed us to minimize the risks

associated with HIT in health care [50]. Indeed, interactions with medical data are never free of risks, and literature shows many examples of poor IT implementation that have worsened the care process [51-53]. By starting with a low ecological validity evaluation, our model helps resolve many issues at an early stage without the risk of compromising the integrity of data [54]. Indeed, the usability test helped us to detect some design problems that could have been considerable sources of errors in the field. Then, by performing limited tests under supervision, we ensured that the app was also adapted in real conditions, particularly for the existing workflow. Finally, the on-site evaluation is indispensable to ensure the acceptance of the app in real conditions and provided information about the care providers' resistance as well as the organizational problems that can hinder the deployment of the app at a larger scale.

It is worth noting that even though end users were involved early in the process, the perception of the app productivity remained pretty low. This may be due to the instructions provided to the caregivers during the study. Indeed, as data manipulations undertaken through the app could have potential repercussions on the patients' safety, we asked the participants to check the correct action or entry of the data in the EHR at the end of each shift. Therefore, care providers may have had the impression of having to repeat their work and thus had a poor perception of productivity. However, as the SDLC model optimizes the use of resources, minimizes risks, and maximizes acceptance, we can confidently recommend the use of our SDLC model in other medical settings.

Limitations

We acknowledge that health care settings, needs, and workflow can differ considerably and that our study was conducted in a single institution. Therefore, generalizability to other health care setting may be limited. The EHR system is also home-grown and also limits generalizability in terms of feasibility, cost, and complexity assessments.

We were also particularly fortunate to have support from both top- and bottom-level stakeholders, who were involved at all stages. We emphasize the importance of user involvement and feedback in iterative cycles throughout the process to help ensure that the design and development are as tailored to the needs and workflow as possible.

Conclusions

Knowledge about design and development of mHealth interventions is often scattered in the literature. In this study, we aimed to present the design and development of an mHealth intervention for caregivers according to a longitudinal methodology that ranges from the initial requirement identification process to the final product, with iterative development and testing processes. To the best of our knowledge, this paper is one of the first to describe the full process of design and development of an app in hospital settings using an SDLC model and to report its benefits and limitations. Each step of this process is necessary to ensure the creation of a useful and effective app that can truly support user needs, within a given workflow process. Although a close collaboration with clinicians throughout the entire project facilitated the development of a tailored solution, it was also important to involve all stakeholders, in particular, the IT security officers. In the health care setting, ensuring the adoption of an IT tool requires a solution that addresses the strong security constraints, while maintaining ease of use and good usability. Furthermore, we tried to anticipate how to potentially scale up this project to an institutional level, contingent on the results of the final testing phase.

Acknowledgments

The authors thank their colleagues from HUG, Pascal Ducloux, Jean-Yves Wojtasikiewicz, Pierre-André Berret, and Thomas Weinhold, who provided insight and expertise that greatly assisted the research. The authors also thank Professor Arnaud Perrier, CMO, for his support throughout this project. This research was funded by the "Private Foundation of the University Hospitals of Geneva."

Conflicts of Interest

None declared.

References

1. Lin AC, Chang MH, Chen MY, Yu T, Chou LC, Tsai D, et al. Mobility patterns of doctors using electronic health records on iPads. In: Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics. New York, New York, USA: ACM Press; 2007 Presented at: International Conference on Bioinformatics, Computational Biology and Biomedical Informatics; September 22-25, 2013; Washington, DC, USA p. 933-939. [doi: 10.1145/2506583.2512366]
2. Moody LE, Slocumb E, Berg B, Jackson D. Electronic health records documentation in nursing: nurses' perceptions, attitudes, and preferences. Comput Inform Nurs 2004 Dec;22(6):337-344 [FREE Full text] [Medline: 15602303]
3. Fahimi F, Abbasi Nazari M, Abrishami R, Sistanizad M, Mazidi T, Faghihi T, et al. Transcription errors observed in a teaching hospital. Arch Iran Med 2009 Mar;12(2):173-175 [FREE Full text] [Medline: 19249889]
4. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. Transl Behav Med 2014 Dec;4(4):363-371 [FREE Full text] [doi: 10.1007/s13142-014-0293-9] [Medline: 25584085]

5. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T 2014 May;39(5):356-364 [FREE Full text] [Medline: 24883008]
6. Samples C, Ni Z, Shaw RJ. Nursing and mHealth. Int J Nurs Sci 2014 Dec;1(4):330-333 [FREE Full text] [doi: 10.1016/j.ijnss.2014.08.002]
7. Duhm J, Fleischmann R, Schmidt S, Hupperts H, Brandt SA. Mobile electronic medical records promote workflow: physicians' perspective from a survey. JMIR Mhealth Uhealth 2016 Jun 06;4(2):e70 [FREE Full text] [doi: 10.2196/mhealth.5464] [Medline: 27268720]
8. Kim J, Lee Y, Lim S, Kim JH, Lee B, Lee JH. What clinical information is valuable to doctors using mobile electronic medical records and when? J Med Internet Res 2017 Dec 18;19(10):e340 [FREE Full text] [doi: 10.2196/jmir.8128] [Medline: 29046269]
9. Prgomet M, Georgiou A, Westbrook JI. The impact of mobile handheld technology on hospital physicians' work practices and patient care: a systematic review. J Am Med Inform Assoc 2009;16(6):792-801 [FREE Full text] [doi: 10.1197/jamia.M3215] [Medline: 19717793]
10. Carlson E, Catrambone C, Oder K, Nauseda S, Fogg L, Garcia B, et al. Point-of-care technology supports bedside documentation. J Nurs Adm 2010 Sep;40(9):360-365. [doi: 10.1097/NNA.0b013e3181ee4248] [Medline: 20798618]
11. Gachet D, De Buenaga M, Puertas E. Mobile access to patient clinical records and related medical documentation. In: Proceedings of the International Conference on Ubiquitous Computing: Applications, Technology and Social Issues. 2006 Jan Presented at: International Conference on Ubiquitous Computing: Applications, Technology and Social Issues; May 8-13, 2016; Madrid, Spain URL: https://pdfs.semanticscholar.org/d875/ba25d29152c00907780868de4624703987f6.pdf
12. Wager KA, Schaffner MJ, Foulois B, Swanson Kazley A, Parker C, Walo H. Comparison of the quality and timeliness of vital signs data using three different data-entry devices. Comput Inform Nurs 2010;28(4):205-212. [doi: 10.1097/NCN.0b013e3181e1df19] [Medline: 20571371]
13. Briggs J, Curry R. Towards the integration of near-patient collection of information for the delivery of healthcare services. ACM SIGHIT Rec 2012 Sep 01;2(2):31-35. [doi: 10.1145/2384556.2384560]
14. Yee T, Needleman J, Pearson M, Parkerton P, Parkerton M, Wolstein J. The influence of integrated electronic medical records and computerized nursing notes on nurses' time spent in documentation. Comput Inform Nurs 2012 Jun;30(6):287-292. [doi: 10.1097/NXN.0b013e31824af835] [Medline: 22411414]
15. Christino MA, Matson AP, Fischer SA, Reinert SE, Digiovanni CW, Fadale PD. Paperwork versus patient care: a nationwide survey of residents' perceptions of clinical documentation requirements and patient care. J Grad Med Educ 2013 Dec;5(4):600-604 [FREE Full text] [doi: 10.4300/JGME-D-12-00377.1] [Medline: 24455008]
16. Ammenwerth E, Spötl HP. The time needed for clinical documentation versus direct patient care. A work-sampling analysis of physicians' activities. Methods Inf Med 2009;48(1):84-91. [Medline: 19151888]
17. Hendrich A, Chow MP, Skierczynski BA, Lu Z. A 36-hospital time and motion study: how do medical-surgical nurses spend their time? Perm J 2008;12(3):25-34 [FREE Full text] [Medline: 21331207]
18. Fleischmann R, Duhm J, Hupperts H, Brandt SA. Tablet computers with mobile electronic medical records enhance clinical routine and promote bedside time: a controlled prospective crossover study. J Neurol 2015 Mar;262(3):532-540 [FREE Full text] [doi: 10.1007/s00415-014-7581-7] [Medline: 25476692]
19. Motulsky A, Wong J, Cordeau J, Pomalaza J, Barkun J, Tamblyn R. Using mobile devices for inpatient rounding and handoffs: an innovative application developed and rapidly adopted by clinicians in a pediatric hospital. J Am Med Inform Assoc 2017 Apr 01;24(e1):e69-e78. [doi: 10.1093/jamia/ocw107] [Medline: 27554824]
20. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018 May 01;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
21. Ehrler F, Wipfli R, Teodoro D, Sarrey E, Walesa M, Lovis C. Challenges in the implementation of a mobile application in clinical practice: case study in the context of an application that manages the daily interventions of nurses. JMIR Mhealth Uhealth 2013 Jun 12;1(1):e7 [FREE Full text] [doi: 10.2196/mhealth.2344] [Medline: 25100680]
22. Kaplan B, Harris-Salamone KD. Health IT success and failure: recommendations from literature and an AMIA workshop. J Am Med Inform Assoc 2009;16(3):291-299 [FREE Full text] [doi: 10.1197/jamia.M2997] [Medline: 19261935]
23. Hughes M. Do 70 per cent of all organizational change initiatives really fail? J Change Manag 2011 Dec 16;11(4):451-464. [doi: 10.1080/14697017.2011.630506]
24. Cresswell K, Sheikh A. Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review. Int J Med Inform 2013 May;82(5):e73-e86. [doi: 10.1016/j.ijmedinf.2012.10.007] [Medline: 23146626]
25. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
26. Sligo J, Gauld R, Roberts V, Villa L. A literature review for large-scale health information system project planning, implementation and evaluation. Int J Med Inform 2017 Dec;97:86-97. [doi: 10.1016/j.ijmedinf.2016.09.007] [Medline: 27919399]

27. Heeks R. Health information systems: failure, success and improvisation. Int J Med Inform 2006 Feb;75(2):125-137. [doi: 10.1016/j.ijmedinf.2005.07.024] [Medline: 16112893]
28. Ho WY. Nurses' Use of iPads for Clinical Documentation: A Usability Study. Canada: University of British Columbia; 2013. URL: https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0165730
29. Barra DC, Paim SM, Sasso GT, Colla GW. [Methods for the development of mobile applications in health: integrational review of literature]. Texto Contexto Enferm 2018 Jan 08;26(4):12 [FREE Full text] [doi: 10.1590/0104-07072017002260017]
30. Ehrler F, Lovis C, Blondon K. Addressing the complexity of mobile app design in hospital setting with a tailored software development life cycle model. Stud Health Technol Inform 2016;228:200-204. [Medline: 27577371]
31. Malik S, Nigam C. A comparative study of different types of models in software development life cycle. Int Res J Eng Technol 2017 Oct 08;4(10) [FREE Full text]
32. Rastogi V. Software development life cycle models-comparison, consequences. Int J Comput Sci Inform Secur 2015;6(1):168-172 [FREE Full text]
33. Munassar NM, Govardhan A. A comparison between five models of software engineering. Int J Comput Sci Issues 2010;7(5):94-101 [FREE Full text]
34. Nawaz A. A comparison of card-sorting analysis methods a comparison of card-sorting analysis methods. 2012 Presented at: The 10th Asia Pacific Conference on Computer Human Interaction; August 28-31; Matsue, Japan p. 583-592 URL: https://openarchive.cbs.dk/bitstream/handle/10398/8587/Ather_Nawaz_2012.pdf?sequence=1
35. Ehrler F, Weinhold T, Joe J, Lovis C, Blondon K. A mobile app (BEDSide mobility) to support nurses' tasks at the patient's bedside: usability study. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e57 [FREE Full text] [doi: 10.2196/mhealth.9079] [Medline: 29563074]
36. Borycki EM, Monkman H, Griffith J, Kushniruk AW. Mobile usability testing in healthcare: methodological approaches. Stud Health Technol Inform 2015;216:338-342. [Medline: 26262067]
37. Svanaes D, Alsos OA, Dahl Y. Usability testing of mobile ICT for clinical settings: methodological and practical challenges. Int J Med Inform 2010 Apr;79(4):e24-e34. [doi: 10.1016/j.ijmedinf.2008.06.014] [Medline: 18789753]
38. Nayebi F, Desharnais JM, Abran A. The state of the art of mobile application usability evaluation. 2012 Apr 29 Presented at: 25th IEEE Canadian Conference on Electrical and Computer Engineering; April 29-May 2, 2012; Montreal, QC, Canada. [doi: 10.1109/CCECE.2012.6334930]
39. Brooke J. SUS—A quick and dirty usability scale. In: Jordan PW, Thomas P, Weerdmeester BA, McClelland AL, editors. Usability Evaluation in Industry. London: Taylor and Francis; 1996.
40. Ehrler F, Ducloux P, Wu DT, Lovis C, Blondon K. Acceptance of a mobile application supporting nurses workflow at patient bedside: results from a pilot study. Stud Health Technol Inform 2018;247:506-510. [Medline: 29678012]
41. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
42. Johnson CM, Johnson TR, Zhang J. A user-centered framework for redesigning health care interfaces. J Biomed Inform 2005 Feb;38(1):75-87 [FREE Full text] [doi: 10.1016/j.jbi.2004.11.005] [Medline: 15694887]
43. Bendixen RM, Fairman AD, Karavolis M, Sullivan C, Parmanto B. A user-centered approach: understanding client and caregiver needs and preferences in the development of mHealth apps for self-management. JMIR Mhealth Uhealth 2017 Sep 26;5(9):e141 [FREE Full text] [doi: 10.2196/mhealth.7136] [Medline: 28951378]
44. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
45. Maher M, Kaziunas E, Ackerman M, Derry H, Forringer R, Miller K, et al. User-centered design groups to engage patients and caregivers with a personalized health information technology tool. Biol Blood Marrow Transplant 2016 Feb;22(2):349-358 [FREE Full text] [doi: 10.1016/j.bbmt.2015.08.032] [Medline: 26343948]
46. Oulasvirta A, Rattenbury T, Ma L, Raita E. Habits make smartphone use more pervasive. Pers Ubiquit Comput 2011 Jun 16;16(1):105-114. [doi: 10.1007/s00779-011-0412-2]
47. Morera EP, de la Torre Díez I, Garcia-Zapirain B, López-Coronado M, Arambarri J. Security recommendations for mHealth apps: elaboration of a developer's guide. J Med Syst 2016 Jun;40(6):152. [doi: 10.1007/s10916-016-0513-6] [Medline: 27147515]
48. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. J Med Syst 2015 Jan;39(1):181. [doi: 10.1007/s10916-014-0181-3] [Medline: 25486895]
49. Al Ayubi SU, Pelletier A, Sunthara G, Gujral N, Mittal V, Bourgeois FC. A mobile app development guideline for hospital settings: maximizing the use of and minimizing the security risks of. JMIR Mhealth Uhealth 2016 May 11;4(2):e50 [FREE Full text] [doi: 10.2196/mhealth.4424] [Medline: 27169345]
50. Keller JP. Patient safety implications with the rapid adoption of IT-based health technologies. Digit Med 2017;3(3):115-119. [doi: 10.4103/digm.digm_20_17]
51. Kim MO, Coiera E, Magrabi F. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. J Am Med Inform Assoc 2017 Dec 01;24(2):246-250. [doi: 10.1093/jamia/ocw154] [Medline: 28011595]

52. Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004;11(2):104-112 [FREE Full text] [doi: 10.1197/jamia.M1471] [Medline: 14633936]
53. Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005 Mar 9;293(10):1197-1203. [doi: 10.1001/jama.293.10.1197] [Medline: 15755942]
54. Kushniruk AW, Triola MM, Borycki EM, Stein B, Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. Int J Med Inform 2005 Aug;74(7-8):519-526. [doi: 10.1016/j.ijmedinf.2005.01.003] [Medline: 16043081]

Abbreviations

EHR: electronic health record
HIS: health information system
HIT: health information technology
HUG: University Hospitals of Geneva
IT: information technology
IV: intravenous
mHealth: mobile health
PIN: Personal Identification Number
QR: Quick Response
REST: Representational State Transfer
SDLC: software development life cycle
SUS: System Usability Scale
UTAUT: unified theory of acceptance and use of technology
WOW: workstation on wheels

Edited by CL Parra-Calderón; submitted 19.10.18; peer-reviewed by A Klingberg, JH Lee; comments to author 01.12.18; revised version received 18.01.19; accepted 12.02.19; published 11.04.19.

Please cite as:
Ehrler F, Lovis C, Blondon K
A Mobile Phone App for Bedside Nursing Care: Design and Development Using an Adapted Software Development Life Cycle Model
JMIR Mhealth Uhealth 2019;7(4):e12551
URL: http://mhealth.jmir.org/2019/4/e12551/
doi:10.2196/12551
PMID:30973339

Original Paper

A Digital Cognitive Aid for Anesthesia to Support Intraoperative Crisis Management: Results of the User-Centered Design Process

Stefanie Schild[1], MSc; Brita Sedlmayr[2], PhD; Ann-Kathrin Schumacher[1], MSc; Martin Sedlmayr[3], PhD; Hans-Ulrich Prokosch[1], PhD; Michael St.Pierre[4], MD; German Cognitive Aid Working Group[5]

[1]Department of Medical Informatics, Biometrics and Epidemiology, Chair of Medical Informatics, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany

[2]Center for Evidence-based Healthcare, Carl Gustav Carus University Hospital, Technische Universität Dresden, Dresden, Germany

[3]Carl Gustav Carus Faculty of Medicine, Institute for Medical Informatics and Biometry, Technische Universität Dresden, Dresden, Germany

[4]Anästhesiologische Klinik, Universitätsklinikum Erlangen, Erlangen, Germany

[5]Berufsverband Deutscher Anästhesisten, Nürnberg, Germany

Corresponding Author:
Stefanie Schild, MSc
Department of Medical Informatics, Biometrics and Epidemiology
Chair of Medical Informatics
Friedrich-Alexander University Erlangen-Nürnberg
Wetterkreuz 13
Erlangen, 91058
Germany
Phone: 49 9131 85 26785
Email: stefanie.schild@fau.de

Abstract

Background: Stressful situations during intraoperative emergencies have negative impact on human cognitive functions. Consequently, task performance may decrease and patient safety may be compromised. Cognitive aids can counteract these effects and support anesthesiologists in their crisis management. The Professional Association of German Anesthesiologists set up a project to develop a comprehensive set of digital cognitive aids for intraoperative emergencies. A parallel development for several software platforms and stationary and mobile devices will accommodate the inhomogeneity of the information technology infrastructure within German anesthesia departments.

Objective: This paper aimed to provide a detailed overview of how the task of developing a digital cognitive aid for intraoperative crisis management in anesthesia was addressed that meets user requirements and is highly user-friendly.

Methods: A user-centered design (UCD) process was conducted to identify, specify, and supplement the requirements for a digital cognitive aid. The study covered 4 aspects: analysis of the context of use, specification of user requirements, development of design solutions, and evaluation of design solutions. Three prototypes were developed and evaluated by end users of the application. Following each evaluation, the new requirements were prioritized and used for redesign. For the first and third prototype, the System Usability Scale (SUS) score was determined. The second prototype was evaluated with an extensive Web-based questionnaire. The evaluation of the third prototype included a think-aloud protocol.

Results: The chosen methods enabled a comprehensive collection of requirements and helped to improve the design of the application. The first prototype achieved an average SUS score of 74 (SD 12), indicating good usability. The second prototype included the following main revisions: 2-column layout, initial selection of patient type (infant, adult, or parturient), 4 offered search options, and the option to check off completed action steps. Its evaluation identified the following major revision points: add quick selection for resuscitation checklists, design the top bar and tabs slightly larger, and add more pictograms to the text. The third prototype achieved an average SUS score of 77 (SD 15). The evaluation of the think-aloud protocol revealed a good intuitiveness of the application and identified a missing home button as the main issue.

Conclusions: Anesthesiology—as an acute medical field—is particularly characterized by its high demands on decision making and action in dynamic, or time-critical situations. The integration of usability aspects is essential for everyday and emergency suitability. The UCD process allowed us to develop a prototypical digital cognitive aid, exhibiting high usability and user satisfaction in the demanding environment of anesthesiological emergencies. Both aspects are essential to increase the acceptance

of the application in later stages. The study approach, combining different methods for determining user requirements, may be useful for other implementation projects in a highly demanding environment.

(JMIR Mhealth Uhealth 2019;7(4):e13226) doi:10.2196/13226

KEYWORDS

anesthesiology; checklist; crew resource management, healthcare; emergency treatment; ergonomics; human factors; practice guideline; reference books, medical; resuscitation; user-computer interface

Introduction

Background

Intraoperative emergencies require rapid, coordinated management in time-critical and stressful settings. Anesthesiologists are faced with the twofold expectation that their situational assessment should include all possible differential diagnoses, whereas their management should always be grounded in current evidence-based recommendations. Unfortunately, stressful situations have a negative impact on human cognitive functions (eg, attention, working memory, or prospective memory) and analysis-driven decision making [1]. Challenged by an uncommon emergency, task performance can decrease even more, as humans are not optimized to retrieve rarely used information. As a result, omissions of critical steps, practice variability, and noncompliance to established guidelines increase [2].

Encouraged by promising results from other high-profile industries in which very complex problem situations have to be solved quickly, the past decade has witnessed a growing interest of anesthesiologists in paper-based or electronic cognitive aids that can counteract the possible deleterious effects of stress listed above [3,4]. Cognitive aids can remind clinicians of important diagnostic causes, guide them on the basis of current evidence-based practices through a sequence of complex steps, and prevent them from omitting key actions [5]. Furthermore, these aids should anticipate common pitfalls of the particular emergency, provide prioritized and explicit instructions to prevent them, and contain important local information (eg, phone numbers, depositories of critical drugs, and precalculated drug dosages) that may help increase the speed and fluidity of performance. A process of user-centered design (UCD) is necessary to prevent poorly designed aids from distracting and negatively affecting clinicians. Ideally, a national body should attempt, in a formal consensus approach, to create a more comprehensive set of perioperative cognitive aids than what is currently available [6].

Objectives

As there are currently no officially endorsed cognitive aids for intraoperative emergencies available in Germany, the Professional Association of German Anesthesiologists (BDA) set up a project to develop such a comprehensive set of digital cognitive aids for intraoperative emergencies. Within the scope of this study, clinicians from different university hospitals (German Cognitive Aid Working Group) worked together with human factors engineers and software developers. To ensure wide dissemination of the cognitive aid to German anesthesia departments, consideration had to be given to the inhomogeneity of information technology infrastructures within German anesthesia departments. As a result, it was decided to develop an electronic cognitive aid as a mobile health app for multiple software platforms as well as stationary and mobile devices.

In this paper, we have answered the following research questions:

1. What demands do clinicians have on the digital representation of cognitive aids for emergency situations?
2. How can we best develop a user interface for a cognitive aid to be applied in the time-critical and stressful environment of intraoperative crisis management in anesthesia?

Methods

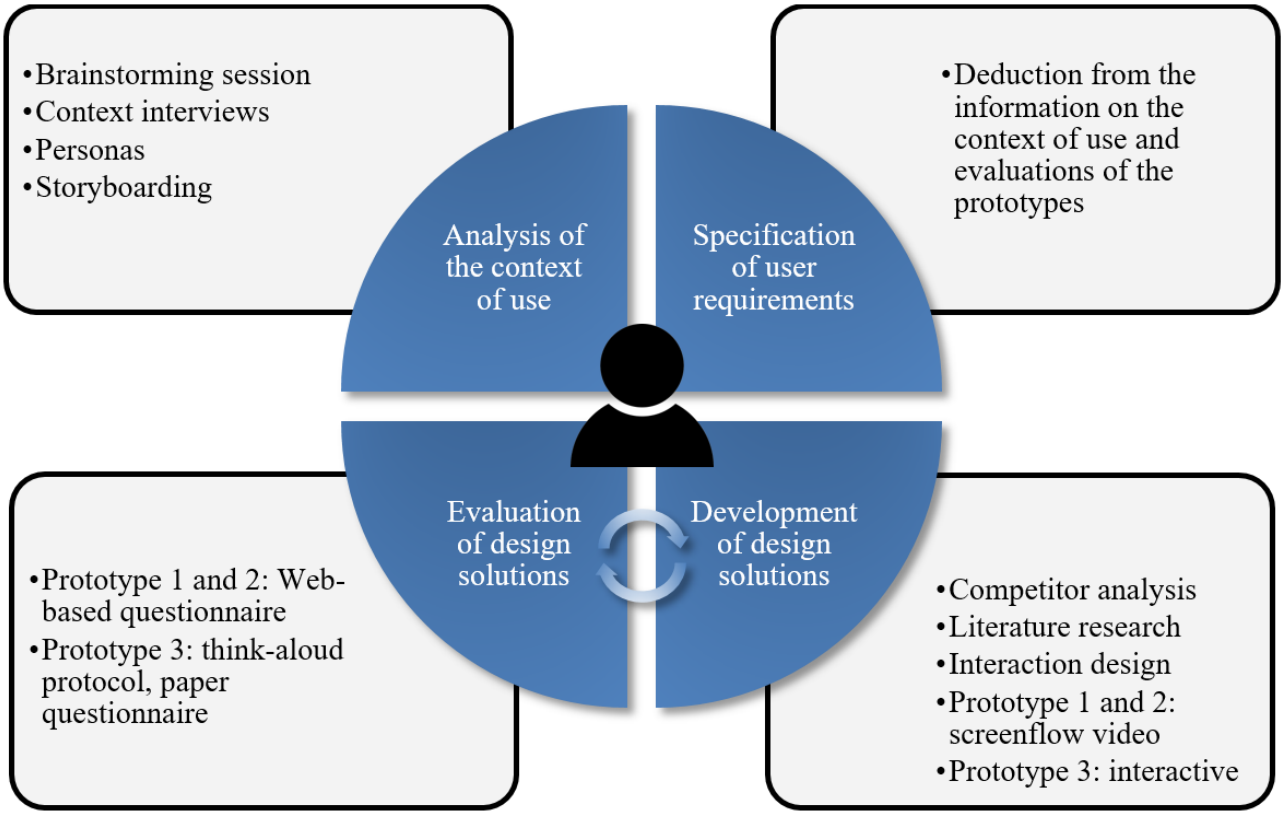
User-Centered Design

The development of this cognitive aid is based on the UCD process of the International Organization for Standardization (ISO) norm 9241-210 [7], which has recently been performed for the development of other applications in the medical context [8-10]. This process comprises the analysis of the context of use, the definition of user requirements, the development of prototypes, and tests with representative users.

For this study, a thorough analysis of the physical and organizational environment, the application context, and the technical and task-specific requirements of the end users was performed. These requirements were implemented in an initial prototype and evaluated with a user test. The evaluation was used to query the user requirements regarding their prioritization and, if necessary, to adapt them accordingly or add new derived ones. With the user requirements revised in this way, a second prototype was created, which in turn was evaluated with a user test. This iteration step was repeated with a third prototype.

The study process design [7] is presented in Figure 1. Owing to its scope, only the more significant parts of the process are described in this paper.

Figure 1. Study process design.



Analysis of the Context of Use

A brainstorming session was held with 7 members of the German Cognitive Aid Working Group present at an initial meeting to identify users' expectations regarding a digital cognitive aid and develop a first set of ideas concerning its functionality. Furthermore, context interviews with 4 anesthesiologists and 3 anesthetic nurses were conducted at the anesthesia departments of 3 project partners (University Hospital Erlangen, University Hospital Dresden, and Trauma Hospital Berlin). Personas, that is, written representations of the intended end users of the application, were developed to give a clear picture of the characteristics of the users within the project group.

On the basis of the results of the above methods, storyboards were developed for 3 different use cases (intraoperative management of malignant hyperthermia, management of ventricular fibrillation, and a local user editing the content). To visualize the respective stories, the interaction between users and the application was illustrated by images and accompanying text. The 3 storyboards were shared with all members of the project group to get general feedback on the stories. They were asked to indicate how well the application met their needs to derive additional requirements for the application. Feedback was given in free-text form and aggregated afterward. Subsequently, the storyboards were slightly modified by the additional requirements identified.

Specification of User Requirements

The collected information by the various applied methods was consolidated and recorded in individual documents. The main requirements document contained the results of the user context analysis and evaluation of the first prototype together with the specification of the respective derived and prioritized user requirements. In addition, 2 further documents included the evaluation results of the second and third prototype and the respective derived or revised requirements.

If the evaluation result of a prototype regarding the need for revision was ambiguous in some aspects, relevant issues were identified and prepared both textually and visually. Subsequently, the points of discussion were jointly reviewed with the project group until a consensus was reached.

Development of Design Solutions

An initial interaction design was developed that described the functions and processes in the interaction between the system and user. For the development, requirements identified by the analysis of the context of use and a literature research as well as international recommendations for navigation and user dialogue design [11-13] were taken into account. In addition, the results of a competitor analysis and existing paper-based checklists served as sources of information. The medical content was based on a checklist provided by the German Cognitive Aid Working Group for the intraoperative emergency of malignant hyperthermia. On the basis of the initial interaction design, the first prototype was built in *Balsamiq Mockups 3* (Balsamiq Studios, LLC) [14]. Subsequently, a demonstration video was created, describing the individual screens of the prototype regarding their functionality and navigation options, and was made available online.

A revised interaction design was developed, considering the results of the evaluation of the first prototype and the specifications of the main requirements document. On the basis

of the revised interaction design, the second prototype was built using *PowerPoint*. The development was realized by the creation of a separate slide for each view and each function call within the application. Furthermore, an audio track was recorded for every slide, describing the respective screen regarding the functionality and navigation options. Subsequently, 3 separate demonstration videos were produced (for desktop/tablet general structure, desktop/tablet exemplary run with checklists, and smartphone general structure), which were made available online.

The third prototype was created for tablets. The process was identical to the second prototype concerning the graphical generation. This time, however, most of the elements were hyperlinked to other slides, allowing logical navigation between the different screens of the prototype. As a result, once the PDF had been opened with a PDF reader, the user could interactively navigate the prototype.

Evaluation of Design Solutions

Web-Based Questionnaire

To evaluate the first and second prototype, corresponding Web-based questionnaires were developed. Except for the open questions, the items were answered on a 5-point Likert scale, ranging from 1=strongly disagree to 3=neither agree nor disagree (neutral) to 5=strongly agree. For some items, respondents were asked to justify their response should the rating be neutral or negative (ie, rating 1 to 3). The evaluation of the first prototype included the System Usability Scale (SUS) [15,16]. The final questionnaires (adapted from the German versions) can be found in Multimedia Appendix 1.

For evaluation purposes, the developed questionnaires were provided via the platform *SoSci Survey* [17], including a link to the respective demonstration videos. The links to the SoSci Survey portal were distributed to the German Cognitive Aid Working Group via email. Project partners were encouraged to forward the links to interested colleagues to increase the number of respondents. For the analysis, the average rating and the SDs were calculated for all closed questions. The answers to open questions were grouped by topic, and the number of participants mentioning the topic was counted. The identified issues to improve the cognitive aid were prioritized to indicate the most urgent areas for revision. The prioritization was based on the frequency of mention, the severity of the problem, and the requirements to better adapt the application to the medical workflow without overloading it. The results of the evaluations were summarized in tables and text form, and aspects for revision of the application were listed by thematic classification.

Think-Aloud Protocol and Paper Questionnaire

The third prototype was evaluated by means of a laboratory-based think-aloud protocol as a well-established method to capture all actions and thoughts when using a system [18-20]. The think-aloud study was led by a researcher of the Chair of Medical Informatics and conducted at a single institution (Department of Anesthesiology, University Hospital Erlangen). The department's library provided an undisturbed and quiet environment for the study. The related institutional review board approved the investigation. Participants were end users who were neither familiar with the project nor with the interface design. The individual meetings were scheduled to last for about 20 to 30 min.

All participants were introduced to a medical emergency (malignant hyperthermia) and asked to imagine themselves being in the situation. In addition, the test persons were confronted with 4 tasks that had to be solved using the prototype. After all test conditions were explained, participants were asked to work on the tasks in the given order. At a test session, all verbal statements and screen actions were videotaped via the device.

After completing the think-aloud protocol, a test person was asked to complete a paper questionnaire related to the previous use of the prototype. The questionnaire included the SUS. Following the evaluation, the recorded videos were transcribed and the results were thematically summarized and compiled together with the evaluation of the questionnaires in tabular form.

The test instructions and tasks for the think-aloud protocol as well as the questionnaire (adapted from the German versions) can be found in Multimedia Appendix 2.

Results

Analysis of the Context of Use

Anesthesiologists in the inpatient and outpatient area were identified as the primary target group for the digital cognitive aid. Anesthesia nurses as part of the anesthesia team are considered secondary users. Figure 2 shows an excerpt of the storyboard for emergency malignant hyperthermia (adapted from the German version), which describes the use of the cognitive aid on a tablet device. In the presented scenario, an anesthesia nurse takes on the role of a reader.

Specification of the User Requirements

The main requirements document, including the results of the user context analysis and evaluation of the first prototype, has 52 pages and is available in German only. The document can be requested from the corresponding author on reasonable demand.

Figure 2. Scenario of the cognitive aid (use case malignant hyperthermia, use of the application via a tablet device).

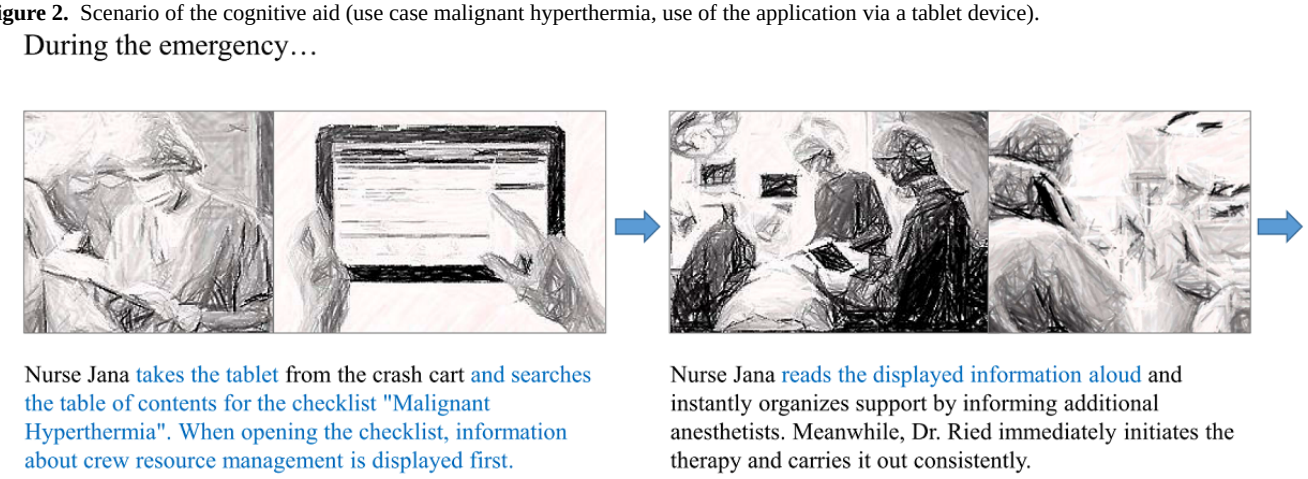


Figure 3. First prototype start page (only available in German).

Development of Design Solutions

First Prototype

On the basis of the initial interaction design, the first prototype of the digital cognitive aid was created as individual screenshots of a desktop application. At the time of the creation, colors were deliberately used very sparingly, as in this first step the focus was on functionality and screen layout of the cognitive aid. Figure 3 illustrates the start page of the application. From here, the respective emergency information can be found via the different search options offered on the left (keyword search, search by the airway, breathing, circulation, disability, exposure (ABCDE) approach, alphabetical search, search via body navigator, and search for symptoms). Once a checklist is opened, this area can be used to navigate through various sections of the emergency information. The central and largest area of the 3-column page structure always displays the essential information. Here, the user receives a short description of the various available functions. This is also the section where the search results are displayed, and after opening a checklist, the corresponding emergency action steps are displayed. The content of the right-hand area is variable: on the start page, the area has a chronological function and displays the recently accessed checklists. Once a checklist is called up, additional information about the emergency action steps is displayed at this position. In the bar at the top right of the screen, the patient's weight can be entered for individual drug dosage suggestions and an emergency call list can be accessed.

Second Prototype

The second prototype was built as a desktop or tablet and smartphone version. The prototypes were intentionally kept as similar as possible and differed only in parts of the navigation and screen layout. The revised interaction design included the following major changes compared with the initial interaction design: 2-column layout for the desktop or tablet version, selection of patient type (infant, adult, or parturient) before selecting a search option, 4 offered search options (keyword or free-text search, alphabetical search, ABCDE approach search, or body navigator search), possibility to check off completed

action steps, illustration of checklist content by different tabs, and display of the elapsed time since opening a checklist.

Third Prototype

The third prototype was modified in accordance to the revised and supplemented requirements resulting from the evaluation of the second prototype. Figure 4 illustrates the start page of the third prototype with the opened alphabetical search. In the left area, different search options to find emergency checklists can be selected, and except for the keyword search, the patient type has to be chosen first. On the right, a corresponding letter can be marked, to which the respective search results are listed in the area below.

An example of how a checklist (malignant hyperthermia) is displayed in the cognitive aid prototype is shown in Figure 5. The checklist title is centered at the top, and the time elapsed since the checklist was opened is positioned to the left, together with the selected patient type and optional patient weight. To the right, an emergency call list and the menu icon are depicted. In the middle of the screen, the emergency action steps are displayed on the left side, with different sections being selectable via the tabs above. Completed action steps can be marked as performed by ticking off. Additional information on individual items is provided in the right-hand area. Furthermore, related symptoms and differential diagnoses can be accessed via tabs on the right. In the bottom bar, the navigation path to the checklist search and information concerning the last update of the checklist content are displayed.

Additional screenshots of the third prototype (adapted from the German version) can be found in Multimedia Appendix 3.

Figure 4. Third prototype, patient type infant selected, alphabetical search (adapted from the German version).

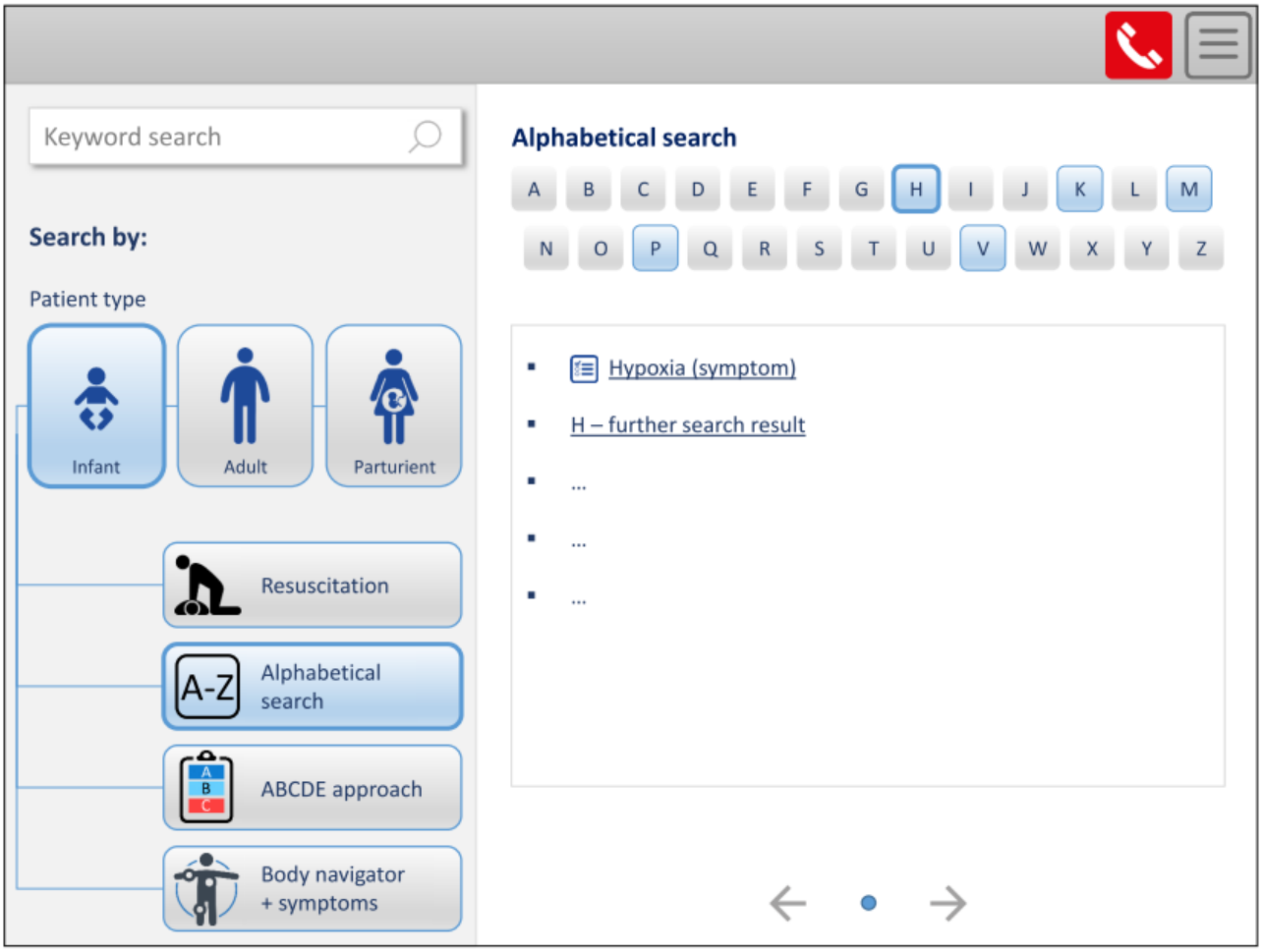
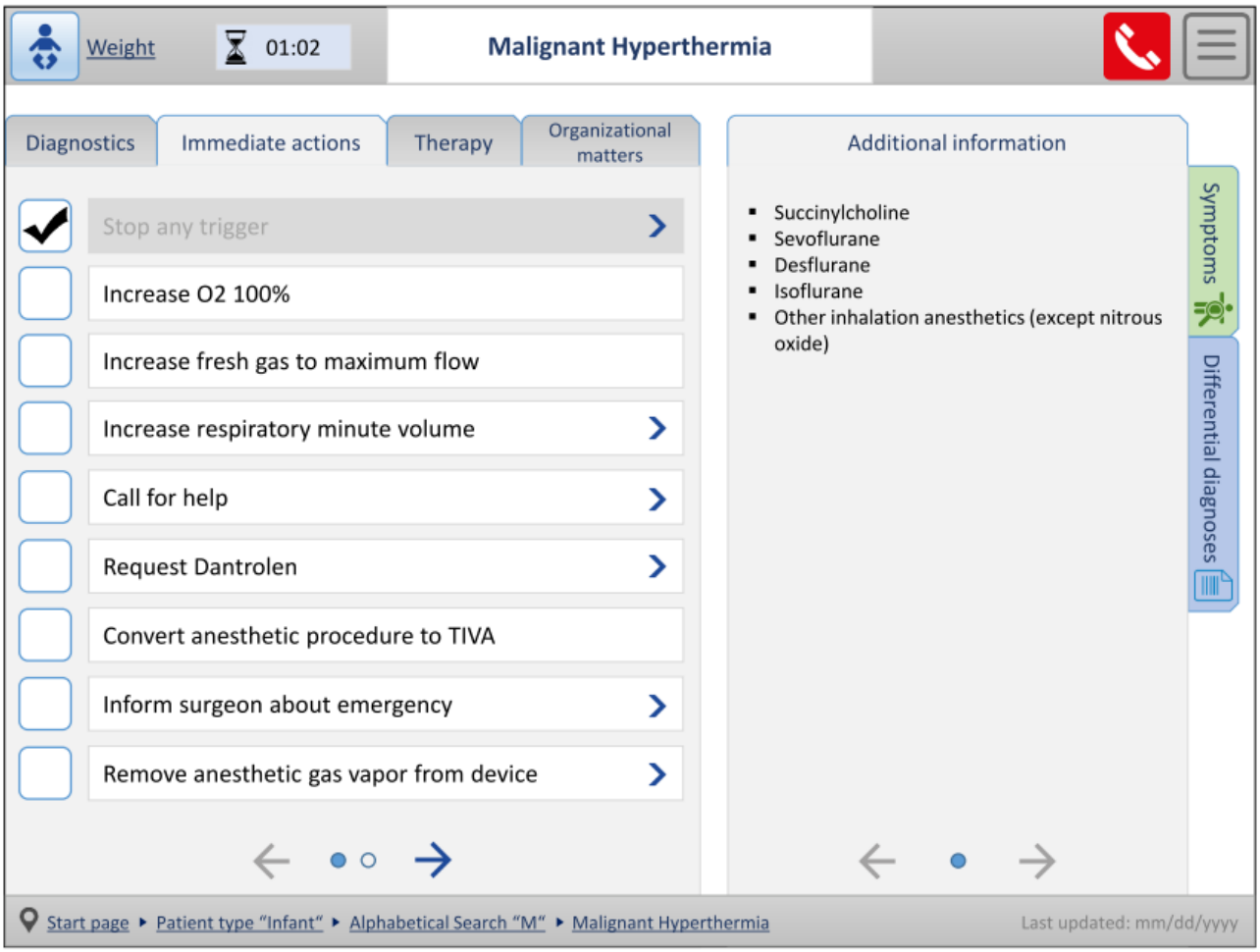


Figure 5. Third prototype, checklists view malignant hyperthermia (adapted from the German version).



Evaluation of Design Solutions

First Prototype: Web-Based Questionnaire

Sample description: A total of 12 anesthesiologists participated in the Web-based survey, whose affiliation was either the Hannover Medical School, the Hospital Munich, the University Hospital Erlangen, the Trauma Hospital Berlin, or not specified. The participant structure consisted of 5 senior physicians, 6 resident physicians, and 1 specialist, with an average professional experience of 12 years (range: 1 to 25 years). Half of the participants had already gained experience with paper checklists, but no comparable systems or applications had been used so far. The general computer literacy level was rated medium to high.

The evaluation of the SUS scores resulted in an average value of 74 (SD 12), indicating good usability of the prototype. Thereby, 2 participants rated the application less than 60 (considerable usability problems), 6 participants between 60 and 80 (borderline to good usability), and 4 participants over 80 (good to excellent usability).

The analysis of the questionnaire resulted in a prioritized list of identified topics and requirements to be clarified. After discussion with the project group, the following final list of modifications and additional requirements emerged:

- The additional information should continue to be listed in the right-hand area.
- No scrolling should be offered, as otherwise the user could lose track or miss important points.
- A combination search of symptoms is not desired as the application is not intended to be a decision support system.
- The ABCDE approach and symptom search provide redundant information. As a result, the symptom search should be completely removed.
- The body navigator should not allow the simultaneous selection of multiple organ systems.
- An additional subdivision of the search into the patient categories *Infant*, *Adult*, and *Parturient*, intended as a preliminary query right after the start of the application, should be included. After selecting one of these categories, an automatic query of the patient's weight should appear, which is only optional.
- On the start page, one of the search categories should be displayed in the middle area instead of a short explanation of the individual functions.

Table 1. Average time required per task during the think-aloud protocol.

Task	Time (seconds), mean (SD)
Find the checklist for malignant hyperthermia using the alphabetical search and open it	19 (11)
Mark the first 3 emergency action steps as completed	6 (4)
Navigate back to the start page and use the body navigator to find other emergencies or symptoms that can cause circulatory problems	44 (8)
Open the malignant hyperthermia checklist and inform yourself about possible differential diagnoses via the respective tab	20 (6)

Second Prototype: Web-Based Questionnaire

Sample description: A total of 8 anesthesiologists participated in the Web-based survey, whose affiliation was either the Hannover Medical School, the University Hospital Erlangen, the University Hospital Heidelberg, or not specified. The participant structure consisted of 2 senior physicians, 4 resident physicians, and 2 specialists, with an average professional experience of 10 years (range: 1 to 25 years). On average, respondents rated their computer literacy level as medium (advanced user). Furthermore, 2 of the participants already worked with paper-based checklists and 5 with electronic checklists (eg, apps or own checklists of the university hospital). In case of an emergency, the intranet (n=3) or colleagues (n=2) were mainly used for further information. The desktop computers (n=7) or smartphones (n=6) are mainly used to search for information in everyday life. Furthermore, the participants preferred to use a digital cognitive aid for the desktop computer (n=6) and smartphone (n=6).

The page layout, design, search options, positioning of the elements, and the way the information is presented were predominantly positively evaluated for both prototype variants (desktop or tablet and smartphone). Especially, the uniform design of the different variants was appreciated, as it allowed easy orientation and comparable handling across different devices. The presentation of an open checklist was perceived to be helpful, as it resembled an open book with indexes.

The revised requirements that resulted from a joint discussion in the project group on the results of the questionnaire were, among others, the following:

- Add quick selection for cardiopulmonary resuscitation checklists to the search options.
- When selecting an organ system in the body navigator, the associated symptoms as well as the corresponding checklists should both be listed.
- Crew resource management (CRM) and literature tab should only be available via the menu button.
- The size of the top bar and the tabs should be slightly increased.
- Tabs and times should additionally be marked with pictograms.
- A stopwatch should only be displayed when one of the resuscitation emergencies is selected.

Third Prototype: Think-Aloud Protocol and Paper Questionnaire

Sample description: A total of 9 anesthesiologists participated in the think-aloud protocol and completed the questionnaire. The participants had an average professional experience of 8 years (range: 1 to 20 years). All respondents rated their computer literacy level as medium. Furthermore, 6 of the participants already worked with paper-based checklists and 4 with electronic checklists. The majority stated that they were familiar with the use of a tablet in everyday life (mean 3.33 [SD 1.56]). It was almost a unanimous agreement among participants that they could well imagine using a tablet in the operating room (mean 4.44 [SD 0.96]).

The evaluation of the SUS scores resulted in an average value of 77 (SD 15). One participant rated the application less than 60, 4 participants between 60 and 80, 4 participants over 80, and 1 of them 95 (almost perfect application).

Overall, the application was rated positively by the participants regarding the concept (n=2), design (n=3), function (n=2), intuitiveness (n=5), and structure (n=3). Furthermore, it was easy to understand after a short practice (n=3). Participants noticed that the button for start page navigation was too small (n=4), a home button was missing (n=3), the areas of the body navigator were confusing (n=4), training was necessary (n=2), and too many clicks and tabs existed (n=2).

The participants needed an average of 88 seconds (SD 19) to complete all tasks set for the think-aloud protocol. The results of the individual tasks are listed in Table 1.

Discussion

Principal Findings

Discussion of Results

This study addresses our research questions on the clinicians' demands on the digital representation of a cognitive aid for emergency situations and how its user interface can be best developed that meets user requirements, is highly user-friendly, and is easy to navigate.

Within a UCD process, requirements were identified and refined step by step, resulting in a user-centered development of a digital cognitive aid for intraoperative crisis management in anesthesia.

Storyboards were developed to create a common vision for developers and end users about the context of use and enabled the specification of the first user requirements. The content for the storyboards was based on initial requirements defined in a

brainstorming session with the project group and context interviews that identified the characteristics of future users as well as the peculiarities of the anesthetic workplace. In addition, a competition analysis and literature search contributed to the development of a first interaction design, facilitating the creation of a first prototype. A user evaluation of the first prototype helped to define and prioritize further requirements. Considering these new requirements, a revised interaction design was developed. On the basis of this design, a second prototype was created, which in turn was evaluated in a user test. Subsequently, this process was repeated with a third prototype.

The evaluation of the SUS scores of the first and third prototype proved that the iterative process described above improved usability. Furthermore, the number of requirements to be revised was reduced from step to step. The evaluation of the third user test also revealed that the application was highly intuitive even without previous instruction. The search options allowed the user to find a checklist in a relatively short time and to check off the action steps quickly and easily. In addition, some of the test persons indicated the feeling of being familiar with the application after a short training period. This test result provides a solid basis for a subsequent test of the application under real-life conditions, associated with the high stress of an emergency situation. For this purpose, the project plan stipulates the evaluation of the cognitive aid during simulated scenarios of an intraoperative crisis. This evaluation may identify further requirements for the application that could not be determined with the previous evaluation methods.

The development of the cognitive aid needs to focus primarily on the main user of the application, which in the context of the German health care system would be the physician anesthesiologist. In contrast, anesthesia nurses as part of the anesthesia team play an assisting role and are considered as secondary users. However, the usage requirements identified were largely identical for both user groups. Major differences between both user groups only exist with regard to the handling of the cognitive aid. As the use of a digital cognitive aid for the management of intraoperative crises is not yet established in Germany, the question is yet unresolved whether anesthesia nurses will be given the task of taking the team through the checklist (ie, as readers) or whether they will continue to merely carry out the required treatment steps. In a simulation study, the presence of a reader improved the performance in crisis situations and the completeness of task execution [21].

At present, knowledge concerning the actual use of cognitive aids for crisis management in anesthesia is limited. Chances are that only a few anesthesia departments have implemented the use of paper-based checklists in their crisis management approach. Available data are mainly derived from simulation studies [5]. Although the use of the World Health Organization Surgical Safety Checklist is mandatory in Germany [22], this type of checklist only refers to the perioperative standardization of procedures. However, for the treatment of intraoperative emergencies, there exists no such mandatory regulation. Emergency checklists are therefore only available as a voluntary hospital-based measure for patient safety.

This study was able to identify the following main requirements for a digital cognitive aid: intuitive operation and error tolerance of the design, short navigation paths, full functionality for offline use, and the provision of an educational concept for those anesthesia departments that decide to implement the cognitive aid. These results are confirmed by data from other studies. In the manual by Crosland et al for the development of digital emergency checklists, the aspects of error tolerance and navigation are already provided with short implementation conditions [23]. Marshall et al concluded that the roles for the use of a cognitive aid must be clearly defined to help all team members to perform their tasks in a coordinated order [24]. The implementation of a cognitive aid should always be preceded by appropriate introductory training: on the one hand, because users can experience the physical limits of the application in the real environment and, on the other hand, because familiarity with the application increases both effectiveness and frequency of use. To overcome the concern that an application might lose its functionality in case of internet connection failure [25], it is recommended to ensure functional independence from the internet [26].

There are not many recent publications on digital cognitive aids for crisis management in anesthesia. However, it is striking that the majority of digital emergency checklists found on this topic relate mainly to the field of resuscitation. One of these publications is the tablet-based *Decision Support Tool* application for the management of advanced life support activities developed by McEvoy et al [27]. The functionality is comparable with that of the developed third prototype. Thus, the application also features a 2-column layout and includes an initial selection of the patient type and to check off action steps. In comparison, the application mainly differs in the navigation concept, which is far less clear and is associated with poorer use of the screen space. Information about the development process or the evaluation of user satisfaction is not available. An innovative concept of the cognitive aid described in this paper is that it aims to cover all emergencies, non-normal situations, and symptoms encountered during an intraoperative crisis in the future. Along with non-normal events, the list is expected to include 80 to 100 topics. In addition, an editor functionality is considered, which would enable the user to tailor the checklists to local conditions (eg, trade names of available drugs, depositories of critical drugs, location of the nearest defibrillator, and emergency telephone numbers).

Another point that distinguishes this paper from the publications mentioned is that none of the applications address issues of CRM, such as provision of role clarity, distribution of work load, identification of all available resources, reevaluation of the situation, and clear communication. This finding was also confirmed by Evans et al, who conducted an evaluation of the content and usability of the checklists and complained about the lack of explicit statements on responsibility [28]. The requirement analysis for the cognitive aid confirmed the initial idea that the integration of team management aspects was indeed a desirable feature. Although this requirement was already considered in the first prototype (as a placeholder), the development of a concrete concept for the implementation of CRM support in the application turned out to be a major

difficulty. This challenge became apparent in the different development stages of the prototypes. The first version included a switch for activating or deactivating CRM prompts implemented as pop-ups. A separate tab was provided for CRM in the second prototype, which could be used to read relevant notes. In the third version, CRM was no longer available as a switch or button, but it was intended to integrate the relevant information directly into the action points. As the elaboration of medical contents is a task in itself and is therefore not a part of this study, the elaboration of CRM support will not be further discussed. It should be noted, however, that a regulated distribution of tasks is essential for effective emergency management to be supported by a cognitive aid.

Discussion of Methods

The prototypical development of the digital cognitive aid for anesthesia to support intraoperative crisis management was based on the UCD process [7] described by the following main subactivities: analysis of the context of use, specification of usage requirements, development of design solutions and evaluation from the user's perspective, and extension of usage requirements. With the UCD, the end users play a key role in the development process right from the beginning, enabling optimal adaptation of the application to their specific needs.

The lack of usability can lead to rejection of the system or, even worse, to treatment errors that jeopardize patient safety [29]. The working context of the cognitive aid may further aggravate usability problems as a result of the cognitive limitations that occur during an intraoperative emergency.

The majority of diagnosis and treatment errors in anesthesia occur as a result of a mismatch between situational requirements and the actor's mental model of the current situation (ie, the so-called *human error*) [30]. Fortunately, human performance in time-critical situations can be enhanced by the use of medical applications that have been developed with user-oriented methods [31].

Finally, a user-oriented development process is a prerequisite for approval as a medical device [32]. Development according to the UCD process is therefore an established and thus suitable method for the conception of a digital cognitive aid for crisis management in anesthesia.

The usability of a software product always depends on the real operating conditions, making an analysis of the working context an indispensable prerequisite [7]. Herczeg recommends a combination of empirical and analytical methods for a comprehensive analysis [33]. In this study, the characteristics and requirements of the user groups and the user environment were analyzed by means of several different methods. The knowledge gained helped to understand anesthesiologists' working practices and needs as well as the complexity of the anesthesia workplace in routine and emergency situations.

By user tests of the different prototypes during the iterative UCD process, each development step was evaluated regarding its usability. Thus, the application was checked with each step for suitability to the intended purpose and adapted accordingly. This enables a high user acceptance of a future application.

Limitations

Throughout the entire development process, participants were only anesthesiologists from hospitals but not from outpatient departments. As a result, it is possible that the requirements for intraoperative emergencies in the outpatient and inpatient settings only partially overlap.

In addition, only physicians as the primary user group participated in evaluations. The requirements of nurses may therefore differ from those identified.

Furthermore, all evaluations of the prototypes took place in theory, that is, without dealing with a real emergency situation. However, these disadvantages are counterbalanced by future plans to evaluate the cognitive aid during performance of an anesthesia team in a simulation study.

Conclusions

Anesthesiology—as an acute medical field—is particularly characterized by its high demands on decision making and action in uncertain, dynamic, or time-critical situations. The fact that emergency checklists for anesthesia have so far only been available as paper versions has prompted the BDA to develop a national digital cognitive aid for crisis management in anesthesia.

Applications intended for use in stressful and time-critical situations pose special demands on their development. The integration of usability aspects is essential for everyday and emergency suitability. For this reason, a user-centered development was chosen according to the UCD process model as defined in ISO 9241-210 [7].

The UCD process allowed us to develop a prototypical digital cognitive aid, exhibiting high usability and user satisfaction in the demanding environment of anesthesiological emergencies. Both aspects are essential to increase the acceptance of the application in later stages.

The presented study approach, combining many different methods for determining user requirements, may be useful for other implementation projects in a highly demanding environment.

Acknowledgments

This study was performed in fulfillment of the requirements for obtaining the degree *Doctor rerum biologiae humanae* from the Friedrich-Alexander Universität Erlangen-Nürnberg (SS). The authors sincerely thank all study participants for taking the time to give their valuable feedback despite their busy schedules. Furthermore, the authors thank Christian Gulden, Christina Schüttler, and Lena Griebel for providing valuable input and comments to the manuscript. This research was supported by the Funk Stiftung (grant number RM-FS3-2017-1).

Authors' Contributions

MSP initiated the overall project and designed it together with BS and MS. BS supervised the methodological approach and conducted the context interviews. AKS developed the first prototype and collected and analyzed the data. BS, SS, and MS developed the second prototype, and BS and SS collected and analyzed the data. SS developed the third prototype and collected and analyzed the data. BS and MSP wrote the first draft of the requirements analysis. SS wrote the first draft of the manuscript. BS and MSP revised the first draft. BS and HUP provided valuable input and comments to the final version. SS wrote the final version. German Cognitive Aid Working Group: Jan Baus, Hendrik Eismann, Oliver Happel, Axel Heller, Christopher Neuhaus, Torsten Richter, and Mark Weinert. MSP is group leader. The group contributed as mentioned in the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaires first and second prototype.

[PDF File (Adobe PDF File), 622KB - mhealth_v7i4e13226_app1.pdf]

Multimedia Appendix 2

Think-aloud protocol and questionnaire third prototype.

[PDF File (Adobe PDF File), 428KB - mhealth_v7i4e13226_app2.pdf]

Multimedia Appendix 3

Screenshots third prototype.

[PDF File (Adobe PDF File), 1MB - mhealth_v7i4e13226_app3.pdf]

References

1. St.Pierre M. Stress. In: St.Pierre M, Hofinger G, Simon R, editors. Crisis Management in Acute Care Settings: Human Factors and Team Psychology in a High-Stakes Environment. Cham: Springer International Publishing; 2016:183-207.
2. Wen LY, Howard SK. Value of expert systems, quick reference guides and other cognitive aids. Curr Opin Anaesthesiol 2014 Dec;27(6):643-648. [doi: 10.1097/ACO.0000000000000130] [Medline: 25254573]
3. Goldhaber-Fiebert SN, Howard SK. Implementing emergency manuals: can cognitive aids help translate best practices for patient care during acute events? Anesth Analg 2013 Nov;117(5):1149-1161. [doi: 10.1213/ANE.0b013e318298867a] [Medline: 24108251]
4. Merry AF, Mitchell SJ. Advancing patient safety through the use of cognitive aids. BMJ Qual Saf 2016 Dec;25(10):733-735. [doi: 10.1136/bmjqs-2015-004984] [Medline: 26729917]
5. Marshall S. The use of cognitive aids during emergencies in anesthesia: a review of the literature. Anesth Analg 2013 Nov;117(5):1162-1171. [doi: 10.1213/ANE.0b013e31829c397b] [Medline: 24029855]
6. Gaba DM. Perioperative cognitive aids in anesthesia: what, who, how, and why bother? Anesth Analg 2013 Nov;117(5):1033-1036. [doi: 10.1213/ANE.0b013e3182a571e3] [Medline: 24149497]
7. International Organization for Standardization. Geneva; 2010. ISO 9241-210:2010-03, Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems URL: https://infostore.saiglobal.com/preview/is/en/2010/i.s.eniso9241-210-2010.pdf?sku=1441363 [accessed 2019-03-26] [WebCite Cache ID 779vwERqs]
8. Giunti G, Mylonopoulou V, Rivera Romero O. More Stamina, a gamified mHealth solution for persons with multiple sclerosis: research through design. JMIR Mhealth Uhealth 2018 Mar 02;6(3):e51 [FREE Full text] [doi: 10.2196/mhealth.9437] [Medline: 29500159]
9. Gabrielli S, Dianti M, Maimone R, Betta M, Filippi L, Ghezzi M, et al. Design of a mobile app for nutrition education (TreC-LifeStyle) and formative evaluation with families of overweight children. JMIR Mhealth Uhealth 2017 Apr 13;5(4):e48 [FREE Full text] [doi: 10.2196/mhealth.7080] [Medline: 28408361]
10. Curtis KE, Lahiri S, Brown KE. Targeting parents for childhood weight management: development of a theory-driven and user-centered healthy eating app. JMIR Mhealth Uhealth 2015;3(2):e69 [FREE Full text] [doi: 10.2196/mhealth.3857] [Medline: 26088692]
11. DIN EN ISO 9241-12:1998, Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs)-Part 12: Presentation of Information (ISO 9241-12:1998). Geneva: International Organization for Standardization; 1998.
12. DIN EN ISO 9241-110:2008-09, Ergonomics of Human-System Interaction-Part 110: Dialogue Principles (ISO 9241-110:2006). Geneva: International Organization for Standardization; 2006.

13. International Organization for Standardization. Geneva; 2008. DIN EN ISO 9241-151:2008-09, Ergonomics of human-system interaction - Part 151: Guidance on World Wide Web user interfaces (ISO 9241-151:2008) URL: https://www.sis.se/api/document/preview/909821/ [accessed 2019-03-26]
14. Balsamiq Sudios. 2017. URL: https://balsamiq.com/products/ [accessed 2018-12-01] [WebCite Cache ID 774FpwSNN]
15. Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry. London: Taylor & Francis Ltd; 1996.
16. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
17. Leiner DJ. SoSci Survey. 2017. URL: https://www.soscisurvey.de/ [accessed 2018-12-01] [WebCite Cache ID 774GhkQb9]
18. Nitsch M, Dimopoulos CN, Flaschberger E, Saffran K, Kruger JF, Garlock L, et al. A guided online and mobile self-help program for individuals with eating disorders: an iterative engagement and usability study. J Med Internet Res 2016 Jan 11;18(1):e7 [FREE Full text] [doi: 10.2196/jmir.4972] [Medline: 26753539]
19. Boudreaux ED, Brown GK, Stanley B, Sadasivam RS, Camargo CA, Miller IW. Computer administered safety planning for individuals at risk for suicide: development and usability testing. J Med Internet Res 2017 May 15;19(5):e149 [FREE Full text] [doi: 10.2196/jmir.6816] [Medline: 28506957]
20. Srikesavan C, Williamson E, Cranston T, Hunter J, Adams J, Lamb SE. An online hand exercise intervention for adults with rheumatoid arthritis (mySARAH): design, development, and usability testing. J Med Internet Res 2018 Jun 27;20(6):e10457 [FREE Full text] [doi: 10.2196/10457] [Medline: 29950288]
21. Burden AR, Carr ZJ, Staman GW, Littman JJ, Torjman MC. Does every code need a "reader?" improvement of rare event management with a cognitive aid "reader" during a simulated emergency: a pilot study. Simul Healthc 2012 Feb;7(1):1-9. [doi: 10.1097/SIH.0b013e31822c0f20] [Medline: 22113440]
22. German Society for Anesthesiology and Intensive Care Medicine. 2015. [Joint Federal Committee (GBA) makes surgical checklists mandatory] URL: https://www.dgai.de/aktuelles/309-gemeinsamer-bundesausschuss-gba-schreibt-op-checklisten-verbindlich-vor.html [accessed 2018-11-16] [WebCite Cache ID 73ym5h1ed]
23. Crosland M, Wang W, Ray J, Michelson S, Hutto CJ. Cognitive ergonomics applied to the eQRH: developing an electronic quick reaction handbook for use during aviation emergencies. In: Hale KS, Stanney KM, editors. Advances in Neuroergonomics and Cognitive Engineering. Cham: Springer International Publishing; 2017:27-38.
24. Marshall SD. Helping experts and expert teams perform under duress: an agenda for cognitive aid research. Anaesthesia 2017 Dec;72(3):289-295 [FREE Full text] [doi: 10.1111/anae.13707] [Medline: 27804114]
25. Watkins SC, Anders S, Clebone A, Hughes E, Zeigler L, Patel V, et al. Paper or plastic? Simulation based evaluation of two versions of a cognitive aid for managing pediatric peri-operative critical events by anesthesia trainees: Evaluation of the society for pediatric anesthesia emergency checklist. J Clin Monit Comput 2016 Jun;30(3):275-283. [doi: 10.1007/s10877-015-9714-7] [Medline: 26067401]
26. Clebone A, Burian BK, Watkins SC, Gálvez JA, Lockman JL, Heitmiller ES, Members of the Society for Pediatric Anesthesia QualitySafety Committee. The development and implementation of cognitive aids for critical events in pediatric anesthesia: the society for pediatric anesthesia critical events checklists. Anesth Analg 2017 Dec;124(3):900-907. [doi: 10.1213/ANE.0000000000001746] [Medline: 28079584]
27. McEvoy MD, Hand WR, Stoll WD, Furse CM, Nietert PJ. Adherence to Guidelines for the Management of Local Anesthetic Systemic Toxicity Is Improved by an Electronic Decision Support Tool and Designated Reader. Reg Anesth Pain Med 2014;39(4):299-305 [FREE Full text] [doi: 10.1097/AAP.0000000000000097] [Medline: 24956454]
28. Evans D, McCahon R, Barley M, Norris A, Khajuria A, Moppett I. Cognitive Aids in Medicine Assessment Tool (CMAT): preliminary validation of a novel tool for the assessment of emergency cognitive aids. Anaesthesia 2015 Aug;70(8):922-932 [FREE Full text] [doi: 10.1111/anae.13015] [Medline: 25758401]
29. Hallbeck MS, Koneczny S, Büchel D, Matern U. Ergonomic usability testing of operating room devices. Stud Health Technol Inform 2008;132:147-152. [Medline: 18391276]
30. Reason J. Safety in the operating theatre-Part 2: human error and organisational failure. Qual Saf Health Care 2005 Feb;14(1):56-60 [FREE Full text] [Medline: 15692005]
31. Bartelmes M, Neumann U, Lühmann D, Schönermark MP, Hagen A. Methods for assessment of innovative medical technologies during early stages of development. GMS Health Technol Assess 2009;5:5. [doi: 10.3205/hta000077] [Medline: 21289902]
32. Federal Ministry of Health. 2011. The Act on Medical Devices URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/M/MPG_englisch.pdf [accessed 2018-11-16] [WebCite Cache ID 73ylN6o4c]
33. Herczeg M. [Software Ergonomics (de Gruyter Studium)]. Berlin, Boston: De Gruyter; 2018.

Abbreviations

ABCDE: airway, breathing, circulation, disability, exposure

BDA: Professional Association of German Anesthesiologists
CRM: crew resource management
ISO: International Organization for Standardization
SUS: System Usability Scale
UCD: user-centered design

Edited by G Eysenbach; submitted 22.12.18; peer-reviewed by B Metelmann, B Smaradottir; comments to author 23.01.19; revised version received 05.02.19; accepted 09.02.19; published 29.04.19.

Please cite as:
Schild S, Sedlmayr B, Schumacher AK, Sedlmayr M, Prokosch HU, St.Pierre M, German Cognitive Aid Working Group
A Digital Cognitive Aid for Anesthesia to Support Intraoperative Crisis Management: Results of the User-Centered Design Process
JMIR Mhealth Uhealth 2019;7(4):e13226
URL: http://mhealth.jmir.org/2019/4/e13226/
doi:10.2196/13226
PMID:31033445

Original Paper

The Role of Data Type and Recipient in Individuals' Perspectives on Sharing Passively Collected Smartphone Data for Mental Health: Cross-Sectional Questionnaire Study

Jennifer Nicholas[1], PhD; Katie Shilton[2], PhD; Stephen M Schueller[1,3], PhD; Elizabeth L Gray[4], MS; Mary J Kwasny[4], PhD; David C Mohr[1], PhD

[1]Center for Behavioral Intervention Technologies, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[2]College of Information Studies, University of Maryland, College Park, MD, United States

[3]Department of Psychological Science, University of California - Irvine, Irvine, CA, United States

[4]Biostatistics Collaboration Center, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Corresponding Author:
Jennifer Nicholas, PhD
Center for Behavioral Intervention Technologies
Department of Preventive Medicine, Feinberg School of Medicine
Northwestern University
750 N Lake Shore Drive
10th Floor, CBITs
Chicago, IL, 60611
United States
Phone: 1 3125034156
Email: j.nicholas@northwestern.edu

Abstract

Background: The growing field of personal sensing harnesses sensor data collected from individuals' smartphones to understand their behaviors and experiences. Such data could be a powerful tool within mental health care. However, it is important to note that the nature of these data differs from the information usually available to, or discussed with, health care professionals. To design digital mental health tools that are acceptable to users, understanding how personal sensing data can be used and shared is critical.

Objective: This study aimed to investigate individuals' perspectives about sharing different types of sensor data beyond the research context, specifically with doctors, electronic health record (EHR) systems, and family members.

Methods: A questionnaire assessed participants' comfort with sharing six types of sensed data: physical activity, mood, sleep, communication logs, location, and social activity. Participants were asked about their comfort with sharing these data with three different recipients: doctors, EHR systems, and family members. A series of principal component analyses (one for each data recipient) was performed to identify clusters of sensor data types according to participants' comfort with sharing them. Relationships between recipients and sensor clusters were then explored using generalized estimating equation logistic regression models.

Results: A total of 211 participants completed the questionnaire. The majority were female (171/211, 81.0%), and the mean age was 38 years (SD 10.32). Principal component analyses consistently identified two clusters of sensed data across the three data recipients: "health information," including sleep, mood, and physical activity, and "personal data," including communication logs, location, and social activity. Overall, participants were significantly more comfortable sharing any type of sensed data with their doctor than with the EHR system or family members (P<.001) and more comfortable sharing "health information" than "personal data" (P<.001). Participant characteristics such as age or presence of depression or anxiety did not influence participants' comfort with sharing sensed data.

Conclusions: The comfort level in sharing sensed data was dependent on both data type and recipient, but not individual characteristics. Given the identified differences in comfort with sensed data sharing, contextual factors of data type and recipient appear to be critically important as we design systems that harness sensor data for mental health treatment and support.

(JMIR Mhealth Uhealth 2019;7(4):e12578) doi:10.2196/12578

KEYWORDS

mHealth; privacy; personal sensing; digital mental health; depression; anxiety; mobile phone

Introduction

Personal sensing, also referred to as context sensing and digital phenotyping [1], is the acquisition and use of data from networked sensors (as in a smartphone) for the detection of behaviors, psychological states, and environmental conditions [2]. Personal sensing shows great promise within mental health research [3]. Sensed data have already been used in a number of mental health conditions including schizophrenia [4], bipolar disorder [5], social anxiety [6], and depression [7]. For example, in schizophrenia, changes in mobility and social behavior, measured using global positioning system (GPS) and communication log data, were found to proceed clinical relapse [4]. Mobility and location data, measured using information extracted from GPS data [8,9] or the number of cell tower connections [5], have also been reported to identify and predict episodes of depression in bipolar disorder and are associated with severity of unipolar depression symptoms [9-11]. Sleep duration can also be successfully inferred using sensed data acquired from smartphones [12-14] and is related to depression severity [15]. Finally, data on subjectively reported mood, collected via ecological momentary assessments, that often accompany sensed data have demonstrated validity, correlating highly with clinician-assessed mood scales [5]. However, as demonstrations of the potential of sensed data to support mental health care and behavior change increase, questions arise regarding the acceptability of collecting different types of sensed data and the people who have access to that information.

Attitudes about privacy related to digitally collected data are theorized to rely on two major variables: contextual factors and individual characteristics. Research indicates that contextual factors may be the primary influence on people's reasoning about privacy [16,17]. A robust framework—privacy as contextual integrity—defines the contextual factors that influence people's privacy judgements and willingness to share data, such as data type and sensitivity, data use, transmission principles and constraints (eg, confidentiality or anonymity), and data recipient [18,19]. In mental health research, the *sensitivity* of data is high [20]. The *types* of data collected for mental health are broad, ranging from mood, communication logs, and social activity to GPS data. Finally, within the mental health field, potential data *recipients* beyond researchers include doctors, electronic health record (EHR) systems, and family members. Importantly, the digital data privacy literature emphasizes the need for individuals to understand (and preferably control) who has access to their personal data [16,21,22].

Despite the importance of privacy considerations in personal sensing, few studies have explored how contextual factors such as data type or recipient influence the acceptability and appropriate use of smartphone sensor data in the treatment and management of mental health. The closest example is Klasnja and colleagues' [23] study of sensed data related to physical activity. Participants' concerns about personal sensing varied depending on the type of sensor data collected. Accelerometer data were not considered sensitive, so their daily recording and storage were not of concern. However, perspectives on the collection and storage of GPS data were mixed, and raw audio data were considered very sensitive, with most participants indicating that they would not allow continuous recording. Although these results lend support to the privacy as contextual integrity framework and provide insight into privacy perspectives in personal sensing, many contextual factors differ between sensed data collected for physical activity and mental health purposes, which potentially impact the extensibility of the findings to the mental health context.

In addition to contextual factors, individual characteristics of users may influence attitudes towards privacy. Although, to our knowledge, no research has assessed differences in perspectives of sensed data privacy between people living with and those not living with a mental health condition, research indicates that data privacy and confidentiality are among the primary concerns of individuals with a mental health condition when considering the use of apps to support their mental health [24,25]. Research also suggests that older individuals may have more concerns about the collection and sharing of sensed data than younger people [26]. In a study of app privacy permissions, including access to sensors, participants characterized as "unconcerned" by permissions (those who had a high comfort level with sharing sensitive information across numerous settings) were more likely to be younger. Characteristics of the individual should therefore be considered in conjunction with context when exploring perspectives regarding sensed data sharing in mental health.

As advances in personal sensing aim to integrate the passive identification of behavioral indicators of common mental health disorders such as depression and anxiety with existing mental health services, it is critical to understand how context and individual characteristics influence individuals' perspectives regarding the use and sharing of sensed data. Understanding such perspectives is vital to guiding the design and successful implementation of digital mental health systems. The aims of this study were (1) to investigate the acceptability of sharing sensed mental health data beyond researchers, specifically with doctors, EHR systems, and family members; (2) to determine the acceptability of use of different types of sensed data beyond the research context by doctors, EHR systems, and family members; and (3) to explore the impact of age and presence of anxiety or depression on the acceptability of sharing data.

Methods

Participants

Data were collected from a convenience sample of individuals participating in a 6-week personal sensing study that required them to download an app that collected mobile sensor data

including activity, light, GPS location, and communication logs, and to complete daily questionnaires regarding sleep and wake times [12]. In that study, participants were eligible if they were aged 18 years or above, were able to read and understand English, owned an Android smartphone, and had access to Wi-Fi connectivity for at least a 3-hour period each day. Participants were excluded if they were diagnosed with any psychotic disorder or screened positive for a substance use disorder (Alcohol Use Disorder Identification Test [27] score ≥16 or Drug Abuse Screening Test-10 [28] score ≥6), suicidal ideation (Patient Health Questionnaire - 9 item [PHQ-9] [29], item 9 rating ≥1 or Beck Depression Inventory-II [30], item 9 rating ≥2), or bipolar disorder (mood disorder questionnaire [31] question 1 rating ≥7, endorsed question 2, and responded 2 or 3 for question 3). Individuals who shared their smartphone with others were also excluded.

Based on the results of the PHQ-9 and Generalized Anxiety Disorder - 7 item (GAD-7) [32] screening questionnaires, participants were selected to create roughly four equal groups across depression and anxiety symptoms: nondepressed or anxious (PHQ-9 score<10; GAD-7 score<10), depressed (PHQ-9 score ≥ 10; GAD-7 score <10), anxious (PHQ-9 score <10; GAD-7 score ≥10), and depressed and anxious (PHQ-9 score ≥ 10; GAD-7 score ≥10).

The study was approved by the Northwestern University Institutional Review Board. Participant responses from only the screening and baseline questionnaire were considered in this study. Within the larger study, participants were compensated for their participation. Compensation depended on both the length of their participation in the study and the number of daily questionnaires answered, and ranged between US $25 and US $270.40.

Measures

Demographic information (for example, age and sex) and data on the presence of depression or anxiety, determined using the PHQ-9 and GAD-7, were collected by self-report.

The acceptability of sharing sensed data was measured through a series of questions with regard to three potential recipients: the participant's doctor, representing a known individual in the health care system; the participant's EHR system, a generic destination in the health care system that would broaden access to potentially unknown people; and the participant's family members. Response options were recorded on an ordinal scale:

0 - I would not use any app that gave these data to my doctor/electronic health record/family member;

1 - I'd be uncomfortable but would consider using an app that did this;

2 - It wouldn't matter to me;

3 - I'd like an app that gave these data to my doctor/electronic health record/family members.

Participants rated their comfort with sharing five classes of sensed activities or states (physical activity, sleep, mood, social activity, and location [places visited and patterns of movement]) and one raw sensed data type (communication logs [number of calls made or texts sent]), with each of the three recipients.

Data Analysis

Given the ordinal (0, 1, 2, and 3) nature of the responses on the acceptability scale and the relatively limited sample size, it was not appropriate to treat these data as continuous. Although these outcomes are appropriate for ordinal logistic regression, we did not assume proportional odds for our models, and hence, survey responses were dichotomized to indicate whether the participant was comfortable sharing their data. Responses 0 and 1 (*I would not use any app that gave these data to my doctor/electronic health record/ family member* and *I'd be uncomfortable but would consider using an app that did this*, respectively) were coded as not comfortable, and responses 2 and 3 (*It wouldn't matter to me* and *I'd like an app that gave these data to my doctor/electronic health record/ family members*, respectively) were coded as comfortable.

First, we performed a series of three principal component analyses, one for each data recipient, to identify clusters of sensor data types according to participants' comfort with sharing them. Second, we determined relationships between recipients and the types of sensor data using generalized estimating equation logistic regression models of each participant's 18 dichotomized responses (3 types of recipients × 6 types of sensor data, which were clustered into two groups based on the principal component analysis results). Third, we explored the influence of age and mental health on participants' comfort with sharing sensed data by adding age, depression (defined as PHQ-9 score ≥ 10 and GAD-7 score <10), and anxiety (defined as PHQ-9 score <10 and GAD-7 score ≥10) as covariates in the model. Models were assessed using analysis of variance and Wald tests. All analyses were performed using R (v3.4.3; R Foundation for Statistical Computing, Vienna, Austria) with a type I error rate of 0.05.

Results

Participants

A total of 211 eligible participants were enrolled and completed the privacy survey. The majority were female (171/211; 81.0%), and the mean (SD) age was 38 (10.32) years, ranging from 18 to 66 years. A total of 83% (176/211) of participants identified as Caucasian; 13.3% (28/211), as African American; and 9.0% (19/211), as Hispanic or Latino. Further details of the sample are shown in Table 1.

Table 1. Participant characteristics.

Characteristic	Statistics
Gender, n (%)	
Male	36 (17.1)
Female	171 (81.0)
Another	3 (1.4)
Age (years), mean (SD)	38.09 (10.32)
Race and ethnicity[a], n (%)	
Black or African American	28 (13.3)
American Indian or Alaska native	6 (2.8)
Asian	10 (4.7)
White	176 (83.4)
Hispanic or Latino	19 (9.0)
Highest level of education, n (%)	
Some high school	4 (1.9)
Completed high school	25 (11.8)
Some college	77 (36.5)
Completed associate's or bachelor's degree	77 (36.5)
Master's degree	23 (10.9)
Doctoral degree or professional doctorate	5 (2.4)
Employment status, n (%)	
Employed	130 (61.6)
Unemployed	44 (20.9)
Disability	17 (8.1)
Retired	4 (1.9)
Other	16 (7.6)
Mental health status, n (%)	
Healthy (PHQ-9[b] score<10 and GAD-7[c] score<10)	59 (28.0)
Depressed (PHQ-9 score≥10 and GAD-7 score<10)	55 (26.1)
Anxious (PHQ-9 score<10 and GAD-7 score≥10)	44 (20.9)
Depressed and anxious (PHQ-9 score≥10 and GAD-7 score≥10)	53 (25.1)

[a]Race and ethnicity categories are not mutually exclusive.

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

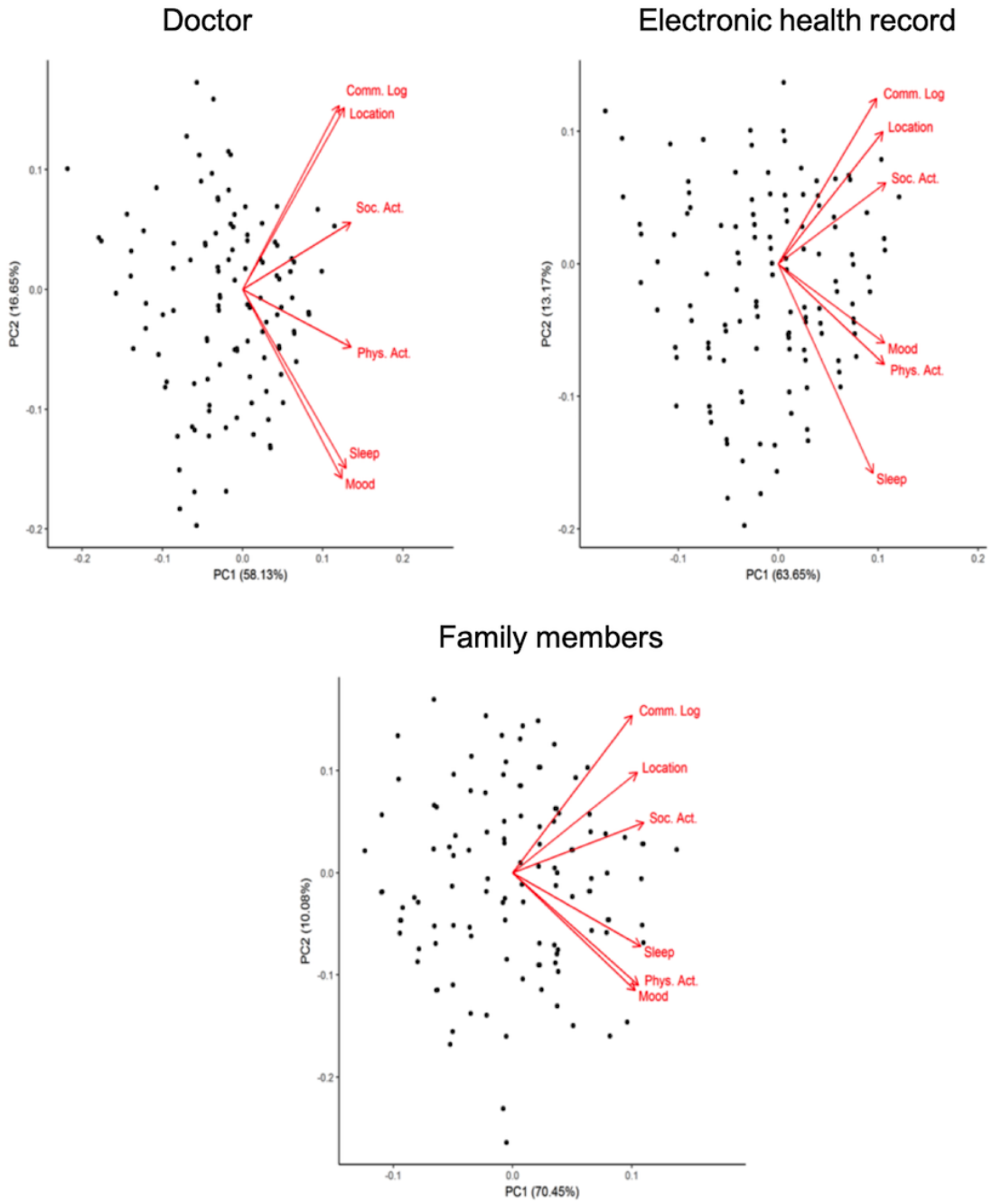
[c]GAD-7: Generalized Anxiety Disorder - 7 item.

Sensor Data Type

Principal component analyses were performed on participant responses to their comfort level with sharing the six types of sensor data with each of the three data recipients. Although sensor clusters were largely consistent across data recipients, the relative strengths of each variable across recipients differed. The two clusters extracted across each of the recipients could best be described as "health information," including sleep, mood, and physical activity, and "personal data," including communication logs, location, and social activity.

For doctor recipients, the two principle components explained 74.8% of the variability in the responses (Figure 1). Participants were most comfortable sharing health information with doctors and least comfortable sharing personal data, particularly communication logs and location. The same two components for the EHR system explained 76.8% of the variability in the responses (Figure 1). Participants were least comfortable sharing personal data with their EHR system. For family members, the two components explained 80.5% of the responses (Figure 1); however, the two groups were less distinct in terms of comfort with sharing personal data with family members than with doctors or EHR systems.

Figure 1. Principal component analysis of participant comfort with sharing data from six different sensors with their doctor, electronic health record, and family members. Comm Log; communications log; PC: principal component; Phys Act: physical activity; Soc Act: social activity.



Sensor Data Sharing

Comfort level in sharing the different types of sensor data with each of the three recipient groups is displayed in Table 2. On the whole, participants indicated they were comfortable sharing their sensor data apart from communication logs and location data, and the comfort levels were higher when such data were shared with their doctor than with their family members.

Table 2. Participants' comfort level sharing various sensor data with different recipients. The values indicate participants who were comfortable with sharing data.

Sensor data type	Doctor, n (%)	Electronic health record, n (%)	Family members, n (%)
Sleep	188 (89.1)	183 (86.7)	135 (64.0)
Mood	173 (82.0)	151 (71.6)	108 (51.2)
Physical activity	162 (76.8)	152 (72.0)	109 (51.7)
Communication logs	135 (64.0)	110 (52.1)	115 (54.5)
Location	122 (57.8)	105 (49.8)	97 (46.0)
Social activity	154 (73.0)	129 (61.1)	116 (55.0)

Generalized estimating equation logistic regression models were fit on the binary outcome of participants' comfort with sharing sensor data, with the type of sensor data, the recipient of sensor data, presence of depression, presence of anxiety, and age (modeled continuously) used as covariates. The interaction of depression and anxiety was considered, but was not significant (P=.25). As responses on the sensor data type were highly correlated, the type was grouped according to the results of the principal component analyses mentioned above, with sleep, mood, and physical activity combined as "health information" and communication logs, location, and social activity grouped as "personal data." Data recipient was a significant predictor of a participant's comfort with sharing data (P<.001; Table 3). Individuals were significantly more comfortable sharing their sensed data with their doctor than with family members (P<.001) or their EHR system (P<.001), regardless of the data type. As suggested by the principal component analyses, overall, participants were significantly more likely to be comfortable sharing "health information" than "personal data" (P<.001). The model results also indicated a significant interaction between data type and recipient (P<.001). Comfort with sharing data was more strongly associated with data recipient for health information than for personal data, such that the difference in comfort with sharing data between doctors and family members was greater for health information than for personal data. We failed to detect any difference in comfort with sharing data with regard to depression (P=.12), anxiety (P=.14), or age (P=.67) in our study.

Table 3. Model summary of participants' comfort with sharing "health information" (sleep, mood, and physical activity) and "personal data" (communication logs, location data, and social activity) with their doctor, the electronic health record system, or family members.

Covariate	Estimate	SE	Wald statistic	Pr(>|W|)
Intercept	0.745	0.468	2.529	.112
Age	–0.002	0.011	0.041	.840
Anxiety	0.309	0.211	2.158	.142
Depression	–0.333	0.214	2.416	.120
Health information[a]	0.953	0.126	57.621	<.001[b]
Personal data: Recipient - EHR[c,d]	–0.451	0.074	36.722	<.001[b]
Personal data: Recipient - family[e]	–0.554	0.108	26.460	<.001[b]
Health information: Recipient - EHR[c]	0.086	0.103	0.690	.406
Health information: Recipient - family[e]	–0.801	0.141	32.453	<.001[b]

[a]Health information versus personal data.

[b]These values are significant.

[c]EHR versus doctor as recipient.

[d]EHR: electronic health record.

[e]Family versus doctor as recipient.

Discussion

Principal Findings

This study explored the attitudes of participants toward sharing of personal data gathered from smartphone sensors with three potential data recipients in the context of mental health care involving the use of digital interventions. We found that the level of comfort with sharing sensed data was dependent on both data *type* and *recipient*, but not individual characteristics. This result is in accordance with the contextual integrity framework [18,33], which states that privacy expectations are influenced by contextual factors such as data type and sensitivity, the relationship between the individual and the data recipient, and context-specific information norms [33]. We found that participants had similar levels of comfort with sharing sleep, mood, and physical activity data ("health information") and were more comfortable sharing them than communication

logs, location, and social activity data ("personal data"). Moreover, participants were significantly more comfortable sharing sensed data with their doctor than with the EHR system or family members.

This difference between comfort in sharing data with doctors and that in sharing data with the EHR system particularly emphasizes the nuanced role that the data recipient plays in privacy concerns. The difference is possibly dependent on the existing relationship between the individual and the data recipient, and the important role of trust in privacy [33]. Participants may attribute more trust to a specific person with whom a relationship discussing health has been established (ie, their doctor), rather than the health system more generally, represented by the EHR system. Indeed, previous research has shown that when a trusting relationship is not established, participants resist sharing sensed physical activity data with health providers [34]. In our study, the fact that participants did not always extend trust and willingness to share their data with the EHR system may indicate that they have concerns about who can access their data.

The role of the recipient in comfort with data sharing, therefore, has important implications for the use of apps that acquire and transmit sensed information related to mental health, especially given the variety of people who are often involved in mental health treatment and management. Providers as well as technological system designers must be aware that although individuals may be comfortable sharing their sensed data with their doctor, they may not be comfortable sharing it more widely, even with people who are already involved in their mental health management such as other health professionals, via the EHR system, or family members. As apps are integrated into clinical care, upfront and ongoing conversations regarding the distribution of sensed data will become increasingly critical, as will provider education about sensed data and the ability for providers and individuals to manage sharing options.

In line with contextual integrity, participants' comfort with sharing sensed data significantly differed by data type. The way in which these different data types fit within existing information norms in the doctor-patient context could possibly explain the observed differences. Sleep, mood, and physical activity data types may closely align with doctor-patient information norms, as they are often discussed with providers. On the other hand, the data types that participants were least comfortable in sharing—communication logs, location, and social activity—are not commonly discussed with doctors. Thus, discussing these data types may violate existing norms and therefore be less willingly shared with providers.

Another explanation for the observed difference in comfort with sharing across data types is that communication logs, location, and social activity data may carry additional sensitivities. For example, just four points of GPS data can reveal the identity of up to 95% of individuals [35]; personal communications document a wider variety of behaviors, habits, and beliefs than data like sleep patterns, and sharing social activity data might put the privacy of other individuals at risk. Such differences related to data type, information norms, and the unique sensitivities associated with certain data are critical to understand and heed as we begin to create systems that harness sensor data for mental health treatment and support.

Comfort with data sharing could not be predicted based on the individual characteristics of age and presence of depression or anxiety, supporting an important tenet of contextual integrity: Contextual factors more strongly influence individuals' privacy preferences than individual characteristics [36]. The lack of differences in comfort with sharing sensed data between individuals with depression or anxiety and those without a mental health condition found in our study is interesting and important, given the stigmatized nature of these common mental health conditions [37] and the risks associated with sharing. However, further research should aim to elucidate whether any differences in the nature of concerns regarding data privacy exist between people with common mental health disorders and those without such disorders, in order to mitigate any specific concerns of this population.

Irrespective of people's mental health status, ethicists have raised concerns that personal sensing projects challenge traditional research ethics' tenets of informed consent and risk mitigation, because data collection is both unobtrusive (easily forgotten and not easily avoided) and pervasive (recording many aspects of a participants' daily habits for long periods) [19,38], and the inferential harms of passively collected data are often poorly understood [39]. Our findings suggest that personal sensing projects should use contextual factors to guide research design and should revise participant consent processes to address these ethical concerns. For example, individuals should be afforded the opportunity to select specific allowances for data sharing based on factors such as type, purpose, recipient, and sensitivity, rather than providing a blanket consent. Further, researchers should not assume that the acceptability of using sensed data is easily generalized between different research contexts or types of data collected without first considering the comparability of the contextual norms (roles, data types, transmission principles, and uses). Beyond these considerations, data collectors must also ensure protections, such as data deletion, deidentification, and restrictions on sharing.

Limitations

Although this study reveals important differences in the comfort level of individuals sharing sensed data based on the recipient and data type, a number of limitations should be considered when interpreting the results. First, survey participants consented to participate in a wider study that collected data from a number of smartphone sensors over a 6-week period. Therefore, the views of individuals who were deeply uncomfortable with sharing sensed data were likely not represented within the sample. This sampling bias reduces the generalizability of these findings to the general population. However, these findings do represent sensitivities of people who are open to using personal sensing apps. As a further caution to generalizability, we note that the majority of the sample comprised white, employed, and well educated people. Given that privacy is experienced by different populations in distinct ways [40], further work should examine privacy associated with mental health across a broader section of the community.

This study asked general questions about individuals' willingness to share sensed data and did not explore richer contextual factors such as perceived benefits or risks that are often considered by individuals when making decisions about privacy [39]. Therefore, the behavior and decisions made by individuals when deciding whether to share sensed data may vary from what is outlined here. The importance of richly contextualizing information about the collection of sensed data for mental health was highlighted by a recent study [34] on veterans with posttraumatic stress disorder. Researchers found that the lack of clarity of purpose was a primary reason given by participants for not using a wearable fitness tracker to support treatment. Lack of purpose also contributed to uncertainty and increased discomfort about the collection of sensed GPS and Bluetooth data in a feasibility study of a passive data–collection app [41]. Individuals' existing beliefs about the importance of privacy and data control have also been shown to interact with contextual factors when people make privacy decisions [36]. These beliefs were not explored in the current study. Future research that more richly contextualizes the collection and sharing of sensed data and explores existing beliefs using vignettes, semistructured interviews, or other contextualized methods would provide deeper insight into why the reported differences in comfort with sharing sensed data exist.

Finally, considering the importance of context in our study, it is worth noting that the data in this study were collected before a number of highly publicized data privacy scandals took place, most notably, that of Cambridge Analytica, in which data were misused, and that of Strava, where released data had unforeseen consequences for disclosure. These events brought to light the importance of contextual factors such as unforeseen harms and data re-identification, purpose, and recipient. Individuals' attitudes toward passive data collection may have since changed, again highlighting the need for further contextualized research regarding the privacy of sensed data for mental health.

Conclusions

In line with the contextual integrity framework, participants' comfort with sharing sensed data was dependent on the type of data collected and the intended recipient of those data. Given these differences, research and treatment protocols and systems designed to use sensed data must consider differences in individuals' comfort depending on contextual factors. These differences represent important considerations, as systems are developed to integrate sensed data into health systems and use these data to encourage behavior change and mental health management. The reported insights will help establish data sharing norms for personal sensing and manage or mitigate privacy concerns as we develop systems to collect, share, and use sensed data to support mental health treatment.

Acknowledgments

This research was supported by the National Institute of Mental Health (R01 MH111610 and P20 MH090318 to DM).

Conflicts of Interest

None declared.

References

1. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
2. Schatz BR. National Surveys of Population Health: Big Data Analytics for Mobile Health Monitors. Big Data 2015;3(4):219-229 [FREE Full text] [doi: 10.1089/big.2015.0021] [Medline: 26858915]
3. Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. JAMA 2017;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
4. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J. Relapse Prediction in Schizophrenia Through Digital Phenotyping: A Pilot Study. Neuropsychopharmacology 2018;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]
5. Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE, Kessing LV. Smartphone Data as Objective Measures of Bipolar Disorder Symptoms. Psychiatry Res 2014;217(1-2):124-127. [doi: 10.1016/j.psychres.2014.03.009] [Medline: 24679993]
6. Chow PI, Fua K, Huang Y, Bonelli W, Xiong H, Barnes LE, et al. Using Mobile Sensing to Test Clinical Models of Depression, Social Anxiety, State Affect, and Social Isolation Among College Students. J Med Internet Res 2017;19(3):e62 [FREE Full text] [doi: 10.2196/jmir.6820] [Medline: 28258049]
7. Mohr DC, Zhang M, Schueller SM. Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning. Annu Rev Clin Psychol 2017;13:23-47. [doi: 10.1146/annurev-clinpsy-032816-044949] [Medline: 28375728]
8. Osmani V. Smartphones in Mental Health: Detecting Depressive and Manic Episodes. IEEE Pervasive Comput 2015 Jul;14(3):10-13. [doi: 10.1109/MPRV.2015.54]
9. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study. J Med Internet Res 2015;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]

10. Canzian L, Musolesi M. Trajectories of Depression: Unobtrusive Monitoring of Depressive States by Means of Smartphone Mobility Traces Analysis. : ACM Press; 2015 Presented at: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing; Sept 7-11; Osaka, Japan.
11. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The Relationship Between Mobile Phone Location Sensor Data and Depressive Symptom Severity. PeerJ 2016;4:e2537 [FREE Full text] [doi: 10.7717/peerj.2537] [Medline: 28344895]
12. Saeb S, Cybulski TR, Schueller SM, Kording KP, Mohr DC. Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles. J Med Internet Res 2017;19(4):e118 [FREE Full text] [doi: 10.2196/jmir.6821] [Medline: 28420605]
13. Doryab A, Min J, Wiese J, Zimmerman J, Hong J. Detection of Behavior Change in People with Depression. 2014 Presented at: AAAI workshop: Modern Artificial Intelligence for Health Analytics; Jul 27-28; Quebec, Canada.
14. Rabbi M, Ali S, Choudhury T, Berke E. Passive and In-situ Assessment of Mental and Physical Well-being using Mobile Sensors. Proc ACM Int Conf Ubiquitous Comput 2011;2011:385-394 [FREE Full text] [doi: 10.1145/2030112.2030164] [Medline: 25285324]
15. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. StudentLife: Assessing Mental Health, Academic Performance and Behavioral Trends of College Students using Smartphones. StudentLife; 2014 Presented at: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; Sept 13-17; Seattle, WA.
16. Adams A. Users' Perception of privacy in Multimedia Communication. 1999 Presented at: CHI'99 Extended Abstracts on Human Factors in Computing Systems; May 15-20; Pittsburgh, PA.
17. Barth A, Datta A, Mitchell J, Nissenbaum H. Privacy and Contextual Integrity: Framework and Applications. 2006 Presented at: Proceedings of the 2006 IEEE Symposium on Security and Privacy; May 21-24; Berkeley/Oakland, CA.
18. Nissenbaum H. Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford, CA: Stanford University Press; 2009.
19. Zimmer M. Addressing Conceptual Gaps in Big Data Research Ethics: An Application of Contextual Integrity. Social Media+Society 2018;4(2):1-11. [doi: 10.1177/2056305118768300;4]
20. Dehling T, Gao F, Schneider S, Sunyaev A. Exploring the Far Side of Mobile Health: Information Security and Privacy of Mobile Health Apps on iOS and Android. JMIR Mhealth Uhealth 2015;3(1):e8 [FREE Full text] [doi: 10.2196/mhealth.3672] [Medline: 25599627]
21. Pentland A, Lazer D, Brewer D, Heibeck T. Using Reality Mining to Improve Public Health and Medicine. Stud Health Technol Inform 2009;149:93-102. [Medline: 19745474]
22. Shilton K. Four Billion Little Brothers? Commun ACM 2009 Nov 01;52(11):48. [doi: 10.1145/1592761.1592778]
23. Klasnja P, Consolvo S, Choudhury T, Beckwith R, Hightower J. Exploring Privacy Concerns About Personal Sensing. 2009 Presented at: Proceedings of the 7th International Conference on Pervasive Computing; May 11-14; Nara, Japan.
24. Proudfoot J, Parker G, Hadzi Pavlovic D, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. J Med Internet Res 2010 Dec 19;12(5):e64 [FREE Full text] [doi: 10.2196/jmir.1475] [Medline: 21169174]
25. Nicholas J, Boydell K, Christensen H. Beyond Symptom Monitoring: Consumer Needs for Bipolar Disorder Self-management Using Smartphones. Eur Psychiatry 2017 Dec;44:210-216. [doi: 10.1016/j.eurpsy.2017.05.023] [Medline: 28692910]
26. Lin J, Liu B, Sadeh N, Hong J. Modeling Users Mobile App Privacy Preferences: Restoring Usability in a Sea of Permission Settings. 2014 Presented at: SOUPS'14 Proceedings of the Tenth USENIX Conference on Usable Privacy and Security; Jul 09-11; Menlo Park, CA.
27. Bohn MJ, Babor TF, Kranzler HR. The Alcohol Use Disorders Identification Test (AUDIT): Validation of a Screening Instrument for Use in Medical Settings. J Stud Alcohol 1995;56(4):423-432. [Medline: 7674678]
28. Skinner HA. The Drug Abuse Screening Test. Addict Behav 1982;7(4):363-371. [Medline: 7183189]
29. 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] [Medline: 11556941]
30. Beck A, Steer R, Brown G. Manual for the Beck Depression Inventory-II. San Antonio, USA: Psychological Corporation; 1996.
31. Hirschfeld RM, Williams JB, Spitzer RL, Calabrese JR, Flynn L, Keck PE, et al. Development and Validation of a Screening Instrument for Bipolar Spectrum Disorder: The Mood Disorder Questionnaire. Am J Psychiatry 2000;157(11):1873-1875. [Medline: 11058490]
32. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. Arch Intern Med 2006;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
33. Nissenbaum H. Privacy as contextual integrity. Wash L Rev 2004;79:101-139 [FREE Full text]
34. Ng A, Reddy M, Zalta AK, Schueller SM. Veterans' Perspectives on Fitbit Use in Treatment for Post-Traumatic Stress Disorder: An Interview Study. JMIR Ment Health 2018;5(2):e10415 [FREE Full text] [doi: 10.2196/10415] [Medline: 29907556]
35. De Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. Unique in the Crowd: The Privacy Bounds of Human Mobility. Sci Rep 2013;3:1376 [FREE Full text] [doi: 10.1038/srep01376] [Medline: 23524645]

36. Martin K, Shilton K. Why experience matters to privacy: How context-based experience moderates consumer privacy expectations for mobile applications. J Assn Inf Sci Tec 2015 May 05;67(8):1871-1882 [FREE Full text] [doi: 10.1002/asi.23500]
37. Corrigan P, Watson A. The Paradox of Self-Stigma and Mental Illness. Clinical Psychology: Science and Practice 2002;9:35-53 [FREE Full text] [doi: 10.1093/clipsy.9.1.35]
38. Vitak J, Shilton K, Ashktorab Z. Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. : ACM Press; 2016 Presented at: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; Feb 27-Mar 02; San Francisco, CA.
39. Beckwith R, Mainwaring S. Privacy: Personal information, threats, and technologies. 2005 Presented at: Proceedings International Symposium on Technology and Society, 2005 Weapons and Wires: Prevention and Safety in a Time of Fear ISTAS; Jun 8-10; Los Angeles, CA.
40. Marwick A, Boyd D. Understanding Privacy at the Margins. International Journal of Communication 2018;12:1157-1165 [FREE Full text]
41. Boonstra TW, Nicholas J, Wong QJ, Shaw F, Townsend S, Christensen H. Using Mobile Phone Sensor Technology for Mental Health Research: Integrated Analysis to Identify Hidden Challenges and Potential Solutions. J Med Internet Res 2018;20(7):e10131 [FREE Full text] [doi: 10.2196/10131] [Medline: 30061092]

Abbreviations

EHR: electronic health record
GAD-7: Generalized Anxiety Disorder - 7 item
GPS: global positioning system
PHQ-9: Patient Health Questionnaire - 9 item

Edited by C Dias; submitted 23.10.18; peer-reviewed by M Chen, T Wykes; comments to author 05.12.18; revised version received 29.01.19; accepted 05.02.19; published 05.04.19.

Please cite as:
Nicholas J, Shilton K, Schueller SM, Gray EL, Kwasny MJ, Mohr DC
The Role of Data Type and Recipient in Individuals' Perspectives on Sharing Passively Collected Smartphone Data for Mental Health: Cross-Sectional Questionnaire Study
JMIR Mhealth Uhealth 2019;7(4):e12578
URL: http://mhealth.jmir.org/2019/4/e12578/
doi:10.2196/12578
PMID:30950799

Original Paper

Barriers to and Facilitators of the Use of Mobile Health Apps From a Security Perspective: Mixed-Methods Study

Leming Zhou[1], DSc, PhD; Jie Bao[1], MSc; Valerie Watzlaf[1], PhD; Bambang Parmanto[1], PhD

Department of Health Information Management, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Leming Zhou, DSc, PhD
Department of Health Information Management
University of Pittsburgh
6021 Forbes Tower
3600 Forbes Ave at Meyran Ave
Pittsburgh, PA, 15260
United States
Phone: 1 412 383 6653
Fax: 1 412 383 6655
Email: lmzhou@gmail.com

Abstract

Background: A large number of mobile health (mHealth) apps have been created to help users to manage their health or receive health care services. Many of these mHealth apps have proven to be helpful for maintaining or improving their users' health. However, many people still choose not to use mHealth apps or only use them for a short period. One of the reasons behind this lack of use is the concern for their health information security and privacy.

Objective: The goal of this study was to determine the relationship between users' characteristics and their security and privacy concerns and to identify desired security features in mHealth apps, which could reduce these concerns.

Methods: A questionnaire was designed and validated by the research team. This questionnaire was then used to determine mobile app users' security and privacy concerns regarding personal health data in mHealth apps as well as the security features most users' desire. A semistructured interview was used to identify barriers to and facilitators of adopting mHealth apps.

Results: In total, 117 randomly selected study participants from a large pool took part in this study and provided responses to the validated questionnaire and the semistructured interview questions. The results indicate that most study participants did have concerns about their privacy when using mHealth apps. They also expressed their preferences regarding several security features in mHealth apps, such as regular password updates, remote wipe, user consent, and access control. An association between their demographic characteristics and their concerns and preferences in security and privacy was identified; however, in most cases, the differences among the different demographic groups were not statistically significant, except for a few very specific aspects. These study participants also indicated that the cost of apps and lack of security features in mHealth apps were barriers for adoption, whereas having free apps, strong but easy-to-use security features, and clear user protection privacy policies might encourage them to use mHealth apps in their health management.

Conclusions: This questionnaire and interview study verified the security and privacy concerns of mHealth app users, identified the desired security and privacy features, and determined specific barriers to and facilitators of users adopting mHealth apps. The results can be used to guide mHealth app developers to create apps that would be welcomed by users.

(JMIR Mhealth Uhealth 2019;7(4):e11223) doi:10.2196/11223

KEYWORDS

confidentiality; privacy; mobile apps; questionnaire

Introduction

Background

In recent years, both ownership of smartphones and the number of available mobile health (mHealth) apps have increased dramatically. According to one study performed by the Pew Research Center in 2018, 95% of Americans owned a mobile phone, and in 2018, 77% of them were smartphones [1] as opposed to 2011, when the share of smartphones among American adults was just 35%. In the same period, many mHealth apps were created and published on app stores. Specifically, by October 2017, roughly 325,000 mHealth apps were available on major app stores [2].

For patients (in a general sense, ie, people who want to maintain or improve their health), mHealth apps can be used to perform tasks such as wellness management, encouraging and monitoring behavior change, health data collection, disease management, self-diagnosis, medication reminders, and rehabilitation schedule management [3,4]. A number of research studies have been performed on mHealth apps and their results indicate that well-designed mHealth apps can empower patients, improve medication adherence, and reduce the cost of health care [5-8].

However, the adoption of mHealth apps in personal health care is still limited. The growth rate of mHealth app downloads dropped dramatically from more than 35% in 2015 to roughly 7% in 2016 [2]. Moreover, it has been shown that after smartphone users download mHealth apps, close to half of them stop using mHealth apps for various reasons such as hidden costs, high data entry burden, loss of interest, and security and privacy concerns [9].

In the context of patient health data, security and privacy are always linked since any unauthorized access to patient health data (security breach) is a violation of patient privacy. Here, security is the state of being protected against the unauthorized use of patient health information, whereas privacy is the freedom from unauthorized intrusion.

There are various types of mHealth apps; some collect health information from patients, whereas others simply provide general guidelines for maintaining a healthy lifestyle and information about certain diseases. If an mHealth app does not handle any patient health data, it typically does not trigger security and privacy concerns; therefore, the mHealth apps discussed in this study are the ones that handle patient health data.

Concern about health data security and privacy is one important reason people choose not to use mHealth apps for their own health care [3,9-11]. More specifically, users are not certain what type of data are collected and stored by mHealth apps, who can access the self-entered and sensor-collected data, and what purpose data are used for [10,12]. Security and privacy concerns about mHealth apps are greater when the apps are for issues associated with stigma, social isolation, or discrimination such as HIV/AIDS, sexual orientation, and mental disease [13-17]. All of these concerns are not surprising since millions of patients' health records have been compromised because of hacking or other incidents in recent years in the United States [18]; many mHealth apps do not have the necessary security features to protect users' health data [19-21]; at the same time, many smartphone users (including patients and health care providers) do not even use the most basic authentication features (such as a passcode) to prevent access to private data on their phones [22,23]. In addition, a recent study has indicated that in 2015 around 70% of the 600 most commonly used mHealth apps did not provide privacy policies; many current mHealth app developers do not provide privacy policies in their apps either [24,25].

One possible way to reduce users' concerns about privacy in mHealth apps is to determine the specific concerns of mHealth app users, evaluate the association between these concerns and users' characteristics (such as demographics, experience with technology, and health care needs), and identify specific mHealth app features that may enhance their trust in the apps so that they will start to use mHealth apps in their own health care and management. In other words, investigating the barriers to and facilitators of using mHealth apps may lead to finding a way to increase users' adoption of mHealth apps. Before going into the details of this study, a brief review of previous studies on this topic is provided below.

Previous Studies

A number of studies have been conducted to identify users' attitudes toward and perceptions of mHealth apps using focus groups, questionnaires, and interviews [9,10,26-28]. Below is a summary of the findings in a few of these studies.

In 2015, Krebs and Duncan distributed a cross-sectional survey throughout the United States to determine the usage of mHealth apps among mobile phone owners and the reasons behind their choice about whether or not to use mHealth apps [9]. There were 1604 respondents in the study, and more than 40% of these mobile phone users reported that they had chosen not to download mHealth apps. One of the reasons given was security and privacy concerns. The ones who had chosen to download mHealth apps, on the other hand, seemed to trust in the security of the app. Individuals more likely to use health apps tended to be younger, have higher incomes, be more educated, be Latino/Hispanic, and have a body mass index in the obese range.

Atienza et al used a mixed-methods approach (survey and focus group studies) to determine consumer attitudes toward and perceptions of mHealth privacy and security [10]. The conclusion was that user attitudes regarding mHealth privacy and security were highly contextualized. They were related to the type, place, time, purpose, and person accessing the health information. They found that people in similar demographic groups may have quite different opinions on privacy.

Peng et al conducted focus groups and individual interviews with 44 smartphone owners to determine user perceptions of mHealth apps [26]. People in all demographic groups (age, gender, and income) revealed that they did not like to share health data in the app via social networking features. They had concerns about how the information might be exploited by a third party. They were willing to share selected information with a small number of people if necessary. Besides security and privacy concerns, one other major barrier the participants

of the study mentioned was the cost of the app. Many people only used free mHealth apps.

Dennison et al conducted 4 focus group studies with students and staff at a university in the United Kingdom to assess the opportunities for and challenges to getting young adults to use smartphone apps in supporting health behavior change [27]. Study results indicated that young, healthy young adults have some interest in apps for behavior change. However, participants expressed concerns about the security of the data in the app. They were afraid that the data might get into the hands of third parties. They also felt it was intrusive when apps use a context-sensing approach to generate reminders or suggestions. They particularly did not like the app using the Global Positioning System to track their locations.

Prasad et al arranged 8 focus groups to identify privacy concerns related to mHealth apps [28]. The participants were young college students (aged between 19 and 30 years), elderly hospital outpatients (aged between 80 and 85 years), and residents of a retirement community (aged between 65 and 100 years). First, elderly participants were more comfortable sharing health information with their doctors than family members. The young participants were willing to share their medication information with their doctors but no other information such as their location and their social interactions with others. Second, some participants were afraid that some information collected by the mobile device might get compromised during transmission and storage. They wanted to have control over the disclosure of their data because they included private information. Young participants did not want the mobile device to collect their information without their consent.

One comprehensive review covered privacy in mobile technology for personal health care and discussed topics such as security regulations, technologies, threats, and possible solutions [29]. One review of privacy and security in mHealth apps briefly summarized the studies on security research in mHealth systems and provided general recommendations for creating secure mHealth apps [30]. Another study provided more detailed security recommendations for mHealth apps [31]. These recommendations are theoretical, with a major focus on the sensitivity of information itself. Although these recommendations for mHealth app developers are surely helpful in terms of making mHealth apps more secure, we also need to take end-user concerns into account and consider the usability of the mHealth app [32]. After all, no matter how secure the app is, if the end-users do not like it and do not use it, it will not contribute to the improvement of users' health and well-being.

In other words, although there are several studies that have revealed the existence of security and privacy concerns from mobile app users, significant diversity in attitudes regarding mHealth privacy/security exists for different demographic groups. Thus, one may need to specifically customize security features for different purposes and different users to address users' individual concerns regarding mHealth privacy and security in apps.

Objectives

In this study, we used a questionnaire and held interviews to assess the association between users' demographic characteristics and their security and privacy concerns, and more importantly, the specific security features they desire to have in mHealth apps and the features or language that may encourage them to use mHealth apps.

The purpose of this questionnaire and interview study is to collect data and answer the following 4 questions:

1. What are mobile app users' opinions or concerns about their personal data security and privacy?
2. What are mobile app users' opinions and concerns about their data security and privacy in mHealth apps?
3. What are the security and privacy features they desire to see in mHealth apps?
4. What are the barriers to and facilitators of the use of mHealth apps?

Methods

Questionnaire and Interview Question Development

Step 1. Literature Search and Review

We used the literature collected in our previous research studies [33-38] and performed keyword searches "(security OR privacy)" AND "(questionnaire OR survey OR interview)" for published studies in PubMed, IEEE Xplore, ACMD Digital Library, and INSPEC. We also used the same keywords to perform searches in Google. From the obtained search results, we identified a few hundred statements relevant to information security and privacy.

Step 2. Creating a Draft of the Questionnaire

Each of the research team members went through these identified statements to determine their relevance and clarity in terms of the study purpose on a scale of 1 to 4, where 1 means no relevance or clarity, whereas 4 means high relevance or clarity. If 3 or 4 team members rated the relevance of a statement 1 or 2, it was removed from the questionnaire. If one of the team members rated the clarity of a statement 1 or 2, the wording of the statement was adjusted. The research team had multiple face-to-face meetings to discuss the rating and wording of statements. After this step, 24 statements remained in the questionnaire.

Step 3. Refining the Draft Questionnaire

The research team used the information from the literature and past experience to refine the draft questionnaire by adding, removing, and adjusting statements. Previous studies have indicated that users or patients are particularly interested in issues such as the locations at which their data are stored, who can access their data, the specific approaches used in handling their data, and the purpose for accessing their data [10]. Therefore, in this questionnaire, we specifically included questions related to these topics. There are many security and privacy features available in mobile apps, such as informed consent, privacy policy, access control in general, remote wipe, role-based access control, encryption, and multifactor

authentication. Therefore, we created one statement for each of these topics as well. At the end of this step, there were 17 statements in the questionnaire, and they were arranged into 3 categories: opinion on personal data security and privacy, opinion on security and privacy in mHealth apps, and desired security and privacy features in mHealth apps.

We also wrote interview questions based on the information found in our literature review and experience gained from our work in the past. For instance, possible facilitators could be apps that are free and have a low data entry burden, a clear patient protection privacy policy, an intuitive user interface, and strong but easy-to-use security features, whereas possible barriers are the opposite of these desired features (eg, paid apps, heavy data entry burden, unclear privacy policy, hard-to-use user interface or security features).

Step 4. Pilot-Testing the Questionnaire

After we all agreed on the content validity of the statements and interview questions in this study, we then distributed the first version of the questionnaire and interview questions to 14 graduate students in an information security class. These graduate students reviewed this version and provided their comments on some statements and questions. We made changes on the statements and questions according to their suggestions. For instance, almost all the students indicated that before they took the information security class, they had had no idea what role-based access control was or did not know the details about encryption. Almost all of them also indicated that they did not like to use multifactor authentication even though they knew that feature would protect their highly sensitive information. A mobile app with multifactor authentication would simply discourage them from using the app. Therefore, statements corresponding to those security features (role-based access control, encryption, and multifactor authentication) were removed from the questionnaire because if respondents did not understand those features, the results obtained would not be reliable. The final questionnaire had 14 statements.

Step 5. Performing Questionnaire and Interview Studies and Psychometric Analysis

In this step, we recruited a group of study participants to conduct studies using these new questionnaire and interview questions. The obtained data were used to evaluate the reliability and validity of the new questionnaire and answer the research questions. The details of the study and the data analysis are presented in the following sections.

Study Design

After the study protocol was approved by the institutional review board (IRB) office at the University of Pittsburgh, we recruited study participants with the following criteria: native English speaker, high school or higher education, aged between 18 and 65 years, capable of communicating with others orally and in writing, and has at least a few years of experience in using smart devices such as smartphone, tablet, or smart watch.

Study participants were recruited through flyers distributed in the Greater Pittsburgh area and through the Pitt + Me website at the University of Pittsburgh, which in January 2019 had more than 193,000 potential study participants registered at the site. Potential participants could indicate their interest in this study by sending a message to the research team or clicking on the link of the study on the Pitt + Me website. They were then screened according to the selection criteria. A list of these eligible subjects was stored in an Excel file. We then randomly selected study participants from this list to conduct the questionnaire and interview study.

Each study participant was given the opportunity to read and sign the IRB-approved consent form before the commencement of the study. At the beginning of the study, the investigators explained the purpose of the study, the procedure of the study, and the data to be collected in the study. Study participation was completely voluntary, and participants could stop participating in the study at any time.

During the study, the study participants were asked to provide answers to demographic questions, statements in the questionnaire, and the interview questions. When the study participants responded to the demographic questions and the questionnaire, the investigators did not provide any explanation on the terms used in the questionnaire. All of the answers to the questionnaire were collected with the Web-based Qualtrics system. When the study participants answered the interview questions, the investigators provided an explanation of some security terms if needed, such as encryptions, user authentication, multifactor authentication, access control, user auditing, and privacy policy. All the answers were noted, categorized, and entered into the Qualtrics system as well.

Statistical Analysis

All statistical analyses were conducted using SPSS version 24 (IBM). The internal consistency of the questionnaire was evaluated using Cronbach alpha. For research or evaluation, a value of .7 to .8 in Cronbach alpha is considered reliable [39].

Descriptive statistics were calculated for all the items in the questionnaire and the interview questions. Statistical significance was determined by $P<.05$. The normality of the data was evaluated with the Shairo-Wilk test. Nonparametric Kruskal-Wallis H (KWH) test was used to determine the significance of differences among multiple categories.

Results

Demographics

In total, 117 participants were recruited in the Greater Pittsburgh area to undertake the survey and interview study. The demographic information is summarized in Table 1.

Table 1. Demographic characteristics of the study participants (N=117).

Demographics	Statistics
Age (years), mean (SD)	31.49 (12.354)
Age (years), n (%)	
18-28	64 (54.7)
29-50	39 (33.3)
51-65	14 (12.0)
Gender, n (%)	
Male	53 (45.3)
Female	64 (54.7)
Race, n (%)	
African American	16 (13.7)
White	76 (65.0)
Asian	25 (21.4)
Education, n (%)	
Below bachelor's	39 (33.3)
Bachelor's	44 (37.6)
Graduate	34 (29.1)
Marital status, n (%)	
Single	80 (68.4)
Married	34 (29.1)
Divorced or separated	3 (2.6)
Occupation, n (%)	
Student	37 (31.6)
Health care provider	9 (7.7)
Customer service	19 (16.2)
Administrative personnel	14 (12.0)
Researcher	14 (12.0)
Other	24 (20.5)
Self-assessed health status, n (%)	
Excellent	30 (25.6)
Very good	49 (41.9)
Good	29 (24.8)
Fair	9 (7.7)
Years of using mobile devices, mean (SD)	6.21 (2.585)
Years of using mobile devices, n (%)	
<3	11 (9.4)
3-5	36 (30.8)
>5	70 (59.8)
Used mHealth apps, n (%)	
Yes	79 (67.5)
No	38 (32.5)
Household income, n (%)	
<US $10,000	20 (17.1)

Demographics	Statistics
US $10,000-US $75,000	62 (53.0)
>US $75,000	22 (18.8)
Decline to answer	13 (11.1)

Responses to the Statements on the Questionnaire

There were 14 statements on this questionnaire, and the study participants were required to select an answer from 1 to 7, corresponding to strongly agree (1), agree (2), somewhat agree (3), neither agree nor disagree (4), somewhat disagree (5), disagree (6), and strongly disagree (7).

The Cronbach alpha of this 14-item questionnaire was .730, which is good for research and exploratory studies. The overall mean for all 14 items was 2.55 (SD 0.658), reflecting that most study participants agreed with these statements to a certain degree. These 14 items were categorized into 3 groups: opinions on personal data, opinions on mHealth apps, and security features that users desire in mHealth apps.

The first category, opinion on personal data security and privacy, had 5 statements: S1, S2, S3, S4, and S5. The Cronbach alpha of this category was .737. The overall mean of the first category was 2.72 (SD 1.06), indicating that the study participants somewhat agreed with the 5 statements about personal information. They had some level of concern about the privacy of their personal data and wanted to have some specific protections on their personal data.

Similarly, the second category, opinion on security and privacy in mHealth apps, had 5 statements as well: S6, S7, S8, S9, and S10. The Cronbach alpha of this category was .785. The overall mean of the second category was 2.78 (SD 1.05), which indicates that the study participants generally accepted using mHealth apps for health care purposes, and most of them also believed that there was some level of privacy protection currently available in mHealth apps.

The third category of statements was about several security and privacy features in mHealth apps, such as informed consent (S11), access control (S12), privacy policy (S13), and remote wipe (S14). These statements fall into the same category in general; however, they are not in the same construct since each of them reflects a specific aspect of security. Therefore, it was not surprising to see that the Cronbach alpha of this category was .346. The overall mean of this group was 2.06 (SD 0.547), which reflects that the study participants desired to have those features in mHealth apps.

Table 2 provides a descriptive summary of the answers to the statements in this questionnaire. A mean of less than 4 means that these study participants agreed with the statement; the smaller the value, the stronger the agreement. A mean of greater than 4 means that the group disagreed with the statement; the bigger the value, the stronger the disagreement. The numbers in Table 2 indicate that these 117 study participants generally agreed with almost all of the statements, some showing stronger agreement, and some showing weaker. The only exception is the reported opinions on the privacy policy. It seems that many study participants did not believe that the content of the privacy policy of a mobile app could influence their decision with respect to app selection. This may be related to the readability of mHealth app privacy policies [38,40,41]. At the same time, almost all of the study participants desired to have the other 3 security features (informed consent, access control, and remote wipe) included in mHealth apps.

Relationship Between Demographic Characteristics and Answers to the Statements

The Shairo-Wilk test on the participants' answers to the 14 statements indicated that the data were not normally distributed (P<.05 in all cases). Therefore, the nonparametric KWH test was used to determine the relationship between study participants' demographic characteristics and their responses to the statements in the questionnaire. The differences were not statistically significant for people with different education levels (below bachelor's, bachelor's, and graduate), health status (excellent, very good, good, and fair), occupations, years of using mobile devices (<3, 3-5, >5), or employment status (employed vs unemployed).

Marital Status

In general, married participants had stronger concerns about information security and privacy and desired to have more stringent security protection. Correspondingly, they also were not very comfortable with their health provider using mHealth apps to manage their health data. However, in most cases, these differences were not statistically significant. A KWH test showed that there was a statistically significant difference in statement S2 between those of different marital status, KWH (1)=4.8, P=.03, with a mean rank score of 61.84 for single participants and 47.28 for married participants. A lower value corresponds to stronger agreement with the statement.

Sex

In most statements, the values from males and females were different, but these differences were not statistically significant (P>.05). The difference was statistically significant on S11 (I should have the right to consent to any sharing of my protected health information collected via mHealth apps). There, the mean rank was 67.93 for male and 51.60 for female, P=.001, KWH (1)=11.3, indicating that females have a stronger desire to have the right to consent to their health data collection.

Table 2. A summary of responses to the statements in the questionnaire. Here, agree corresponds to 1 to 3, neutral corresponds to 4, and disagree corresponds to 5 to 7 (N=117; overall Cronbach alpha=.730).

Statements	Agree (1-3), n (%)	Neutral (4), n (%)	Disagree (5-7), n (%)	Mean (SD)
Opinion on personal data (Cronbach alpha=.737)				
S1. In general, I am concerned about the privacy and security of my personal information in everyday life	97 (82.9)	7 (6.0)	13 (11.1)	2.46 (1.424)
S2. I am concerned about the privacy and security of my personal information l when using an mHealth app	82 (70.1)	12 (10.3)	23 (19.6)	2.91 (1.710)
S3. I am concerned about submitting personal information on an mHealth app because of what others might do with it	67 (57.2)	16 (13.7)	34 (29.1)	3.38 (1.731)
S4. I do not want to store my personal identifiers (such as name, SSN, phone number, email address) in the mHealth apps except for one unique ID number which is only recognizable by authorized personnel	87 (74.4)	19 (16.2)	11 (9.4)	2.65 (1.422)
S5. I would like my personal health data to be transferred to a centralized database via a highly secure process	101 (86.3)	11 (9.4)	5 (4.3)	2.19 (1.245)
Opinion on mHealth apps (Cronbach alpha=.785)				
S6. Overall, I am satisfied with the privacy and security of the mHealth apps I am currently using	80 (68.4)	28 (23.9)	9 (7.7)	2.82 (1.317)
S7. Health care providers have the necessary security and privacy measures in place. These measures provide a reasonable level of protection for information collected from mHealth apps	79 (67.5)	20 (17.1)	18 (15.4)	3.02 (1.333)
S8. I would use mHealth apps for my health care needs	95 (81.2)	10 (8.5)	12 (10.3)	2.50 (1.369)
S9. I want my health care providers to use mHealth apps to store and manage my health information	71 (60.7)	27 (23.1)	19 (16.2)	3.05 (1.479)
S10. I would feel comfortable if my health information was shared among my doctors and therapists for my health care purpose	94 (80.3)	7 (6.0)	16 (13.7)	2.52 (1.643)
Desired features in mHealth apps (Cronbach alpha=.346)				
S11. I should have the right to consent to any sharing of my protected health information collected via mHealth apps	114 (97.4)	1 (0.9)	2 (1.7)	1.39 (0.861)
S12. I would like to know how my health care providers make sure that only the correct personnel have access to the mHealth system I am using	111 (94.9)	6 (5.1)	0 (0)	1.56 (0.814)
S13. I read the privacy policies of mHealth apps. The content of the policies influences my decision of whether to use the app	51 (43.6)	15 (12.8)	51 (43.6)	3.89 (1.902)
S14. I would like to be able to remotely remove all my health data on my mobile device if it is lost or stolen	114 (97.4)	2 (1.7)	1 (0.9)	1.39 (0.719)

Race

As shown in the demographic information, three races were represented in this study: African American, white American, and Asian American. In most (13/14, 93%) of these statements, the responses among these three races do not have a statistically significant difference. There is a statistically significant difference between race groups for one statement (S1) as determined by KWH test, P=.01, χ^2_2=8.5. The mean rank was 49.69 for African Americans, 65.42 for white Americans, and 45.55 for Asian Americans. In other words, Asian Americans have a significantly stronger privacy concern than white Americans in general.

Household Income

The study participants were arranged in 4 groups according to their household income: less than US $10,000, between US $10,000 and US $75,000, greater than US $75,000, and decline to answer. In most statements (12/14, 86%), answers from the 4 groups were similar, but answers to 2 statements (S2 and S3) have statistically significant differences. For S2, χ^2_3=8.9, P=.03 and for S3, χ^2_3=9.4, P=.02. Study participants with less than US $10,000 annual income had the weakest concerns about

security and privacy. They were satisfied with the security and privacy protection provided by current mHealth apps. Therefore, they were willing to use mHealth apps themselves and also wanted their health care providers to use mHealth apps to manage their data and share their data with other doctors or therapists for health care purposes. They did not make selections on mHealth apps according to privacy policy, but they still in general agreed that data should be transmitted securely. Study participants with greater than US $75,000 annual income had the strongest concerns about and desire for security and privacy. They also expected the most stringent security measures to protect their privacy, such as the right of consent, access control, limited data stored, highly secure data transmission, and ability to remotely remove personal data if their mobile device is lost or stolen.

Age Groups

The study participants were arranged into 3 age groups: 18 to 28 years, 29 to 50 years, and 50 to 65 years. The general trend of the answers from these age groups was obvious. Participants in the 51 to 65 years age group had the strongest concern about privacy in mHealth apps and were willing to use mHealth apps for health care purposes but expected to have strong security protection in mHealth apps. Participants in the 18 to 28 years age group had the weakest concern about privacy, were somewhat satisfied with current security protection, and believed that health care providers took the necessary security and privacy measures to provide that protection. They were willing to use mHealth apps in their health care but not as strongly willing as those in the other 2 age groups. Therefore, their desire to strengthen security and privacy measures was relatively weaker compared with the other 2 age groups, but they still desired to have those measures strengthened because they did have concerns about privacy. Participants in the 29 to 50 years age group were somewhere in the middle. They had privacy concerns, did not believe the current practice was sufficient, did not have a strong desire to use mHealth apps in health care; therefore, in their response to the statements, they did not show strong desire to strengthen the security measures in mHealth apps. The differences among these age groups in multiple statements (S3, S5, S6, S8, S10, S11, S12, and S14), however, were not statistically significant.

The KWH test showed that there were statistically significant differences in answers to 6 statements (S1, S2, S4, S7, S9, and S13) among the different age groups (AG1: 18-28, AG2: 29-50, AG3: 51-65). Table 3 shows the test statistics and the mean rank for AG1 and AG3.

Experience Using Mobile Health Apps

In most cases, the means from participants who had used mHealth apps before were smaller, which indicated that these participants had a stronger concern about security and privacy in mHealth apps, but they still wanted to use mHealth apps for their health care needs and desired to have stringent security. Most of these differences were not statistically significant. These 2 groups (participants who had used or had not used mHealth apps before) had a statistically significant difference for one statement: S8 (I would use mHealth apps for my health care needs). The mean rank from the participants who had used mHealth apps before was 54.25, whereas that from the participants who had not used mHealth apps before was 68.87. Here KWH (1)=5.1, P=.02. In other words, participants who had used mHealth apps before still wanted to use mHealth apps for their health care needs, whereas participants who had not used mHealth apps before were still hesitant to commit to such a decision.

Answers to Interview Questions

The first interview question was about password change frequency. Nine (9/117, 7.7%) participants responded they would like to change their passwords every month; 44 (37.6%, 44/117) participants indicated a willingness to change their passwords every 3 months; 33 (28.2%, 33/117) participants accepted changing their passwords every 6 months; 23 (19.7%, 23/117) participants claimed to be willing to change their passwords once a year; 2 participants (1.7%, 2/117) would rather change passwords once every 2 years; and 5 participants (4.3%, 5/117) indicated that they would never change their passwords. One participant stated that he would rather not change it regularly but would like to make the change when he believed necessary.

Table 3. Test statistics in the Kruskal-Wallis test on 6 statements among 3 different age groups.

Statement	Chi-square (*df*)	*P* value	Age groups (mean rank)	
			Age group 1	Age group 3
S1	11.2 (2)	.004	67.36	38.29
S2	16.1 (2)	<.001	69.03	33.18
S4	6.7 (2)	.04	61.04	37.68
S7	7.1 (2)	.03	63.24	37.43
S9	7.5 (2)	.02	64.8	38.39
S13	8.9 (2)	.01	67.37	47.11

Table 4. Barriers to and facilitators of the use of mobile health (mHealth) apps identified in the semistructured interview study (N=117).

Questions and answers	n (%)
Q1. What barriers would prevent the adoption and integration of a mHealth App into your health monitoring and management?	
Price of mobile apps. I only use free mobile health apps.	78 (66.7)
The app sends my data to a remote server without my permission.	73 (62.4)
The app asks me to provide my personal information even when I just want to determine whether the app is good for me.	71 (60.7)
The app does not encrypt my personal data.	69 (59.0)
The app runs slowly even though my mobile device is a recent model.	67 (57.3)
The app does not have clear privacy statements about how it handles my personal data.	59 (50.4)
The app stores my personal data on my mobile device and makes the data easily accessible to anyone who can access my mobile device.	58 (49.6)
The app sends me several alerts each day.	56 (47.9)
Name of mobile app. For instance, I do not use an app where the name implies that I have a certain disease.	48 (41.0)
Price of mobile app. I only use mobile health apps costing less than $5.	26 (22.2)
The app asks me to set up an account with user name and password.	14 (12.0)
Other (eg, The app has two-factor authentication. The app requires social network login.)	3 (2.6)
Q2. What security measures would give you confidence that an mHealth app would protect the confidentiality of patient data?	
Explicit encryption on data stored on my mobile device and the data transmitted to a remote server.	96 (82.1)
User authentication.	96 (82.1)
Remote removal of my personal data on a lost mobile device.	96 (82.1)
Access control.	92 (78.6)
Easy-to-understand privacy policy which clearly indicates that my personal data are well protected.	91 (77.8)
Data transmission via a secure channel.	74 (63.2)
Easily adjustable security settings for different types of data.	74 (63.2)
All health care providers' data access activities are logged and can be audited.	70 (59.8)
One unique account for each patient and each health care provider.	67 (57.3)
Regular password update.	51 (43.6)
Q3. What specific privacy policies of an mHealth app would encourage you to use the app for your own health care purpose?	
Your data will NOT be shared with any unauthorized personnel.	98 (83.8)
Your data will be collected only if you give permission to the app.	94 (80.3)
Your data will be removed from the server if you request it.	93 (79.5)
You have the right to terminate the permission for data collection at any time.	90 (76.9)
Your data will be collected only for health care and/or research purposes.	78 (66.7)

The second interview question was about whom they were willing to share their medical information with. Participants could mention people from multiple categories. One hundred and nine (109/117, 93.2%) participants indicated that they would share their medical information with their health care providers; 81 (69.2%) participants with their family members; 38 (32.5%, 38/117) participants with friends; 12 (10.3%, 12/117) participants at password-protected personal websites, patient support groups, or password-protected online patient forums for patients with similar conditions; and 4 (3.4%, 4/117) participants claimed that they were not willing to share their medical information with anyone else.

The last 3 interview questions were about barriers to and facilitators of the use of mHealth apps. The answers from the study participants are summarized in Table 4.

Cost was a significant barrier among respondents, with a large proportion (78/117, 66.7%) indicating that they would not pay anything for a health app. Other barriers mentioned by more than half of the study participants were lack of encryption and informed consent, poor app performance, and request for personal information during app testing stage. Some study participants also mentioned issues such as unclear policy statements, too many alerts, name of mobile app (for instance, some app names include the name of a disease), and inconvenient user authentication.

The study participants also listed a number of facilitators in terms of mobile app security and privacy features, such as encryption, user authentication, user auditing, remote wipe, a clear user protecting privacy policy, and flexible security settings. Specific to the language of a privacy policy, the study participants indicated that they would like to see that they need to provide permission before the data are collected, to know the specific purpose for collecting the data, and to have the ability to stop the data from being collected.

Discussion

Principal Findings

In this study, we first confirmed that mobile app users had security and privacy concerns when they used mHealth apps in their daily life and identified the level of these concerns in people with different demographic characteristics such as sex, age, gender, education, income, and experience using mobile devices; some differences were statistically significant, whereas others were not. These results are consistent with the results reported in a number of previous studies [9,10,26-28] and the theory of privacy as contextual integrity [42]. Specific to mHealth app development, it means mHealth app developers may need to make certain adjustments to mHealth apps' security features for different user groups in terms of marital status, sex, age, and income.

More importantly, we identified the security features desired in mHealth apps using a questionnaire and interview questions, the encouraging language in privacy policies, and the specific barriers in the mHealth app adoption. The findings can be used to guide the design and development of new mHealth apps. In previous studies [9,10,26-28], the major focus was on eliciting and reporting security and privacy concerns, although the recommendations for mHealth app development were typically brief, and even if they were available, they were not from the research studies themselves but general principles of information security. The frequently mentioned security approaches were requiring user authentication (eg, password), information hiding, and informed consent [14,16]. For previous studies that did provide highly detailed security feature recommendations for mHealth app development [31], those recommendations were theoretical in terms of information security itself and did not take the user's characteristics into consideration. The results of this study offer a clearer picture in terms of security and privacy features desired by users with various characteristics in mHealth apps.

One may argue that today's smartphones already have a number of security features implemented; these include data encryption, device password lock, remote data wipe, remote device locator, and antimalware apps. Moreover, a good use of these security features can provide strong protection to users' privacy and their sensitive data such as health records; however, a recent questionnaire study involving 458 smartphone users clearly indicated that most smartphone users do not use these security features [23], and therefore, the task of data protection still falls to mHealth app developers.

According to the results of the study, the vast majority (111/117, 94.9%) of study participants desired to know how health care providers apply access control to their health data. In other words, they wanted to make sure that only authorized personnel could access their health data. The remaining 6 participants (5.1%, 6/117) did not indicate a preference on this issue. This is a topic related to patient education via the mHealth app. In other words, how can an mHealth app convince users that only authorized personnel can access their health data? To address this concern or answer this question, mHealth app developers need to demonstrate to users that corresponding security and privacy features are included in the mHealth apps.

In this study, most of the study participants had good health status even though we did not use health status as one of the selection criteria during screening. We understand that the findings could be significantly different if those with serious health problems and a strong desire to take advantage of the convenience offered by mHealth apps had been part of the population. For instance, patients who have experienced heart failure or who have had a kidney transplant have been shown to welcome mobile app–based home monitoring and reminder systems [7,43-45]. This is very common in the field of security and privacy. People perform their own risk and benefit assessment when they face a choice. The benefits provided by mHealth apps can be free cost, convenience, real-time health services, saving time, and other monetary incentives [46,47]. If users believe that the benefits outweigh the security and privacy risks, they may choose to sacrifice privacy and enjoy the benefits offered by the service, even though they still have concerns.

Limitations

The study was performed at the University of Pittsburgh, and the study participants were recruited from the Greater Pittsburgh area. Roughly one-third of the study participants were undergraduate and graduate students, and they were from many different states. Therefore, the opinions reported in this study can also reflect the opinion of people in other states of the country, or at least people in that specific age group.

The sample size of this study was not very big, but it was sufficient for the purpose of this study. In most demographic categories, the number of participants was sufficient for the analysis. We believe the current sample size is big enough for us to obtain reliable results since these participants were randomly selected from a few hundred potential participants who explicitly expressed their interest in this study. On the other hand, a larger sample size would make the results more convincing and increase the generalizability of the results.

Most of the study participants were young and healthy. Therefore, we did not see a significant difference in opinion with respect to privacy among health status groups. As we did not have very sick people included in this study, the results cannot be generalized to that population.

This study was performed in the United States, and some results may not be applicable in other countries, especially in countries with significantly different regulations and culture. For instance, in many Asian countries, a family member's health information

is openly shared with family members. Employees are also typically required to have an annual physical exam, and the results are reported to employers. Therefore, people in this situation typically do not have a strong privacy concern since there is no corresponding protection anyway.

Future Research

In the future, we will create an mHealth app with those highly desired security and privacy features identified in this study, and we will determine whether that changes app users' trust in the app. We will also perform analysis on the usage data and their security settings to determine whether they have utilized those security features, whether and how the actual app usage has changed, and which specific security features they commonly choose to disable.

It is also necessary to enhance the security education to mobile app users so that they are well aware of the many readily available security features on their smartphones and can take advantage of these features to protect their data and privacy [38].

Acknowledgments

This work was funded in part by grants from the National Science Foundation (NSF) DGE1438809 and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR) 90RE5018. The contents of this study do not represent the views of the NSF, NIDILRR, or the US government.

Conflicts of Interest

None declared.

References

1. Pew Research Center. 2018. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-03-12] [WebCite Cache ID 76ogP5274]
2. Research 2 Guidance. 2017. mHealth App Economics 2017/2018 - Current status and future trends in Mobile Health URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2019-03-12] [WebCite Cache ID 76ogpnFbK]
3. Kao CK, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
4. Roess A. The promise, growth, and reality of mobile health-another data-free zone. N Engl J Med 2017 Nov 23;377(21):2010-2011. [doi: 10.1056/NEJMp1713180] [Medline: 29116869]
5. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE, McCue MP. iMHere: a novel mHealth system for supporting self-care in management of complex and chronic conditions. JMIR Mhealth Uhealth 2013 Jul 11;1(2):e10 [FREE Full text] [doi: 10.2196/mhealth.2391] [Medline: 25100682]
6. Fairman AD, Dicianno BE, Datt N, Garver A, Parmanto B, McCue M. Outcomes of clinicians, caregivers, family members and adults with Spina Bifida regarding receptivity to use of the iMHere mHealth solution to promote wellness. Int J Telerehabil 2013 Jun;5(1):3-16. [doi: 10.5195/ijt.2013.6116] [Medline: 25945209]
7. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Perceptions and experiences of heart failure patients and clinicians on the use of mobile phone-based telemonitoring. J Med Internet Res 2012;14(1):e25 [FREE Full text] [doi: 10.2196/jmir.1912] [Medline: 22328237]
8. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]
9. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
10. Atienza AA, Zarcadoolas C, Vaughon W, Hughes P, Patel V, Chou WS, et al. Consumer attitudes and perceptions on mHealth privacy and security: findings from a mixed-methods study. J Health Commun 2015 Apr;20(6):673-679. [doi: 10.1080/10810730.2015.1018560] [Medline: 25868685]
11. Kotz D, Gunter CA, Kumar S, Weiner JP. Privacy and security in mobile health: a research agenda. Computer 2016 Jun;49(6):22-30 [FREE Full text] [doi: 10.1109/MC.2016.185] [Medline: 28344359]
12. Wiesner M, Zowalla R, Suleder J, Westers M, Pobiruchin M. Technology adoption, motivational aspects, and privacy concerns of wearables in the German running community: field study. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e201 [FREE Full text] [doi: 10.2196/mhealth.9623] [Medline: 30552085]
13. Proudfoot J, Parker G, Hadzi PD, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. J Med Internet Res 2010 Dec;12(5):e64 [FREE Full text] [doi: 10.2196/jmir.1475] [Medline: 21169174]
14. Kenny R, Dooley B, Fitzgerald A. Developing mental health mobile apps: exploring adolescents' perspectives. Health Informatics J 2016 Jun;22(2):265-275. [doi: 10.1177/1460458214555041] [Medline: 25385165]

15. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Preferences for a mobile HIV prevention app for men who have sex with men. JMIR Mhealth Uhealth 2014 Oct 29;2(4):e47 [FREE Full text] [doi: 10.2196/mhealth.3745] [Medline: 25355249]
16. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Building a mobile HIV prevention app for men who have sex with men: an iterative and community-driven process. JMIR Public Health Surveill 2015;1(2):e18 [FREE Full text] [doi: 10.2196/publichealth.4449] [Medline: 27227136]
17. Di Matteo D, Fine A, Fotinos K, Rose J, Katzman M. Patient willingness to consent to mobile phone data collection for mental health apps: structured questionnaire. JMIR Ment Health 2018 Aug 29;5(3):e56 [FREE Full text] [doi: 10.2196/mental.9539] [Medline: 30158102]
18. Office for Civil Rights. US Department of Health & Human Services. 2017. Breach Portal: Notice to the Secretary of HHS Breach of Unsecured Protected Health Information URL: https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf [accessed 2019-03-13] [WebCite Cache ID 76qUpeqBK]
19. Papageorgiou A, Strigkos M, Politou E, Alepis E, Solanas A, Patsakis C. Security and privacy analysis of mobile health applications: the alarming state of practice. IEEE Access 2018;6:9390-9403. [doi: 10.1109/ACCESS.2018.2799522]
20. Müthing J, Jäschke T, Friedrich CM. Client-focused security assessment of mHealth apps and recommended practices to prevent or mitigate transport security issues. JMIR Mhealth Uhealth 2017 Oct 18;5(10):e147 [FREE Full text] [doi: 10.2196/mhealth.7791] [Medline: 29046271]
21. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
22. Smith KA, Zhou L, Watzlaf VJ. User authentication in smartphones for telehealth. Int J Telerehabil 2017 Nov;9(2):3-12 [FREE Full text] [doi: 10.5195/ijt.2017.6226] [Medline: 29238444]
23. Mylonas A, Kastania A, Gritzalis D. Delegate the smartphone user? Security awareness in smartphone platforms. Comput Secur 2013 May;34:47-66. [doi: 10.1016/j.cose.2012.11.004]
24. Huckvale K, Prieto JT, Tilney M, Benghozi PJ, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015;13:214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
25. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. J Am Med Inform Assoc 2015 Apr;22(e1):e28-e33. [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]
26. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
27. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
28. Prasad A, Sorber J, Stablein T, Anthony D, Kotz D. Exposing privacy concerns in mHealth. In: Proceedings of the 2nd USENIX conference on Health security and privacy. 2011 Presented at: the 2nd USENIX conference on Health security and privacy; August 9; San Francisco, CA p. 2-2 URL: https://www.cs.dartmouth.edu/~trdata/reports/TR2012-711.pdf
29. Avancha S, Baxi A, Kotz D. Privacy in mobile technology for personal healthcare. ACM Comput Surv 2012 Nov 01;45(1):1-54 [FREE Full text] [doi: 10.1145/2379776.2379779]
30. Martínez-Pérez B, de la Torre-Diez I, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. J Med Syst 2015 Jan;39(1):181. [doi: 10.1007/s10916-014-0181-3] [Medline: 25486895]
31. Morera EP, de la Torre Diez I, Garcia-Zapirain B, López-Coronado M, Arambarri J. Security recommendations for mHealth apps: elaboration of a developer's guide. J Med Syst 2016 Jun;40(6):152. [doi: 10.1007/s10916-016-0513-6] [Medline: 27147515]
32. Tjora A, Tran T, Faxvaag A. Privacy vs usability: a qualitative exploration of patients' experiences with secure Internet communication with their general practitioner. J Med Internet Res 2005 May 31;7(2):e15 [FREE Full text] [doi: 10.2196/jmir.7.2.e15] [Medline: 15998606]
33. Watzlaf VJ, Zhou L, Dealmeida DR, Hartman LM. A systematic review of research studies examining telehealth privacy and security practices used by healthcare providers. Int J Telerehabil 2017 Nov;9(2):39-59 [FREE Full text] [doi: 10.5195/ijt.2017.6231] [Medline: 29238448]
34. Takyi H, Watzlaf V, Matthews JT, Zhou L, Dealmeida D. Privacy and security in multi-user health kiosks. Int J Telerehabil 2017;9(1):3-14 [FREE Full text] [doi: 10.5195/ijt.2017.6217] [Medline: 28814990]
35. Watzlaf VJ, Dealmeida DR, Zhou L, Hartman LM. Protocol for a systematic review of telehealth privacy and security research to identify best practices. Int J Telerehabil 2015 Nov;7(2):15-22 [FREE Full text] [doi: 10.5195/ijt.2015.6186] [Medline: 27563383]
36. Watzlaf VJ, Moeini S, Matusow L, Firouzan P. VOIP for telerehabilitation: a risk analysis for privacy, security and HIPAA compliance: part II. Int J Telerehabil 2011;3(1):3-10 [FREE Full text] [doi: 10.5195/ijt.2011.6070] [Medline: 25945177]
37. Watzlaf VJ, Moeini S, Firouzan P. VOIP for telerehabilitation: a risk analysis for privacy, security, and HIPAA compliance. Int J Telerehabil 2010 Oct;2(2):3-14. [doi: 10.5195/ijt.2010.6056] [Medline: 25945172]

38. Zhou L, Parmanto B, Alfikri Z, Bao J. A mobile app for assisting users to make informed selections in security settings for protecting personal health data: development and feasibility study. JMIR Mhealth Uhealth 2018 Dec 11;6(12):e11210 [FREE Full text] [doi: 10.2196/11210] [Medline: 30538088]
39. Landauer TK. Behavioral research methods in human-computer interaction. In: Helander MG, Landauer TK, Prabhu PV, editors. Handbook of Human-Computer Interaction, 2nd edition. Amsterdam, Netherlands: Elsevier; 1997:203-227.
40. Powell AC, Singh P, Torous J. The complexity of mental health app privacy policies: a potential barrier to privacy. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e158 [FREE Full text] [doi: 10.2196/mhealth.9871] [Medline: 30061090]
41. Das G, Cheung C, Nebeker C, Bietz M, Bloss C. Privacy policies for apps targeted toward youth: descriptive analysis of readability. JMIR Mhealth Uhealth 2018 Jan 04;6(1):e3 [FREE Full text] [doi: 10.2196/mhealth.7626] [Medline: 29301737]
42. Nissenbaum H. Privacy as contexual integrity. Wash Law Rev 2004;79(1):101-139 [FREE Full text]
43. Piette JD, Striplin D, Marinec N, Chen J, Aikens JE. A randomized trial of mobile health support for heart failure patients and their informal caregivers: impacts on caregiver-reported outcomes. Med Care 2015 Aug;53(8):692-699 [FREE Full text] [doi: 10.1097/MLR.0000000000000378] [Medline: 26125415]
44. Browning RB, McGillicuddy JW, Treiber FA, Taber DJ. Kidney transplant recipients' attitudes about using mobile health technology for managing and monitoring medication therapy. J Am Pharm Assoc 2016;56(4):450-454.e1. [doi: 10.1016/j.japh.2016.03.017] [Medline: 27450140]
45. McGillicuddy JW, Weiland AK, Frenzel RM, Mueller M, Brunner-Jackson BM, Taber DJ, et al. Patient attitudes toward mobile phone-based health monitoring: questionnaire study among kidney transplant recipients. J Med Internet Res 2013;15(1):e6 [FREE Full text] [doi: 10.2196/jmir.2284] [Medline: 23305649]
46. Ramanathan N, Swendeman D, Comulada WS, Estrin D, Rotheram-Borus MJ. Identifying preferences for mobile health applications for self-monitoring and self-management: focus group findings from HIV-positive persons and young mothers. Int J Med Inform 2013 Apr;82(4):e38-e46. [doi: 10.1016/j.ijmedinf.2012.05.009] [Medline: 22704234]
47. Poikela M, Schmidt R, Wechsung I, Moller S. About Your Smartphone Usage - Privacy in Location-Based Mobile Participation. 2015 Presented at: IEEE International Symposium on Technology in Society (ISTAS); Nov 11-12, 2015; Dublin, Ireland URL: https://ieeexplore.ieee.org/document/7439421

Abbreviations

AG: age group
IRB: institutional review board
KWH: Kruskal-Wallis H
mHealth: mobile health
NIDILRR: National Institute on Disability, Independent Living, and Rehabilitation Research
NSF: National Science Foundation

Edited by G Eysenbach; submitted 05.06.18; peer-reviewed by P Ware, B Ray; comments to author 06.10.18; revised version received 23.02.19; accepted 04.03.19; published 16.04.19.

Please cite as:
Zhou L, Bao J, Watzlaf V, Parmanto B
Barriers to and Facilitators of the Use of Mobile Health Apps From a Security Perspective: Mixed-Methods Study
JMIR Mhealth Uhealth 2019;7(4):e11223
URL: http://mhealth.jmir.org/2019/4/e11223/
doi:10.2196/11223
PMID:30990458

Original Paper

Quality of Blood Pressure Tracking Apps for the iPhone: Content Analysis and Evaluation of Adherence With Home Blood Pressure Measurement Best Practices

Amanda Y Leong[1,2], BScPharm; Mark J Makowsky[1], BSP, PharmD

[1]Faculty of Pharmacy & Pharmaceutical Sciences, University of Alberta, Edmonton, AB, Canada

[2]Saskatchewan Health Authority, Regina Region, Regina, SK, Canada

Corresponding Author:
Mark J Makowsky, BSP, PharmD
Faculty of Pharmacy & Pharmaceutical Sciences
University of Alberta
3-171 Edmonton Clinic Health Academy
11405 87 Avenue
Edmonton, AB, T6G 1C9
Canada
Phone: 1 (780) 492 1735
Fax: 1 (780) 492 1217
Email: makowsky@ualberta.ca

Abstract

Background: Blood pressure (BP) tracking apps may aid in hypertension (HTN) self-management, but app quality may be problematic.

Objective: This study aimed to develop a content-dependent rating system for BP tracking apps and systematically evaluate BP tracking features, content-independent quality, functional characteristics, and educational comprehensiveness of English language iPhone apps developed with the primary purpose of tracking a consumer's BP measurements.

Methods: We created a 28-item checklist reflecting overall app quality and a simplified 2-item checklist to assess adherence with home BP monitoring best practices. Apps with educational information were evaluated for comprehensiveness on a 7-point scale and for consistency with evidence-based guidelines. Higher scores represent better quality and comprehensiveness. We searched the Canadian App Store on June 28, 2016, using the keywords *hypertension* and *blood pressure*. A total of 2 reviewers independently assessed apps according to the standardized template. We determined if paid apps, educational apps, or those rated ≥4 stars were of higher quality.

Results: Of the 948 apps screened, 62 met the inclusion criteria. The mean overall quality score was 12.2 (SD 4.6, out of 28) and 6 apps (10%, 6/62) met the home BP monitoring best practice criteria. In all, 12 apps contained educational content (mean comprehensiveness 2.4, SD 1.6 out of 14), most commonly, background information on HTN. Apps with educational content (mean 15.1, SD 3.8 vs 11.8, SD 4.8; *P*=.03) or a ≥4 star rating (median 19, interquartile range [IQR] 15-20, vs 12, IQR 9-15; *P*=.02) had higher overall quality.

Conclusions: The BP tracking apps reviewed had variable quality and few met the home BP monitoring best practice criteria. When deciding to recommend a specific BP tracking app, we suggest clinicians should evaluate whether the app allows input of duplicate BP readings in the morning and evening for at least seven days and presents the mean BP value for user-specified dates. Greater attention to home BP measurement best practices is required during app development.

(JMIR Mhealth Uhealth 2019;7(4):e10809) doi:10.2196/10809

KEYWORDS

hypertension; mobile apps; self-management; blood pressure monitoring, ambulatory; cross-sectional studies

Introduction

Background

High blood pressure (BP) affects approximately 31% of adults globally [1] and affects 23% of Canadian adults [2]. In 2010, hypertension (HTN) was the leading cause of death and disability-adjusted life years worldwide; in 2015, high systolic BP accounted for 10.7 million deaths and 211.8 million disability-adjusted life years [3,4]. Lifestyle modification and medication management effectively reduce cardiovascular risk in patients with HTN. For example, a multicomponent intervention promoting increased physical activity, weight loss, reduced sodium intake, and the Dietary Approaches to Stop Hypertension (DASH) diet, reduced the HTN prevalence in those with above optimal BP at 18 months compared with those receiving advice alone—22% versus 32%; odds ratio, OR 0.77 (95% CI 0.62-0.97) [5]. Antihypertensive pharmacotherapy reduces the relative risk of myocardial infarction by 20% to 25%, stroke by 30% to 40%, and heart failure by nearly 50% [6].

Guidelines consistently recommend regular home BP monitoring for HTN management, particularly in patients with established HTN, comorbid diabetes or chronic kidney disease, suspected nonadherence, and white coat or masked HTN [7-10]. Although these recommendations are based on weak evidence and expert opinion, a recent systematic review and meta-analysis found that home BP monitoring for 6 months leads to a significant decline in systolic blood pressure (SBP) by 3.9 and diastolic blood pressure (DBP) by 2.4 mmHg versus usual care [11]. Home BP telemonitoring is also associated with larger reductions in office (SBP 4.7 mmHg, DBP 2.5 mmHg) and ambulatory BP versus usual care [12]. Recent guidelines suggest that home BP monitoring may also be used in the diagnosis of HTN [10,13-15].

Health interventions using mobile technology, that is, mobile health (mHealth), are increasingly used to provide patients and health care professionals with additional tools and resources to manage chronic disease [16], including HTN [17-20]. For example, a systematic review found that digital health interventions significantly reduced cardiovascular disease outcomes in primary and secondary prevention populations (relative risk 0.61; 95% CI 0.46-0.80, I^2=22%) but had no influence on SBP (−1.18 mmHg, 95% CI −2.93 mmHg to 0.57 mmHg, I^2=100%) [18]. In contrast, a small study using a mobile-based self-management support system significantly reduced systolic (−7 mmHg) and diastolic (−4.9 mmHg) BP over 8 weeks [21]. Furthermore, a mobile phone–based medication reminder app improved adherence and BP among patients with HTN [19]. Therefore, mHealth apps may serve to enhance BP control in patients with HTN by providing a flexible, convenient platform for patient self-management.

Rationale and Objective

Assessing the content and quality of medical apps designed for consumers may help clinicians recommend reliable and accurate apps as well as promote safe app use by patients, but only 1 published study has previously evaluated HTN apps [22-25]. In 2014, Kumar et al conducted a content analysis of the functional characteristics and consumer interaction metrics for the top 107 HTN-related apps for Apple iPhone and Google Android devices [25]. They reported that a majority of apps are designed to track BP, weight, or body mass index and concluded that greater oversight is needed in medical HTN app development, especially apps qualifying as medical devices. However, they did not conduct a formal evaluation of the quality and usefulness of the BP tracking functionalities, for example, whether available apps allow tracking of duplicate home measurements every morning and evening over a 7-day period with calculation of mean BP excluding the first day readings for clinical decision making as recommended by experts [15,26-29].

To address this gap and assist both clinicians and patients in selecting high-quality apps that could be used in the diagnosis and management of HTN, our objective was to develop a content-dependent rating system for BP tracking apps and to systematically evaluate the BP tracking features, content-independent quality, functional characteristics, and educational comprehensiveness of currently available English language iPhone apps developed with the primary purpose of tracking a consumer's home BP measurements.

Methods

Criteria for Assessment of Blood Pressure Tracking Apps

We used several international HTN clinical practice guidelines statements [26-29] and the systematic review of asthma self-management apps by Huckvale et al to guide development of criteria and domains for our BP tracking app evaluation tool (Table 1; Multimedia Appendix 1) [30,31].

Table 1. Features of the reviewed apps (N=62).

Demographics and features	Statistic
Free, n (%)	38 (61)
No option to upgrade	20 (53)
Option to upgrade	18 (47)
Paid only, mean (SD)	24 (39)
Average cost, mean (SD)	Can $2.54 (1.10)
App store category, n (%)	
Medical	34 (55)
Health & fitness	28 (45)
Presence of a star rating	12 (19)
≥4 stars	5 (8)
Country of origin, n (%)	
United States of America	14 (23)
Germany	8 (13)
Unclear	28 (45)
Other	12 (19)
Content rating, n (%)	
4+	34 (55)
12+	18 (29)
17+	10 (16)
Sponsored and created by, n (%)	
Software company	34 (55)
Pharmaceutical company	1 (2)
Medical/device company	6 (10)
Health organization	1 (2)
Individual person	20 (32)
App can transform the phone into a medical device	0 (0)
BP[a] tracking features, n (%)	
Backdate BP measurements	53 (86)
Duplicate measures QAM[b] and QHS[c] for 7 days	49 (79)
Personal notes or diary for each BP reading	39 (63)
Reminders	20 (32)
BP goal-setting	12 (19)
Sync BP data with a BP monitor	12 (19)
Data validation, n (%)	
Rational BP max & min limits	29 (47)
Flags inverted readings	16 (26)
BP analytical features, n (%)	
BP measurements are categorized	41 (66)
Appropriate course of action suggested in alert ranges	5 (8)
In-app graphing of BP measurements	55 (89)
In-app statistical analysis (overall)	36 (58)

Demographics and features	Statistic
In-app statistical analysis (unchangeable pre-set dates)	30 (48)
In-app statistical analysis (user-specified dates)	6 (10)
Data export, n (%)	50 (81)
Live sharing, n (%)	1 (2)
Cloud-based data backup, n (%)	9 (15)
HON[d] quality	
1. Health care professional involvement	4 (7)
2. Disclaimer—not a replacement for a health care provider	27 (44)
3. Privacy policy present	28 (45)
4a. App updated in past year	28 (45)
4b. BP categories have reference to scientific research	
BP categories not present	22 (36)
BP categories not referenced	31 (50)
BP categories referenced	9 (15)
5. Contact details for support	33 (55)
6. Conflict of interest or sponsorship clearly labeled	28 (45)
7. Advertising clearly distinguishable from content	12 (19)
Expanded HON criteria	
No advertising	47 (76)
Promotes a specific product	5 (8)
Functional characteristics	
BP-tracking	62 (100)
Heart rate tracking	56 (90)
Weight tracking	21 (34)
Medication tracking	11 (18)
Built in educational content	11 (18)
Exercise tracking	8 (13)
Diet tracking	4 (7)
Symptom tracking	3 (5)
Social media	4 (7)
Salt tracking	3 (5)
Lab values tracking	1 (2)
Referral to outside education resources	1 (2)
Cardiovascular risk calculator	0 (0)

[a]BP: blood pressure.

[b]QAM: each morning.

[c]QHS: at night.

[d]HON: Health on the Net.

Blood Pressure Tracking Features

BP tracking features were evaluated in 6 areas regarding the following abilities: to record duplicate morning and evening measurements for at least one week, set a BP goal, set reminders to take BP readings, sync data with a home BP monitor, backdate and time stamp readings, and log personal notes with each reading.

Blood Pressure Data Validation

BP pressure data validation was assessed using 2 items: presence of rational minimum and maximum BP limits (scrolled data

entry) or warning about improbably low or high readings (typed in readings), and warning about inverted entry of systolic and diastolic readings.

Blood Pressure Analytical Features

BP analytical features were assessed in 7 areas regarding the following abilities: to conduct in-app statistical analysis including mean BP readings, perform in-app graphing of BP measurements, categorize BP readings and red flag abnormal readings, suggest an appropriate course of action for readings in alert ranges, export data for sharing with others, automatically share readings, and perform cloud-based data backup. The app evaluation tool awarded 1 point for positive responses and 0 points for negative responses, for most items. Some items were scaled on 2-, 3-, or 4-point scales according to the level of functionality present (Multimedia Appendix 1).

Content Independent Health on the Net Quality

The quality of all apps was assessed using modified content-independent criteria created by the Health on the Net (HON) Foundation and used by Huckvale et al [30-32]. The 8 items were adapted to focus on BP tracking apps and included whether or not the app included a health care professional, as defined by the Health Professions Act on the authorship and development team, a clear purpose or disclaimer that it is not meant to replace the advice of a health care professional, a privacy policy, a recent update (ie, an update in the past 12 months), BP categories referenced using scientific literature, contact information for the app developers, a sponsorship statement and clear labeling of sponsors, and finally, a clear distinction between advertising, if present, and content. One point was awarded for positive responses and 0 points were awarded for negative responses.

Functional Characteristics

Functional characteristics were rated according to 13 items including the presence of a BP log, heart rate log, symptoms log, and cardiovascular risk calculator; trackers for exercise, diet, dietary sodium, weight, medication, and lab values (eg, sodium, potassium, serum creatinine); compatibility with social media platforms and presence of built-in educational material or referral to outside resources for HTN education.

Apps Containing Educational Material

For apps that combined BP logging features with health information, the information in these apps was assessed in 2 domains: comprehensiveness and consistency with evidence-based guidelines.

Comprehensiveness

We assessed 7 topics, whether the app contained information on BP basics, treatment options, how to use treatments, BP self-monitoring technique, a personalized action plan, how to recognize abnormal BP values, and links to health care providers. For each topic, coverage was assessed as Present in entirety (2 points), Partial (1 point), or Not present (zero points).

Consistency

We extracted key messages that were consistent among 3 international guidelines regarding BP measurement techniques (7 items) and lifestyle management (6 items) [14,26,33]. One point was awarded for information consistency and 0 points for information inconsistency, with selected statements.

Home Blood Pressure Monitoring Best Practices

Finally, given the complexity of the proposed 28-item evaluation tool, we evaluated each app against a 2-item home BP monitoring best practices criterion as suggested in major clinical practice guidelines [27-29]. Apps that allowed input of duplicate BP readings in the am and pm for at least seven days and contained in-app statistical analysis, allowing for calculation of mean BP values on user-specified dates, were deemed to meet this criterion [26-29].

A mock patient and predefined set of tasks was developed to ensure consistent, comprehensive evaluation of each app (Multimedia Appendix 2). Our full assessment form went through several iterations and rounds of testing before it was fully implemented in the review.

Search Strategy

In total, 2 search terms, *hypertension* and *blood pressure*, were used to identify English language apps focused on tracking serial BP measurements for adult members of the general public or HTN patients available in the Canadian Apple App Store. A preliminary search was conducted between May 22, 2015, and July 28, 2015, (AL) using an Apple iPad model MC769C/A, iOS version 8.4, and the final search was conducted on June 28, 2016, by a second investigator (MJM) using an Apple iPhone 7, iOS version 10.3.2, from Edmonton, Alberta, Canada.

App Selection and Data Extraction

Apps were included if the title or the app description indicated that the primary function of the app was to track BP measurements over time, the app was intended for use by the general public or HTN patients, was in English, and available in Canada. Preliminary searching indicated a large number of potentially relevant apps. Rather than limit our search to the top 50 apps as done by others [25], we narrowed the scope of the review to apps that focused on BP tracking for uncomplicated HTN. Therefore, we excluded apps advertised as whole health trackers where BP was only 1 of several tracked parameters and those that focused on HTN in the context of comorbid diabetes, chronic kidney disease, or other chronic health conditions. We excluded apps that required the purchase of a proprietary BP monitor as the only means to enter BP readings into the app. Both paid and free apps were analyzed, but when both free and paid versions were available, only free versions of apps were analyzed. Finally, we excluded apps that contained 2 or more technical or functional errors that made the app unusable and apps costing more than Can $19.99.

App name, developer, and cost were extracted for all apps by a single reviewer (MJM). Subsequently, a 2-phase screening process was completed independently by both reviewers. The first screen was based on the information provided in the App Store summary and any linked webpages. The second screen was conducted on the basis of the information in the app after it was downloaded. In cases of disagreement, a third party was asked to assess the App Store summary, and an agreement was

reached by consensus. All apps that passed the screening process had the following descriptive information recorded from the App Store: category, date of last update, version, parental rating, original release date, current version, and average user star rating.

Statistical Analysis

We calculated the scores for each respective domain by summing the individual response scores for each component. The usefulness of an app's BP tracking functionality was calculated as the sum of the BP tracking features, BP data validation, and BP analytical features domain scores. We summed the usefulness score with the HON quality score to create the app overall quality score.

We analyzed measures of central tendency for all variables and scores. Parametric data were presented as mean (SD), whereas nonparametric data were presented as median (interquartile range, IQR). We explored whether paid apps were of higher quality than free apps, whether apps with an educational component were of higher quality than those without an educational component, and whether apps with user star ratings ≥4 were of higher overall quality than those rated less than 4 stars or without a rating. Two-tailed Mann Whitney U or 2 independent sample *t* tests were used as appropriate to compare median domain scores and mean overall quality scores between groups. All data were extracted to Microsoft Excel for Mac 2011 (Microsoft Corp), and statistical analysis was conducted using SPSS (IBM SPSS Statistics for Mac, Version 24.0, IBM Corp).

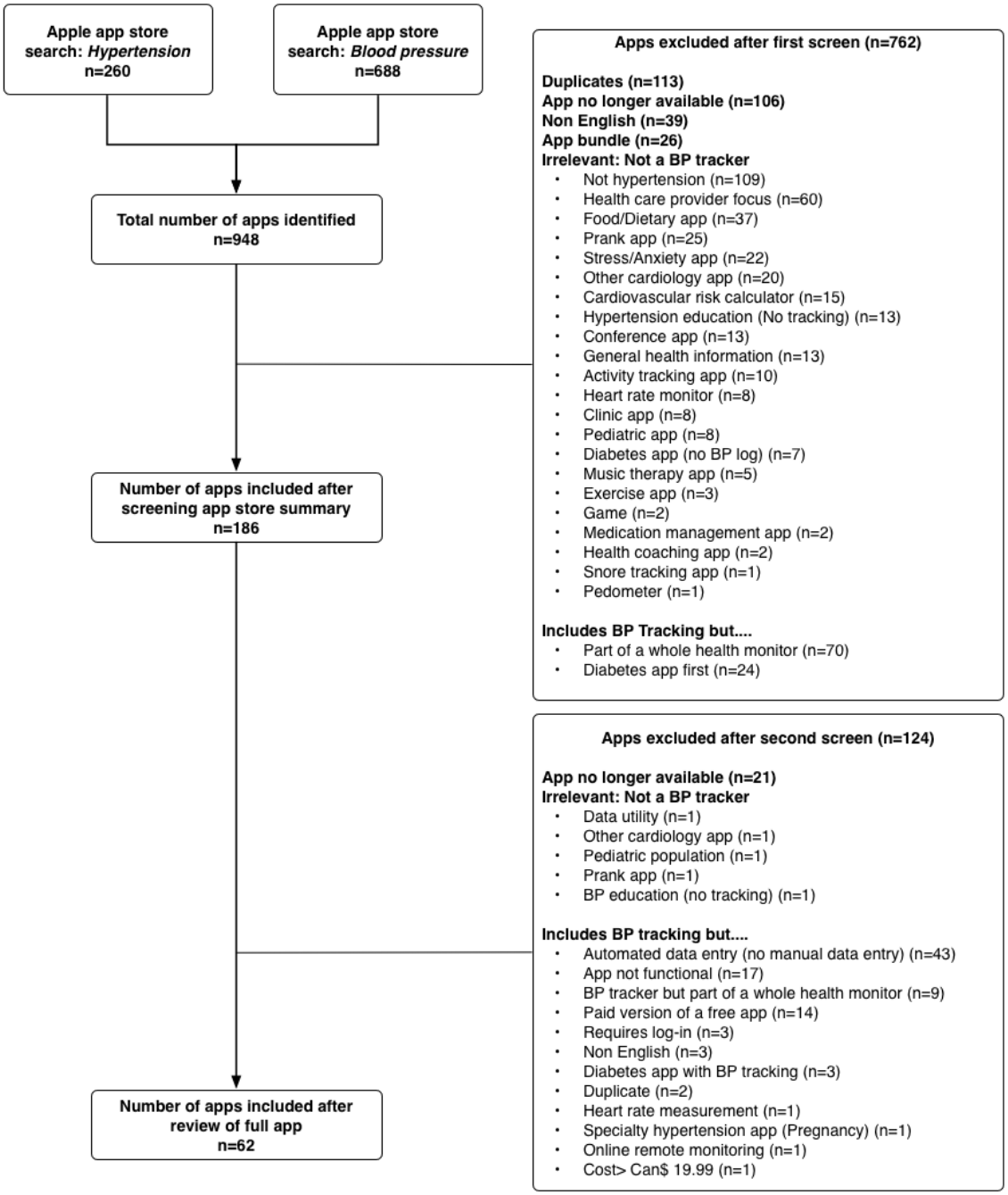
Results

Characteristics of Included Apps

Of the 948 apps screened, 62 (6.5%) met the inclusion criteria (Figure 1). A majority of apps were excluded as they were deemed not relevant to HTN tracking (41.0%; 389/948), they were no longer available for download when the assessment started in May 2017 (13.4%; 127/948), or they were duplicate records (12.1%; 115/948). Of note, several apps with BP tracking functionality were excluded (n=79 whole health monitors; n=43, which required automated transmission of data from a BP monitor; n=27, which were diabetes apps centered on blood glucose tracking; n=17, which were no longer functional; n=1, which cost more than Can $19.99; Figure 1).

Characteristics of the included apps are shown in Table 1. In total, 24 apps required payment and had an average cost of Can $2.54 (SD 1.10). Of the 38 free apps, 18 had an option to upgrade to a paid version/offered in-app purchases. All included apps were categorized as medical or health and fitness apps. Only 19% (12/62) of included apps had a user rating for the current version, and only 8% (5/62) were rated as ≥4 stars. The identified apps were most commonly created in the United States (23%; 14/62) and Germany (13%; 8/62). Most apps were sponsored by/created by a software company (53%; 34/62), an individual person (33%; 20/62), or a medical device company (9%; 6/62). A list of all the included apps is available in Multimedia Appendix 3.

Figure 1. Flow diagram. BP: blood pressure.



Overall Quality Score and Adherence With Home Blood Pressure Monitoring Best Practice Criteria

The mean overall quality score was 12.5 (SD 4.8) out of 28, and it ranged from 2 to 22 (Table 2). In total, 10% (6/62) of included apps met the proposed home BP monitoring best practice criteria.

Usefulness of Blood Pressure Tracking Functions

The mean BP tracking usefulness score was 9.7 (SD 3.7) out of 20 **(**Table 2). The mean BP tracking features domain score was 3.7 (SD 1.6) out of 7, and as shown in Table 1, most apps allowed for the entry of duplicate measures in the morning/evening for 7 days (79%; 49/62), or they allowed backdating of inputted measures (86%; 53/62), whereas a few allowed BP goal setting (19%; 12/62), reminders (32%; 20/62), or allowed automated entry of BP readings (19%; 12/62). The mean data validation domain score was 0.7 (SD 0.7; Table 2) and 47% (29/62) had rational maximum and minimum BP limits, whereas only 26% (16/62) flagged inverted systolic and diastolic BP readings. The mean analytic features domain was 5.3 (SD 2.3; Table 2). Although 58% (36/62) of the apps allowed in-app

analysis of BP readings, only 10% (6/62) allowed for in-app calculation of average BP readings using user dates (Table 1). Most allowed for in-app graphing of BP readings (89%; 55/62), data export (81%; 50/62), or categorization of BP measurements (66%; 41/62), whereas only 8% (5/62) suggested an appropriate course of action when BP readings were in alert ranges.

Table 2. Hypertension app scores for functionality, usefulness, and quality (n=62).

Domain	Potential score range	Mean (SD)	Median (IQR[a])	Minimum	Maximum
Overall quality score[b]	0-28	12.5 (4.8)	12 (9-16)	2	22
BP[c] tracking usefulness score	0-20	9.7 (3.7)	10 (8-13)	2	18
BP tracking features	0-7	3.7 (1.6)	4 (2-5)	1	7
Data validation	0-2	0.7 (0.7)	1 (0-1)	0	2
BP analytical features	0-11	5.3 (2.3)	6 (4-7)	0	10
Health on the Net quality	0-8	2.8 (1.7)	2 (1-4)	0	6
Functional characteristics	0-13	3.0 (1.5)	2.5 (2-4)	1	9
Education comprehensiveness (n=12)	0-14	2.4 (1.6)	2 (1-3)	0	6
BP measurement content (n=3)	0-7	3.3 (1.2)	4 (3-4)	2	4
Lifestyle content (n=7)	0-6	1.7 (1.4)	2 (0-3)	0	3

[a]IQR: interquartile range.

[b]Overall quality score: BP tracking features, data validation, BP analytical features, and Health on the Net quality.

[c]BP: blood pressure.

Health on the Net Quality Scores

The median HON quality score was low (2, IQR 1-4 out of 8, Table 2). As shown in Table 1, just under half (45%; 28/62) of the apps were updated in the past year, contained a privacy policy (44%; 28/62), had contact details or support (55%; 33/62), and had clearly labeled sponsorships (45%; 28/62). A minority had health care professional involvement, a disclaimer, and references to scientific research for BP ranges. Only 15% (9/62) referenced scientific literature for the BP category cut-offs used.

Functional Characteristics

The median number of functional characteristics was 2.5 (IQR 2-4, Table 2). Most apps had at least one additional feature beyond BP and heart rate tracking (53%; 33/62). Most common were weight tracking (34%; 21/62), educational content (18%; 11/62), or medication tracking (18%; 11/62; Table 1). Although 1 app had 9 different functions, few apps (26%; 16/62) had 4 or more built-in functions.

Apps Containing Educational Material

The median educational comprehensiveness in the 12 apps that contained educational content was 2 (IQR 1.5-2) out of 14 (Table 2)**.** Most apps contained basic information about BP and HTN (75%; 9/12), lifestyle management options (58%; 7/12), and encouraged sharing of BP readings with health care professionals (42%; 5/12), whereas few (25%; 3/12) provided information about BP measurement technique (Table 3).

Table 3. Assessment of apps containing educational content.

Features	Statistics, n (%)
Comprehensiveness (n=12)	
Basic BP[a] and hypertension information	9 (75)
Lifestyle treatment options	7 (58)
How to use treatment	2 (17)
BP self-monitoring technique	3 (25)
Personalized action plan (HBPM[b] treatment goal <135/85 mmHg)	0 (0)
Recognition of abnormally high or low BP values	2 (17)
Link to health care provider	5 (42)
BP measurement (n=3)	
Wait 30 min after coffee or smoking	3 (100)
Relax before BP reading	3 (100)
Body positioning	2 (67)
Consistently measure BP the same arm	2 (67)
Validated BP device	0 (0)
BP measured ≥2 times QAM[c] & QHS[d] ≥7 days	0 (0)
Arm automatic monitor with well fitted cuff	0 (0)
Lifestyle content (n=7)	
Advice to quit smoking	4 (57)
Exercise	3 (43)
Body mass index or waist circumference targets	0 (0)
Alcohol restriction	1 (14)
Heart healthy diet	3 (43)
Goal daily sodium: 2 g	1 (14)
Written in plain language	10 (83)
Grammatical and spelling errors present	3 (25)

[a]BP: blood pressure.

[b]HBPM: home-blood pressure measurement.

[c]QAM: each morning.

[d]QHS: at night.

Exploring Influence of Cost, Educational Content, and User Rating

There were no statistically significant differences in functional characteristics, usefulness, HON quality, or overall quality between free and paid apps (Table 4). Apps with an educational component had higher overall quality scores than those without an educational component (mean 15.1, SD 3.8 vs 11.8, SD 4.8; *P*=.03). Although only 5 apps were rated ≥4 stars, those with this rating had a higher overall quality score (median 19, IQR 15-20, vs 12, IQR 9-15; *P*=.02). These appeared to be driven by both higher usefulness and HON quality scores.

Table 4. Comparison of hypertension app domain and overall quality scores by cost, presence of educational content, and user rating (n=62).

Domain	Free (n=38)	Paid (n=24)	*P* value	No education (n=50)	Education (n=12)	*P* value	<4 stars (n=57)	≥4 stars (n=5)	*P* value
Overall quality score	12.7 (5.2)[a]	12.1 (4.0)[a]	.63	11.8 (4.8)[a]	15.1 (3.8)[a]	.03	12 (9-15)[b]	19 (15-20)[b]	.02
BP[c] tracking usefulness score	9.8 (4.1)[a]	9.5 (3.1)[a]	.71	9.2 (3.7)[a]	11.5 (3.0)[a]	.06	10 (7-12)[b]	13 (11-15)[b]	.04
BP tracking features, median (IQR[d])	4 (2-5)	4 (3-4)	.41	4 (2-5)	4 (3-5)	.32	4 (2-5)	5 (4-6)	.12
Data validation. median (IQR)	1 (0-1)	1 (0-1)	.55	1 (0-1)	1 (0-2)	.20	1 (0-1)	1 (1-2)	.47
BP analytical features, median (IQR)	6 (4-7)	5.5 (4-7)	.80	5 (3-7)	7 (5-7)	.09	5 (4-7)	7 (6-8)	.05
HON[e] quality, median (IQR)	2.5 (1-5)	2 (2-4)	.73	2 (1-4)	3.5 (2-5)	.07	2 (1-4)	5 (4-6)	.01
Functional characteristics, median (IQR)	3 (2-4)	2 (2-3)	.44	2 (2-3)	3.5 (3-7)	<.001	2 (2-4)	3 (3-5)	.22
Education comprehensiveness, median (IQR)	2 (2-5)	2 (1-2)	.20	—[f]	—	—	—	—	—

[a]Data are presented as mean (SD).

[b]Data are presented as median (IQR). Aggregate BP usefulness and overall quality scores were normally distributed for each comparison group except by 4-star rating. All single-item domain scores were not normally distributed.

[c]BP: blood pressure.

[d]IQR: interquartile range.

[e]HON: Health on the Net.

[f]Not applicable.

Discussion

Principal Findings

In this review, we found concerning gaps in BP tracking features and large variation in the overall quality of the 62 reviewed apps according to the evaluation tool that we developed. In total, only 6 apps met the home BP measurement best practice criteria, and although most apps allowed the entry of duplicate BP measures in the morning and evening, most did not allow for in-app statistical analysis on the basis of user-specified rather than preset fixed dates based on the date of accessing the app. Of note, no apps automatically calculated the mean BP value on the basis of the last 6 days of readings, as recommended by some organizations, and only apps that allowed users to specify the dates included allowed this calculation. Few apps flagged readings that were incorrectly input as diastolic over systolic, or they suggested an appropriate course of action when BP readings were in alert ranges. In addition, most apps used reference BP ranges set in the Joint National Committee 6 or 7 reports on HTN, which were based on office readings and none were based on the new 2017 American College of Cardiology/American Heart Association hypertension guideline [7,15]. Although some apps allowed for customization of BP target values, the treatment targets were defined on the basis of in-office measures. The content-independent HON quality was low, with the majority scoring poorly on these criteria. None of the included apps included the ability to turn the iPhone into a medical device, and although 6 apps were designed for use with a specific brand of BP monitor, none of the reviewed apps connected with a Hypertension Canada-endorsed BP monitor.

Few of the reviewed apps contained educational content, and in those that did, the materials were not comprehensive. Educational material generally focused on the basic BP background information or treatment options. Most apps were lacking key information on BP measurement technique, and none advocated the use of validated BP measurement devices, duplicate readings twice per day for 7 days, and use of an appropriately sized cuff.

We found that apps with an educational component or an App Store rating ≥4 stars were of higher overall quality compared with those without an educational component or ≤3 stars. Despite this, only 5 out of the top 10 overall quality apps and only 1 of the 6 apps meeting the best practice criteria had star ratings. On the basis of this, we suggest the use of a simplified ranking system that is not only primarily based on consistency with home BP monitoring best practice criteria but also takes into consideration the presence of educational material on BP measurement and the App Store rating (Textbox 1). Such a tool may be directly helpful for clinicians in making recommendations to patients or others regarding BP tracking apps.

Textbox 1. Proposed simplified, content-dependent criteria to evaluate blood pressure-tracking app quality.

Review the app store description, screenshots, and download if necessary.

- Does the app conform to the recognized home blood pressure (BP) best practice criteria? [15,27-29]
  - Does the app allow input of duplicate BP readings for at least seven days?
  - Does the app contain *in-app* statistical analysis, which displays the mean BP values on user-specified dates?
- Does the app contain educational material on BP measurement technique or BP in general?
- Is the app rated 4 or more stars?

Strengths and Limitations

Our study has several strengths, including a comprehensive search to identify all relevant BP tracking apps for the iOS platform, a rigorous app assessment process based on best practices from the app evaluation literature [22], and a duplicate review of all apps by independent investigators [22]. Despite the strengths, there are several limitations. First, we did not review BP tracking apps designed for the Android platform; therefore, we are potentially missing a significant number of unique apps developed exclusively for this platform. Data from 2017 suggest that approximately 46% and 43% of the smartphone market share was held by Android devices in Canada and the United States, respectively [34]. Although Kumar et al found no overlap in the top 5 most popular HTN apps by the number of downloads for iPhone and Android, there was a large degree of overlap in apps identified in Apple iTunes for iPhone and Google Play Store [25]. Second, we could only access the Canadian Apple App Store, and therefore we may not have captured apps available in other countries. Despite this, our review is applicable to clinicians internationally as we have provided both a detailed and simplified content-dependent app ranking system, the latter of which is directly applicable to clinicians in recommending useful home BP monitoring apps, regardless of the country. Third, by excluding apps that absolutely required automated transmission of data from a *smart* BP monitor to populate data into the app (eg, those from Withings, Qardioarm, and iHealth), excluding apps that were focusing on comorbid chronic conditions or were whole-health monitors, our results primarily reflect a specific subset of manual BP tracking apps available in the iOS market before the widespread availably of *smart* BP monitors. Automated *smart* BP monitors that have the ability to automatically populate data from a connected BP monitor into an app may score higher on our assessments. Although multimorbidity is common in primary care, and patients with HTN commonly have other manifestations of coronary heart disease or other chronic conditions [35], we feel justified in excluding these types of apps, as it was felt that consumers wanting to find an app to track their BP would preferentially pick 1 that did so as its primary function. In addition, by excluding apps focused on diabetes, we avoided issues associated with the controversy surrounding BP treatment targets [36]. Although we did not see major differences in paid and free apps, limiting our review to free versions of apps with a paid version may have artificially biased the usefulness and quality scores in a downward direction. Fourth, the recommendations for evaluation of mobile phone apps continue to evolve [23], and we did not use newer validated app-rating scales such as the Mobile Application Rating Scale [37]. However, our 2-item assessment tool using home BP monitoring best practice criteria was based on expert consensus recommendations from major organizations like Hypertension Canada and the American College of Cardiology/American Heart Association, which advocate for recording duplicate BP measures twice daily for at least seven days and using the mean home BP value to diagnose HTN in relation to a cut-off SBP ≥135 mmHg or DBP ≥85 mmHg [15,26,38]. Fifth, given the dynamic nature of apps in terms of content, updates, and changes in availability, the overall app quality scores and any recommendations regarding the use of specific apps are subject to change as apps are updated, new apps are released, and others are removed. Finally, by not involving patients in the app review, important components that may impact effective app use and usability (ie, limited health or technological literacy) were not systematically assessed in our review.

Comparison With Earlier Work

The only other review of HTN self-management apps was performed by Kumar et al, who performed a content analysis of the top 50 apps for HTN for Android and iOS devices [25]. They found that 72% contained a tracking function, 22% had tools to enhance medication adherence, 37% contained general information on HTN, and 8% contained information on the DASH diet. In addition, they found that only 3% of the apps were developed by health care agencies (ie, universities or professional organizations). In contrast, our review only focused on patient-oriented apps that contained BP tracking functions and found that in these apps, 35% contained adherence tools and 18% had built-in educational content. We also found that only a minority of available apps were developed by health care agencies. Kumar et al found that the ability of apps to track BP was significantly associated with the number of app downloads. We did not explore this relationship as Apple does not release iOS app download statistics.

Our finding of a broad variation in the app quality is consistent with previous reviews of health-related mobile phone apps in the areas of diabetes, smoking cessation, pain and pulmonary management. Demidowich et al found that only 4 of 42 Android apps targeted toward diabetes self-management had sufficiently high composite usability scores, which suggested that few apps provided a comprehensive method of diabetes management [39]. In our study, only 7 apps had overall quality scores ≥19 and only 6 met the best practices criteria. Our adapted HON quality score results are similar to those of Huckvale et al, who

found a generally low quality of asthma self-management apps using the HON criteria [30]. In our data, it appeared as though user star ratings may be helpful to identify a quality app, but these data are not robust, and others have found poor correlation between user-star ratings and app usability scores [39].

This study found only 12 apps that contained any sort of educational component, with no app scoring above 50% for educational quality. This demonstrates that these apps were primarily designed to be tracking tools, with little emphasis on comprehensive, high-quality educational material. Similarly, Abroms et al documented that few smoking cessation iPhone apps adhered to key guidelines and provided recommendations or linked users to proven treatments, such as pharmacotherapy, counseling, and/or quit lines [40]. Huckvale et al evaluated 103 asthma self-management apps, 38 of which contained an educational component [30]. Similar to their study, no HTN app addressed all aspects of the guidelines. In their updated 2015 study just over half (57%; 83 of 147) of apps provided educational information about asthma [31]. Therefore, it appears that asthma apps have more robust, evidence-based chronic disease self-management information than apps designed for BP tracking.

Conclusions

A handful of apps explicitly developed and marketed for BP tracking are adherent to home BP monitoring best practices, as set out by clinical practice guidelines, and score highly on overall quality. However, several concerning gaps exist in the current BP tracking apps. Although app store ratings and the presence of educational content may help clinicians or patients choose higher quality apps, many high-quality apps did not have consumer ratings. At minimum, we suggest clinicians should evaluate whether a BP tracking app allows input of duplicate BP readings for at least seven days and presents the mean BP value for user-specified dates. There remain opportunities to improve the overall quality of patient-focused BP tracking apps and incorporate evidence-based HTN education to further optimize patient self-management of HTN.

Acknowledgments

The investigators wish to thank the Rx&D Health Research Foundation Summer studentship and the Lawson Foundation (Grant number GRT 2012-057) for supporting this work. We also wish to acknowledge Ms. Darlene Korn, currently a clinical pharmacist with Alberta Health Services, for assisting with the early phases of app evaluation, which occurred in 2012 to 2013, and Dr. Ann Thompson, Clinical Associate Professor, Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, for acting as the third-party adjudicator.

Authors' Contributions

AL and MM contributed to conception and design. AL and MM contributed to acquisition of data, analysis, and interpretation. AL prepared the first draft of the manuscript. Critical revision of manuscript for intellectual content was done by AL and MM. Final approval of the version to be published was granted by AL and MM.

Conflicts of Interest

MM owns stock in Apple Corporation but has no financial stock or interests in any of the reviewed Apps. AL declares no conflicts of interest.

Multimedia Appendix 1

Criteria for app assessment.

[PDF File (Adobe PDF File), 148KB - mhealth_v7i4e10809_app1.pdf]

Multimedia Appendix 2

Mock patient used for app evaluation.

[PDF File (Adobe PDF File), 83KB - mhealth_v7i4e10809_app2.pdf]

Multimedia Appendix 3

List of included apps.

[PDF File (Adobe PDF File), 84KB - mhealth_v7i4e10809_app3.pdf]

References

1. Mills KT, Bundy JD, Kelly TN, Reed JE, Kearney PM, Reynolds K, et al. Global disparities of hypertension prevalence and control: a systematic analysis of population-based studies from 90 countries. Circulation 2016 Aug 09;134(6):441-450. [doi: 10.1161/CIRCULATIONAHA.115.018912] [Medline: 27502908]

2. Padwal RS, Bienek A, McAlister FA, Campbell NRC, Outcomes Research Task Force of the Canadian Hypertension Education Program. Epidemiology of hypertension in Canada: an update. Can J Cardiol 2016 Dec;32(5):687-694. [doi: 10.1016/j.cjca.2015.07.734] [Medline: 26711315]
3. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
4. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 08;388(10053):1659-1724 [FREE Full text] [doi: 10.1016/S0140-6736(16)31679-8] [Medline: 27733284]
5. Elmer PJ, Obarzanek E, Vollmer WM, Simons-Morton D, Stevens VJ, Young DR, et al. Effects of comprehensive lifestyle modification on diet, weight, physical fitness, and blood pressure control: 18-month results of a randomized trial. Ann Intern Med 2006 Apr 4;144(7):485-495. [Medline: 16585662]
6. Blood Pressure Lowering Treatment Trialists' Collaboration, Turnbull F, Neal B, Ninomiya T, Algert C, Arima H, et al. Effects of different regimens to lower blood pressure on major cardiovascular events in older and younger adults: meta-analysis of randomised trials. Br Med J 2008 May 17;336(7653):1121-1123 [FREE Full text] [doi: 10.1136/bmj.39548.738368.BE] [Medline: 18480116]
7. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure, National High Blood Pressure Education Program Coordinating Committee. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. J Am Med Assoc 2003 May 21;289(19):2560-2572. [doi: 10.1001/jama.289.19.2560] [Medline: 12748199]
8. Weber MA, Schiffrin EL, White WB, Mann S, Lindholm LH, Kenerson JG, et al. Clinical practice guidelines for the management of hypertension in the community: a statement by the American Society of Hypertension and the International Society of Hypertension. J Clin Hypertens (Greenwich) 2014 Jan;16(1):14-26 [FREE Full text] [doi: 10.1111/jch.12237] [Medline: 24341872]
9. Go AS, Bauman MA, Coleman KSM, Fonarow GC, Lawrence W, Williams KA, et al. An effective approach to high blood pressure control: a science advisory from the American Heart Association, the American College of Cardiology, and the Centers for Disease Control and Prevention. J Am Coll Cardiol 2014 Apr 01;63(12):1230-1238 [FREE Full text] [doi: 10.1016/j.jacc.2013.11.007] [Medline: 24246165]
10. Nerenberg KA, Zarnke KB, Leung AA, Dasgupta K, Butalia S, McBrien K, Hypertension Canada. Hypertension Canada's 2018 guidelines for diagnosis, risk assessment, prevention, and treatment of hypertension in adults and children. Can J Cardiol 2018 May;34(5):506-525. [doi: 10.1016/j.cjca.2018.02.022] [Medline: 29731013]
11. Uhlig K, Patel K, Ip S, Kitsios GD, Balk EM. Self-measured blood pressure monitoring in the management of hypertension: a systematic review and meta-analysis. Ann Intern Med 2013 Aug 6;159(3):185-194. [doi: 10.7326/0003-4819-159-3-201308060-00008] [Medline: 23922064]
12. Omboni S, Gazzola T, Carabelli G, Parati G. Clinical usefulness and cost effectiveness of home blood pressure telemonitoring: meta-analysis of randomized controlled studies. J Hypertens 2013 Mar;31(3):455-467. [doi: 10.1097/HJH.0b013e32835ca8dd] [Medline: 23299557]
13. National Institute of Health and Clinical Excellence. Hypertension in adults: diagnosis and management URL: https://www.nice.org.uk/guidance/cg127 [accessed 2019-02-25] [WebCite Cache ID 76Sms4Bfs]
14. Mancia G, Fagard R, Narkiewicz K, Redón J, Zanchetti A, Böhm M, Task Force Members. 2013 ESH/ESC Guidelines for the management of arterial hypertension: the Task Force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). J Hypertens 2013 Jul;31(7):1281-1357. [doi: 10.1097/01.hjh.0000431740.32696.cc] [Medline: 23817082]
15. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison HC, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension 2017 Nov 13;71(6):e13-155. [doi: 10.1161/HYP.0000000000000065] [Medline: 29133356]
16. Silva BM, Rodrigues JJ, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
17. Logan AG. Transforming hypertension management using mobile health technology for telemonitoring and self-care support. Can J Cardiol 2013 May;29(5):579-585. [doi: 10.1016/j.cjca.2013.02.024] [Medline: 23618506]
18. Widmer RJ, Collins NM, Collins CS, West CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. Mayo Clin Proc 2015 Apr;90(4):469-480. [doi: 10.1016/j.mayocp.2014.12.026] [Medline: 25841251]

19. Patel S, Jacobus-Kantor L, Marshall L, Ritchie C, Kaplinski M, Khurana PS, et al. Mobilizing your medications: an automated medication reminder application for mobile phones and hypertension medication adherence in a high-risk urban population. J Diabetes Sci Technol 2013 May 01;7(3):630-639 [FREE Full text] [doi: 10.1177/193229681300700307] [Medline: 23759395]
20. Omboni S, Caserini M, Coronetti C. Telemedicine and m-health in hypertension management: technologies, applications and clinical evidence. High Blood Press Cardiovasc Prev 2016 Sep;23(3):187-196. [doi: 10.1007/s40292-016-0143-6] [Medline: 27072129]
21. Bengtsson U, Kjellgren K, Hallberg I, Lindwall M, Taft C. Improved blood pressure control using an interactive mobile phone support system. J Clin Hypertens (Greenwich) 2016 Feb;18(2):101-108 [FREE Full text] [doi: 10.1111/jch.12682] [Medline: 26456490]
22. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]
23. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. J Telemed Telecare 2016 Oct 18;24(1):22-30. [doi: 10.1177/1357633X16673538] [Medline: 27760883]
24. Parati G, Torlasco C, Omboni S, Pellegrini D. Smartphone applications for hypertension management: a potential game-changer that needs more control. Curr Hypertens Rep 2017 Jun;19(6):48. [doi: 10.1007/s11906-017-0743-0] [Medline: 28501982]
25. Kumar N, Khunger M, Gupta A, Garg N. A content analysis of smartphone-based applications for hypertension management. J Am Soc Hypertens 2015 Feb;9(2):130-136. [doi: 10.1016/j.jash.2014.12.001] [Medline: 25660364]
26. Daskalopoulou SS, Rabi DM, Zarnke KB, Dasgupta K, Nerenberg K, Cloutier L, et al. The 2015 Canadian Hypertension Education Program recommendations for blood pressure measurement, diagnosis, assessment of risk, prevention, and treatment of hypertension. Can J Cardiol 2015 May;31(5):549-568. [doi: 10.1016/j.cjca.2015.02.016] [Medline: 25936483]
27. Celis H, Den Hond E, Staessen JA. Self-measurement of blood pressure at home in the management of hypertension. Clin Med Res 2005 Feb;3(1):19-26 [FREE Full text] [Medline: 15962017]
28. O'Brien E, Asmar R, Beilin L, Imai Y, Mallion J, Mancia G, European Society of Hypertension Working Group on Blood Pressure Monitoring. European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement. J Hypertens 2003 May;21(5):821-848. [doi: 10.1097/01.hjh.0000059016.82022.ca] [Medline: 12714851]
29. Williams B, Poulter NR, Brown MJ, Davis M, McInnes GT, Potter JF, British Hypertension Society. Guidelines for management of hypertension: report of the fourth working party of the British Hypertension Society, 2004-BHS IV. J Hum Hypertens 2004 Mar;18(3):139-185. [doi: 10.1038/sj.jhh.1001683] [Medline: 14973512]
30. 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]
31. 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;13:58 [FREE Full text] [doi: 10.1186/s12916-015-0303-x] [Medline: 25857569]
32. Health on the Net Foundation. The commitment to reliable health and medical information on the internet URL: http://healthonnet.org/HONcode/Pro/Visitor/visitor.html [accessed 2017-09-11] [WebCite Cache ID 76Sn6J8fz]
33. Parati G, Stergiou GS, Asmar R, Bilo G, de Leeuw P, Imai Y, et al. European Society of Hypertension guidelines for blood pressure monitoring at home: a summary report of the Second International Consensus Conference on Home Blood Pressure Monitoring. J Hypertens 2008 Aug;26(8):1505-1526. [doi: 10.1097/HJH.0b013e328308da66] [Medline: 18622223]
34. Device Atlas. Android v iOS market share 2019 URL: https://deviceatlas.com/blog/android-v-ios-market-share [accessed 2019-02-25] [WebCite Cache ID 76SnMrv0k]
35. Guthrie B, Payne K, Alderson P, McMurdo ME, Mercer SW. Adapting clinical guidelines to take account of multimorbidity. Br Med J 2012 Oct 04;345:e6341. [Medline: 23036829]
36. Rabi DM, Padwal R, Tobe SW, Gilbert RE, Leiter LA, Quinn RR, Canadian Hypertensive Education Program, Canadian Diabetes Association. Risks and benefits of intensive blood pressure lowering in patients with type 2 diabetes. Can Med Assoc J 2013 Aug 06;185(11):963-967 [FREE Full text] [doi: 10.1503/cmaj.120112] [Medline: 23734033]
37. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
38. Shimbo D, Abdalla M, Falzon L, Townsend RR, Muntner P. Role of ambulatory and home blood pressure monitoring in clinical practice: a narrative review. Ann Intern Med 2015 Nov 03;163(9):691-700 [FREE Full text] [doi: 10.7326/M15-1270] [Medline: 26457954]
39. Demidowich AP, Lu K, Tamler R, Bloomgarden Z. An evaluation of diabetes self-management applications for Android smartphones. J Telemed Telecare 2012 Jun;18(4):235-238. [doi: 10.1258/jtt.2012.111002] [Medline: 22604278]
40. Abroms LC, Padmanabhan N, Thaweethai L, Phillips T. iPhone apps for smoking cessation: a content analysis. Am J Prev Med 2011 Mar;40(3):279-285 [FREE Full text] [doi: 10.1016/j.amepre.2010.10.032] [Medline: 21335258]

Abbreviations

BP: blood pressure
DASH: Dietary Approaches to Stop Hypertension
DBP: diastolic blood pressure
HON: Health on the Net
HTN: hypertension
IQR: interquartile range
mHealth: mobile health
SBP: systolic blood pressure

Edited by G Eysenbach; submitted 18.04.18; peer-reviewed by N Kumar, S Ong, S Omboni; comments to author 07.08.18; revised version received 02.01.19; accepted 26.01.19; published 12.04.19.

Please cite as:
Leong AY, Makowsky MJ
Quality of Blood Pressure Tracking Apps for the iPhone: Content Analysis and Evaluation of Adherence With Home Blood Pressure Measurement Best Practices
JMIR Mhealth Uhealth 2019;7(4):e10809
URL: https://mhealth.jmir.org/2019/4/e10809/
doi:10.2196/10809
PMID:30977739

Original Paper

Evaluation of Mobile Apps Targeted to Parents of Infants in the Neonatal Intensive Care Unit: Systematic App Review

Brianna Richardson[1,2,3], BScN, RN; Justine Dol[2,3], MSc; Kallen Rutledge[3], MLIS; Joelle Monaghan[3], MN, RN; Adele Orovec[3], BScMS; Katie Howie[3], BScN, RN, PNC; Talia Boates[3], BScN, RN; Michael Smit[4], PhD; Marsha Campbell-Yeo[1,3,5,6], NNP-BC, RN, PhD

[1]School of Nursing, Dalhousie University, Halifax, NS, Canada

[2]Faculty of Health, Dalhousie University, Halifax, NS, Canada

[3]Centre for Pediatric Pain, Izaak Walton Killam Health Centre, Halifax, NS, Canada

[4]School of Information Management, Dalhousie University, Halifax, NS, Canada

[5]Division of Neonatal Perinatal Medicine, Department of Pediatrics, Faculty of Medicine, Dalhousie University, Halifax, NS, Canada

[6]Division of Neonatal Perinatal Medicine, Department of Pediatrics, Faculty of Medicine, Izaak Walton Killam Health Centre, Halifax, NS, Canada

Corresponding Author:
Marsha Campbell-Yeo, NNP-BC, RN, PhD
School of Nursing
Dalhousie University
PO Box 15000, Forrest Bldg
5869 University Avenue
Halifax, NS, B3H4R2
Canada
Phone: 1 9024944283
Email: marsha.campbell-yeo@dal.ca

Abstract

Background: Parents of preterm infants increasingly use their mobile phone to search for health information. In a recent review, websites targeted toward parents with infants in the neonatal intensive care unit (NICU) were found to have poor to moderate quality educational material; however, there is a dearth of literature regarding mobile apps for NICU parents.

Objective: This study aimed to identify and evaluate apps targeting parents of infants in the NICU for quality of information, usability, and credibility.

Methods: We systematically searched the Apple App Store and Google Play using 49 key terms (eg, “preterm infant”) from July 26 to August 18, 2017. English apps targeting NICU parents that cost less than $20 were included. Apps for health care professionals, e-books/magazines, or nonrelevant results were excluded. In total, 3 tools were used for evaluation: Mobile Application Rating Scale (MARS) to measure quality; Patient Education Materials Assessment Tool for Audiovisual Materials (PEMAT-AV) to measure the app’s content usability; and Trust it or Trash It to measure credibility.

Results: The initial search yielded 6579 apps, with 49 apps eligible after title and description screening. In total, 27 apps met the eligibility criteria with 9 apps available in both app stores; of those, the app with the most recent update date was chosen to be included in the analysis. Thus, 18 unique apps were included for final analysis. Using MARS, 7 apps (7/18, 39%) received a good score on overall quality (ie, 4.0 out of 5.0), with none receiving an excellent score. In addition, 8 apps (8/18, 44%) received a PEMAT-AV score between 51% and 75% on the understandability subscale, and 8 apps (8/18, 44%) scored between 76% and 100% on the actionability subscale. Trust It or Trash It deemed 13 apps (13/18, 72%) as *trash* for reasons including no identification of sources or lack of current information, with only 5 (5/18, 28%) deemed trustworthy. Reviewer’s expert evaluation found 16 apps contained content that matched information provided by multiple sources; however, most apps did not meet other objective measurement items to support credibility. When comparing the MARS overall quality and subjective quality scores with trustworthiness of apps, there was no statistically significant difference. A statistically significant difference was found between the 2 MARS quality scores, indicating that, on average, apps were ranked significantly lower on subjective quality compared with overall quality measures.

Conclusions: This evaluation revealed that of the available apps targeting NICU parents, less than half should be considered as acceptable educational material. Over two-thirds of the apps were found to have issues regarding credibility and just over a

quarter were considered good quality. The apps currently available for NICU parents are lacking and of concern in terms of quality and credibility.

(JMIR Mhealth Uhealth 2019;7(4):e11620) doi:10.2196/11620

KEYWORDS

parenting; intensive care units, neonatal; review; mobile health; mHealth; mobile apps; eHealth; education, nonprofessional; infant, premature

Introduction

Background

More than 1 in 10 babies are born preterm (ie, before 37 weeks gestation) worldwide and are often admitted to the neonatal intensive care unit (NICU) to help support their survival, growth, and development [1]. Parents of these infants requiring NICU care typically perceive this experience to be incredibly stressful, emphasized by the unfamiliar medical environment, appearance or behavior of their newborn, and interruptions in developing their role as a parent [2-7]. Unsurprisingly, parents strongly value being informed about their infant's condition by the NICU care team; however, their communication needs often go unmet. This is often because of the amount of information, either too much or too little information, and timing that health care providers deliver the said information [8]. As a result, parents strive to regain a sense of control by seeking additional information from external sources, including other health care providers, written educational materials, and the internet [8]. Parents often consult the internet, even before a health care provider, using the search engine Google or social media to search for advice related to their child's health and well-being [9-11]. More specifically, parents with infants in the NICU have also been found to prefer accessing the internet with their mobile phone when seeking information about their infant's health, with a trend in younger NICU parents preferring to use mobile apps [8,12,13].

Given the prevalence of mobile phone ownership worldwide, leveraging this medium to disseminate health information provides a great opportunity to better support overall health practices [14-16]. Not surprisingly, there has been a surge in the development of mobile health (mHealth) apps [17]. The use of mHealth apps by new parents has been increasing, often during the perinatal period, with mothers reporting an interest in or use of apps to monitor their health and their family's well-being [18-21]. Although mHealth apps have great potential, there are currently no formal quality standards that are required when developing these resources; thus, little is known about the quality of mHealth apps [14,22].

Broadly, concerns have been reported related to the quality of general health content on the internet [23] and considering this highly vulnerable population, it is imperative to evaluate the web-based health resources targeted to parents of infants in the NICU to ensure they are accessing current and evidence-informed information. To address this, we recently conducted a systematic review of websites available through Google that were targeted to NICU parents and found websites overall to be moderate to poor in terms of their reliability and quality of information [24]. A recent review evaluated parenting apps [25]; however, there was no emphasis on apps for parents with infants in the NICU. Building on our systematic review of Google, this review sought to evaluate the quality of current apps to provide further insight into how best to meet the informational needs of NICU parents.

Objectives

The primary objective of this review was to identify and evaluate apps available to parents of infants in the NICU for quality (ie, quality of health information and overall design), usability (ie, clarity and applicability of the health information), and credibility (ie, accuracy and reliability of the content). This study had the following research questions:

1. What is the quality of mobile apps targeted to parents of infants in the NICU?
2. What is the usability of the content within mobile apps targeted to parents of infants in the NICU?
3. Is the information provided by mobile apps targeted to parents of infants in the NICU credible?

In addition, a secondary objective was to explore common databases to determine if any peer-reviewed literature regarding the apps included for full review has been published and if so, what has been reported on them.

Methods

Study Design

Although there are some differences from traditional review methods related primarily to the search strategy, this study followed systematic review methodology, adhering to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards [26]. A protocol for this review was developed a priori and registered through PROSPERO International prospective register of systematic reviews [27]. This study is consistent with numerous recent reviews of mHealth apps using the same methodological approach [25,28-33].

Search Strategy

To identify the appropriate records for this review, the Canadian Apple App Store and Google Play were used as databases to search for mobile apps. In 2017, Apple and Android mobile phones accounted for 99.7% of the new market share [15]. Although not all Android phones access Google Play, it has been rated the leading app store offering approximately 1.6 million apps available for download, with Apple App Store in second place offering 1.5 million apps [34]. Thus, the search was limited to Apple App Store and Google Play, given the

substantial number of apps offered by these stores. This is consistent with other studies reviewing mobile apps [35].

To ensure a comprehensive inquiry, this review conducted a systematic search using a 2-step approach developed in collaboration with a librarian (KR). In step 1, we searched both app stores using 21 relevant key terms (eg, "parenting," "newborn," "preterm infant," "preemie," "premature baby," "neonatal intensive care unit," "NICU," and "neonatology"). This was followed by step 2 where each store was searched again using string keywords by inputting multiple forms of parent (eg, "parent," "caregiver," "guardian," "mother," and "father") in combination with the following terms: "neonatal intensive care unit," "NICU," "special care nursery," and "neonatal." This 2-step approach was utilized because of differing search methods between app stores at the time of this review, with the search algorithm for Apple optimized by using single keyword and the algorithm for Google Play with string keywords [35-37].

Selection Criteria

For feasibility, the search was limited to the top 100 apps identified from each search term applied, which is consistent with previous studies reviewing apps [38]. This limitation was applied as searches within the Apple App Store will continually refresh with additional apps of increasingly less relevancy. Apps that met inclusion criteria and were available in both stores were evaluated separately at the screening stage to assess if there were any substantial differences across operating systems. However, to limit redundancy during the final stage of this review, the most recently updated app as identified in the store description was kept for analysis. Snowball searching through recommended apps was conducted; however, no new apps were found that met the criteria that were not identified in the original search.

Inclusion Criteria

Although the term parent will be used throughout this review, this term will be considered inclusive of guardians, additional family or individuals that provide care to infants in the NICU. We chose against using the term caregiver as it is often considered more suggestive of a health care provider role within this context. Apps targeting parents related to the NICU experience were eligible for inclusion. To ensure a broad reach of available apps, there were no restrictions on the app's purpose. For example, apps could be for awareness, education, or tracking growth data, or reducing stress and enhancing coping. The following inclusion criteria was determined a priori: (1) apps targeted to parents of infants in the NICU; (2) available through Apple App Store or Google Play, accessible in Canada; (3) English language; (4) free or paid apps costing less than Can $20 (a consistent cut-off with similar studies as general users are unlikely to spend more than $10 per app [32]); (5) apps available in the following Apple App Store categories: Health & Fitness, Lifestyle, Medical, and Social Networking; and (6) apps available in the following Google Play categories: Communication, Education, Health & Fitness, Lifestyle, Medical, Parenting, and Social.

Exclusion Criteria

Apps were excluded if they were (1) general parenting apps, not related to the NICU experience; (2) intended for health care professionals; (3) classified as *e-books* by app store description or reviewers.

Screening Process and Data Extraction

After removal of duplicates, the title and store descriptions of all apps identified in the initial search were screened by 6 reviewers (BR, JD, KR, AO, JM, and KH) to determine eligibility for full review. For data extraction, 3 independent reviewers (BR, AO, and JM) were trained on how to use the measurement tools to ensure consistency when evaluating the apps. For full review, 1 reviewer was assigned to evaluate Android apps (AO), another for Apple apps (JM), and a third reviewer evaluated all apps (BR) using devices from both operating systems. Any disagreements were resolved by consensus or with a fourth reviewer (JD). Apps that were eligible for full review were downloaded and evaluated using 4 mobile phones, 2 Android (HTC One M7; LG G4) and 2 Apple (iPhone 5s; iPhone 6s). To address the primary objectives of this review, data were extracted using a structured data retrieval form compiled using the following 3 measurement tools:

Mobile App Rating Scale

The Mobile App Rating Scale (MARS) tool is designed to gather descriptive data and assess the quality of mHealth apps using an objective and reliable method [39]. This tool evaluates an average overall app quality through 4 core subscales: engagement, functionality, aesthetics, and information quality. The tool also offers 2 optional subscales to support the evaluation: app subjective quality and perceived impact of the app on user knowledge and behaviors. See Textbox 1 for the definition of each core subscale. Each item within the core subscales uses a 5-point scale (1=Inadequate, 2=Poor, 3=Acceptable, 4=Good, and 5=Excellent). Subscale scores can be isolated to determine strengths and limitations of the app under evaluation. Subscale mean scores are then summed and averaged to create an overall mean quality assessment. Mean scores are calculated for each subscale and scores classifications are consistent with item responses (ie, 1=Inadequate, 2=Poor, 3=Acceptable, 4=Good, and 5=Excellent). For this study, we report on each scale individually as well as report a total score.

Textbox 1. Definitions of Mobile App Rating Scale core subscales.

1. Engagement: "Fun, interesting, customizable, interactive, well-targeted to the audience"
2. Functionality: "App functioning, easy to learn, navigation, flow logic, and gestural design of app"
3. Aesthetics: "Graphic design, overall visual appeal, color scheme, and stylistic consistency"
4. Information Quality: "Contains high quality information from a credible source"

Textbox 2. Example of Patient Education Materials Assessment Tool for Audiovisual Materials items.

1. Understandability: "The material makes its purpose completely evident."
2. Actionability: "The material clearly identifies at least one action the user can take."

The Patient Education Materials Assessment Tool for Audiovisual Materials

This tool systematically evaluates content usability through the understandability and actionability of audiovisual patient education materials [40]. The Patient Education Materials Assessment Tool for Audiovisual Materials (PEMAT-AV) assesses if materials are understandable and key messages are clear to a diverse population with varying levels of health literacy. By evaluating actionability, the PEMAT-AV assesses if a diverse population can identify what they can or need to do based on the information provided in the education material. Each item has 3 options for scoring (0=Disagree, 1=Agree, and NA=Not Applicable), and overall scores for each subscale are summed and divided into percentage quartiles, with a potential range of 0% to 100%. See Textbox 2 for an example of items to assess understandability and actionability.

Trust It or Trash It

To support assessment of credible and unbiased resources, Trust It or Trash It provides guidance on how to critically evaluate the quality of health information provided in health resources [41]. This tool uses 6 questions to help determine the validity and reliability of the resources: Who wrote the information you are reading? Who provided the facts/Where did the facts come from? Who paid for it? When was it written or updated? How do you know this information pertains to you? Does the information seem reasonable based on what you've read or know? Each question has an associated description to explain how to evaluate appropriately to either a *trust* or *trash* the resource. Resources will be considered as *trash* if receiving this option for at least 3 questions from this tool. Trust It or Trash It is a relatively simple tool that will complement the remaining measurement tools and strengthen the rigor of this review.

The secondary objective, to identify the existence of peer-reviewed publications of included apps, was conducted by searching the app name included in the full review in PubMed and Google Scholar. This search was conducted in conjunction with data extraction in Fall 2017.

Data Analysis

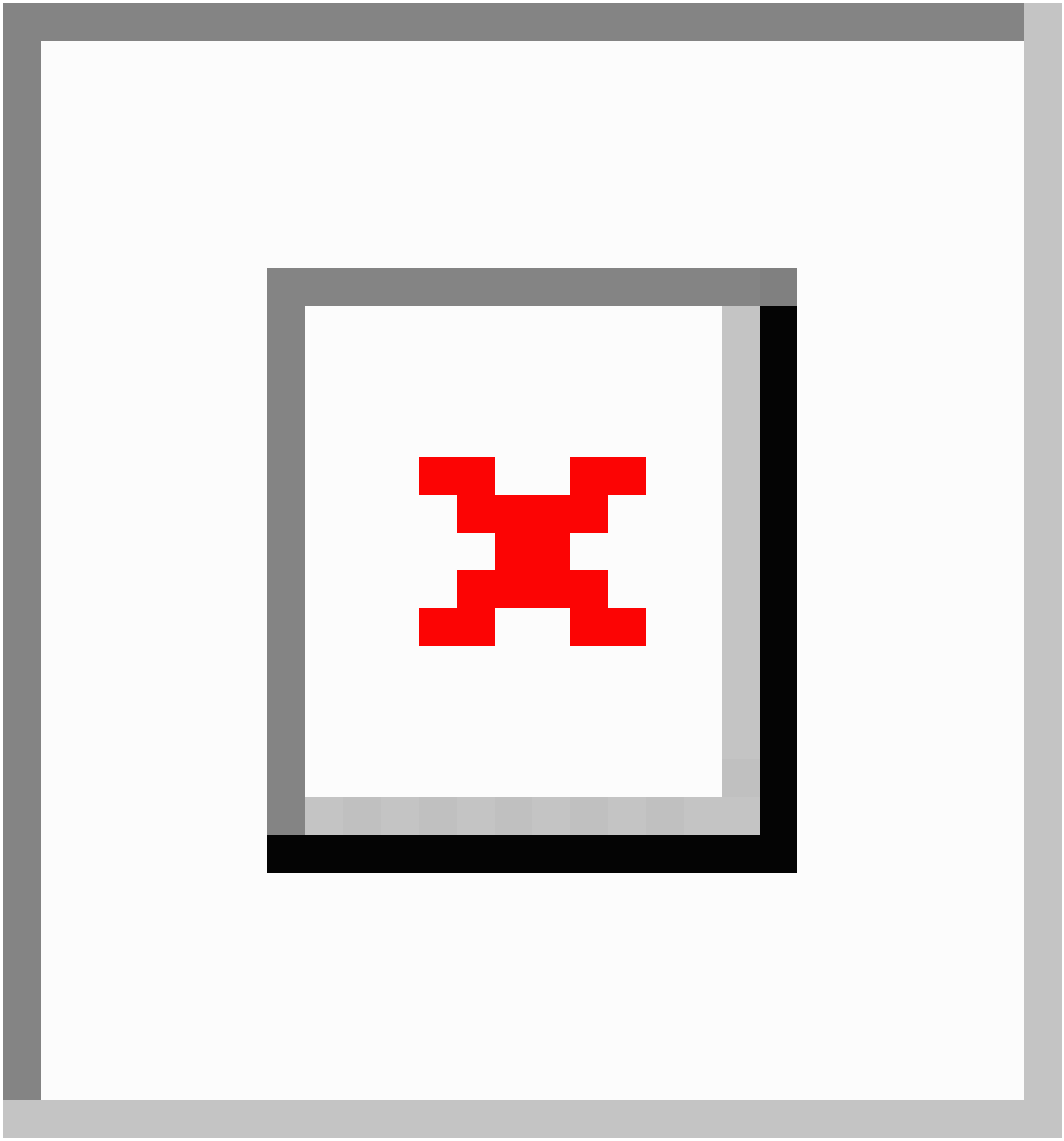
Data analysis was conducted using IBM SPSS 24.0. Descriptive statistics and frequencies were used to summarize the results of the evaluations from each measurement tool. As both the MARS tool and Trust It or Trash It tool evaluated quality, we conducted independent *t* test to compare MARS overall quality scores and subjective quality scores on whether an app was found to be trustworthy, as per the Trust It or Trash It tool. Additionally, a dependent *t* test was conducted to compare between MARS overall quality scores and subjective quality scores.

Results

Screening Process

After systematically searching both stores and manually inputting search findings into a Microsoft Excel spreadsheet over a month in 2017, our initial search yielded 6578 apps. After title and description screening, a total of 49 were assessed for eligibility. A total of 27 apps were included for full review and evaluated by 3 independent reviewers. Moreover, 9 apps were excluded as they were available on both stores; exclusion was based on oldest date as identified in the description, including 6 from Google Play and 3 from Apple App Store. There were 18 unique apps remaining that were included in the final analysis: 6 from Google Play and 12 from the Apple App Store. See Figure 1 for the PRISMA flow diagram outlining the screening process [26].

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of search process. NICU: neonatal intensive care unit.



Descriptive Characteristics of Apps

Of the 18 unique apps, 89% (16/18) had a last update year of 2015 or later, of which 61% (11/18) were updated in 2017. Although 94% (17/18) of apps were free to download, the 1 app that required payment to download was last updated in 2015. Moreover, 3 apps used commercial advertisements and 1 had in-app purchases, costing users Can $20.99 for a subscription to access more content. The costs of in-app purchases were not identified in the app store description. Although this app was included for full review, only content that was freely accessible was evaluated. Half of the apps were developed in the United States (n=9), with the remaining coming from the United Kingdom (n=3), Australia (n=1), New Zealand (n=1), Canada (n=1), Greece (n=1), or were of unknown origin (n=2). The majority of apps appeared to have been developed by a reputable source including nongovernmental organizations (n=12), government agencies (n=3), or university (n=1). App user rankings are represented by star ratings (5-point scale) in the Apple App Store and Google Play. None of the apps from Apple had an overall star rating, stating that they have not received enough reviews or ratings to rank. Moreover, 4 apps from Google Play had enough user ratings to provide an overall score, ranging from 3 to 5 stars with a median of 4 stars. The number of users who rated the apps ranged from 1 to 127. Table 1 provides a complete list of the included apps and their characteristics (for full table, see Multimedia Appendix 1). Using theoretical strategies predefined by the MARS tool, the

purpose of the apps was primarily for providing information or education (n=16), advice or tips (n=12), or monitoring or tracking data (n=10), as shown in Table 2. Although the apps covered a wide range of topics related to the NICU experience, those most commonly addressed were breastfeeding or feeding (n=18), growth and development (n=14), and illness or health issues (n=13). See Table 3 for a list of the topics identified in the apps.

Table 1. Description of apps.

App name, operating system	Developer	Country	Version	Cost (Can $)	Ads/in-app cost
Babble, Apple iOS	Midcentral District Health Board	New Zealand	1.5	Free	No
Baby Growth Tracker, Android	St. Rose Dominican Hospital	United States	1.1	Free	No
Connect2NICU[a], Apple iOS	Connect2 NICU	Unknown	1.1	Free	No
Gift of Life, Apple iOS	Mark Hoewing	United States	1	Free	Yes (Ads)
Integrated Family Delivered Care, Android	Propeller Apps	United Kingdom	1	Free	No
Life's Little Love, Apple iOS	AlexiaTek	Canada	1.0.8	Free	No
myChildren's, Android	Nationwide Children's Hospital	United States	4.2.1.1459-b6334bd	Free	No
MyPreemie App, Apple iOS	Graham's Foundation	United States	1.13	Free	No
My Neonatal Journal, Apple iOS	Rancon	United Kingdom	1.1.1	Free	No
NICU Companion, Apple iOS	Indiana University	United States	2	Free	No
NICU Parent, Apple iOS	PSD Apps	Greece	1.6	Free	Yes (Ads)
Our Journey in the NICU, Apple iOS	Phoenix Children's Hospital	United States	1	Free	No
Pebbles of Hope, Apple iOS	Pebbles of Hope	United States	2	Free	No
Peekaboo ICU Preemie, Android	Jozo Radman	United States	0.0.5	Free	No
Premature Baby Journal, Android	Life's Little Treasures Foundation	Australia	1.1	$2.96	No
Premature Birth, Android	Health Care Tips	Unknown	1	Free	Yes (Ads)
Quantum Caring for Parents (QCP)—NICU, Apple iOS	Caring Essentials	United States	1.1	Free	Yes (in-app purchases)
Tommy's—My Premature Baby, Apple iOS	Tommy's	United Kingdom	1.0.5	Free	No

[a]NICU: neonatal intensive care unit.

Table 2. Theoretical strategies classified by the Mobile Application Rating Scale (N=18).

Theoretical strategies	Apps, n (%)
Information or education	16 (88)
Advice or tips or strategies or skills training	12 (75)
Monitoring or tracking	10 (55)
Assessment	1 (5)
Goal setting	1 (5)
Peer support	1 (5)

Table 3. Full list of topics covered (N=18).

Topic	Statistics, n (%)
Breastfeeding or feeding	18 (100)
Growth and development	14 (77)
Illness or health issues	13 (72)
Overview or expectations	11 (61)
Skin-to-skin care or kangaroo care	11 (61)
Support	10 (55)
Physical health	10 (55)
Experience in the NICU[a]	10 (55)
Bringing baby home	7 (38)
Emotional needs of parents	7 (38)
Long-term outcomes	7 (38)
Complications or risks in the NICU	7 (38)
Parenting or bonding	5 (27)
Pain	2 (11)
Death or loss	2 (11)
Labor or birth	2 (11)
Depression	2 (11)
Preterm birth prevention	1 (5)
Relationships	1 (5)
Entertainment	1 (5)
Parent engagement in care	1 (5)

[a]NICU: neonatal intensive care unit.

Quality Assessment Using Mobile Application Rating Scale

Using the MARS tool, the average overall quality MARS score of the 18 apps ranged from 2.33 to 4.31, with an average of 3.37 (median 3.37). Less than half of apps (7/18, 39%) received an acceptable score (range: 3.26-3.72) on overall quality, with 28% (5/18) receiving a good score (range: 4.06-4.31), and no apps receiving an excellent score. Overall, apps scored low on engagement (1.00-4.6, mean=2.68) and moderate on functionality (2.75-5.00, mean=3.93), aesthetics (1.3-4.67, mean=3.21), and information quality (2.75-4.50, mean=3.65). See Table 4 for results across each core subscales and overall quality. Interestingly, the 5 apps to receive a good score in aesthetics were the only apps to receive a good score in overall quality. Within the engagement subscale, app's capabilities for customization were evaluated with 50% of apps having no (n=7) or insufficient options customization (n=2) and the other 50% of apps having basic (n=5) or numerous (n=4) options for customization. The subjective quality of apps varied, with only 11% (2/18) of apps receiving a good score. As part of the subjective quality assessment, evaluators determined 72% (n=13) of apps would likely be used 3 to 10 times (n=6), 1 to 2 times (n=4), or not used at all (n=3) in the next 12 months. Apps varied in their perceived impact on parental knowledge, attitudes, and behavior, with 22% (4/18) of apps receiving an acceptable score, 11% (2/18) receiving a good score, and 6% (1/18) receiving an excellent score. See Multimedia Appendix 1 for the individual app results across each subscale and overall app quality. Quantum Caring for Parents received the highest overall quality score (mean=4.31), followed by MyPreemie App (4.16), NICU Companion (4.10), Babble (4.08), and Integrated Family Delivered Care (IFDC; 4.06).

Patient Education Materials Assessment Tool for Audiovisual Materials Evaluation

In terms of usability reflected by health literacy, using the PEMAT-AV, 44% (8/18) of the apps received a score between 51% and 75% on understandability and 44% (8/18) were within 76% to 100% on actionability. One app that was solely for tracking and monitoring data was removed from this analysis as the app contained no educational content. See Table 5 for PEMAT-AV results by quartiles.

Table 4. Mobile App Rating Scale average scores.

Core subscales	Apps by average score, n				
	Inadequate	Poor	Acceptable	Good	Excellent
Engagement	2	7	8	1	0
Functionality	0	1	7	8	2
Aesthetics	2	3	8	5	0
Information quality	0	3	6	9	0
Overall app quality	0	6	7	5	0

Table 5. Patient education material assessment tool audiovisual scores.

Categories	Apps by Patient Education Materials Assessment Tool for Audiovisual Materials score[a], n			
	<25%	26%-50%	51%-75%	76%-100%
Understandability	1	4	8	4
Actionability	1	5	3	8

[a]Tracking or monitoring only apps were excluded from analysis (n=1).

Trust It or Trash It

The credibility of apps appears lacking, with 72% (13/18) of apps receiving a *trash* score and only 28% (5/18) deemed as trustworthy. Over 80% (15/18) of apps received a trash score on the following questions: “Who provided the facts or Where did the facts come from?” and “When was it written or updated?” However, with the question “Does the information seem reasonable based on what you’ve read or know?,” nearly all apps (n=16) were considered to match the information found in multiple sources based on the reviewer’s expert evaluation.

Overall Highest Ranking

The app to rank the highest across all measurement outcomes was IFDC. We conducted a subanalysis comparing quality measures between the MARS tool and Trust It or Trash It tool. There was no statistically significant difference in overall quality of apps reported in the MARS between the apps that were deemed trustworthy (mean 3.65, SD 0.75) compared with those that were not (mean 3.26, SD 0.57) in the Trust It or Trash It tool, t_{16}=−1.20, P=.25. Similarly, no statistically significant difference in the subjective quality of apps was found between apps that were trustworthy (mean 2.85, SD 1.22) compared with those that were not (mean 2.33, SD 0.89), t_{16}=−1.01, P=.33. When comparing within the MARS quality scores of apps, on average apps had a higher overall quality score (mean 3.37, SD 0.63) than the subjective quality score (mean 2.47, SD 0.98), t_{17}=−7.05, P<.001.

Secondary Objective

Only 1 app, MyPreemie App, had a relevant publication identified through a search in both PubMed and Google Scholar. The paper on MyPreemie App, published in 2013, was a description article that did not report any outcomes on feasibility, uptake, or impact on parent or newborn outcomes [42]. An abstract discussing MyChildren’s app was retrieved, but it provided no mention of the NICU-specific content [43].

Discussion

Principal Findings

Parents in the NICU are utilizing the internet and mobile apps easily accessible to them through their mobile phones to specifically search for more information about their infant’s health and well-being [12]. To the authors’ knowledge, this is the first and only review of apps targeting parents of infants in the NICU. Consistent with our previous work that found websites targeted to parents with infants in the NICU to be of moderate to low quality[24], this review found the apps currently available for the same population also lacking in quality and credibility. Overall, just over a quarter of apps were considered good quality, less than half acceptable in terms of educational materials, and over two-thirds as having issues related to credibility.

Over a third of the reviewed apps provided information on key topics regarding the NICU experience including breastfeeding or feeding, growth and development, and illness or health issues. However, information on topics that have been identified as prevalent concerns among NICU parents, such as attachment [44], infant pain [45], and death or loss [46,47], were found within only a small number of apps. Unsurprisingly, the topics covered and the proportion of apps related to these topics was comparable with our previous review of websites [24]. Although topics most commonly covered in the apps align with the increasing emphasis on family-centered or family-integrated care in the NICU, there are still gaps in addressing what parents identify as important. Many apps were updated in 2017, the same year when the review was conducted, and the “What’s New” statements in the app store descriptions included mainly technical aspects, such as fixing delays, as app version updates only occur when there are changes to the software code. Thus, it was difficult to determine if the content was updated within the apps, especially as most apps did not disclose their sources.

Quality

Similar findings have been identified from recent systematic reviews of various health-related apps [25,29,32,33] predominantly in terms of quality and credibility. Among other studies using the MARS tool, engagement scores of mHealth apps have been consistently low across reviews of mindfulness-based apps [32], cardiovascular disease symptom monitoring and self-management apps [29], and asthma self-management apps [28]. The focus of these reviews suggests that these apps should be inherently engaging to ensure target users believe they are valuable resources to improve their health and well-being. Additionally, our review found only half of apps had little to no options for customization limiting the user's ability to engage and personalize the app. Using the MARS subjective quality subscale, our expert reviewers evaluated that on average the majority of apps included in this review would likely be used only 0 to 10 times within the next year by NICU parents, which may not be sufficient or valuable for supporting parents throughout their NICU experience. However, this is in contrast to findings that reported almost one hundred percent of NICU parents use the internet or their mobile phones daily [12]. Again, consistent with our evaluation, these reviews found apps to have high scores for their functionality, yet moderate in terms of aesthetics, information quality, and overall quality [28,29,32] This trend suggests that future apps should strive to improve engagement, aesthetics, and information quality to strengthen the overall quality of apps. In addition, a recent review of apps targeting adults with chronic lung disease specifically reported that none of the apps included in their evaluation provided source identification [33]. Although important across all mHealth apps, it is of particular importance for parents with infants in the NICU to have timely access to credible, yet also engaging and understandable information as they often experience feelings of vulnerability and stress during this highly sensitive period [8,48,49].

Of the apps reviewed, most apps received at least an acceptable score in overall quality using the MARS tool. Although the Quantum Caring for Parents app had the highest overall quality MARS score, this app requires a paid subscription (ie, Can $20.99) to access all of the content, which may not be feasible for all parents. Although the other highest-ranking apps were free, it is important to note how the financial support model for these free apps (eg, grant funding) could impact quality over time if the developers are not able to sustain the app. Although outside of the scope of this review, it would be interesting to explore the relationship between app quality and financial support. Only 3 apps received a high score on perceived impact on parental knowledge, attitudes, and behavior change: IFDC, Quantum Caring for Parents, and Peekaboo ICU Preemie (see Multimedia Appendix 1). Interestingly, Peekaboo ICU Preemie was not ranked among the highest in overall quality, but based on expert evaluation, the information within this app was found to likely enhance parents learning and feelings to support their NICU experience, despite receiving lower scores in the MARS core subscales. Unsurprisingly, a statistically significant difference was found between overall quality scores and subjective quality scores, with subjective quality ranking lower on average. This finding is expected as in comparison with the overall quality criteria (ie, the 4 core subscales), the subjective quality items are particularly broad. For example, 4 questions ask for the reviewers' opinion on star rating, recommending the app, anticipated time using the app, and paying for the app. Those receiving a good score in the aesthetics subscale scored highest overall, suggesting these apps may have been developed with greater attention or to be in line with a fundamental theory of design in which *attractive things work better* [50]. The design of web-based resources, both websites or apps, should be a pleasurable experience for users and stimulate an overall positive response when interacting with the resource [51]. Design experts have outlined certain harsh conditions that produce a negative response often apparent in the NICU environment, including bright lights, loud abrupt noises, bodies that appear abnormal, etc [50]. Thus, it is important for these apps to use design techniques to enhance aesthetics as to not add to an already stressful experience in the NICU.

Usability and Credibility

The information within many of the apps was found to be understandable and actionable by the general population as they provided content that used lay language or visual cues to identify important points. Interestingly, subscores related to quality were the most inconsistent across tools. Our subanalyses showed that the Trust It or Trash It tool was not correlated with MARS overall or subjective quality scores. On the basis of expert review using the MARS tool, apps received acceptable to good scores for the information quality subscale with most scoring high on item 15 (ie, *is the app content correct, well written, and relevant to the goal/topic of the app*) [39]. In contrast, when the same apps were measured using the Trust It or Trash It tool, the credibility of most apps was found to be questionable. There are several possible reasons for this difference. One reason may be that The Trust it or Trash Tool consists of 6 items, specific to quality, whereas the MARS tool is a single item. Additionally, the MARS tool does not include important questions related to credibility, such as identification of sources. Given the limitations of the MARS tool, a supplementary assessment of credibility should be considered. However, it is important to emphasize that although it is crucial to use criteria to assess credibility, current mobile apps may not weigh the criteria as strongly during development; that is, app developers may use evidence-based information in the development of content, yet not provide end users with where they obtained that information. Thus, it is possible the content may not necessarily be untrustworthy but simply does not provide a source and therefore, received a negative score using the Trust It or Trash It tool. However, this lack of provision of sources to parents would not be captured when solely evaluating using the MARS tool, resulting in the disconnect between these 2 scales. Parents are often limited in their ability to assess the reliability of health content, and mHealth apps that contain educational information should be held responsible to disclose important aspects related to credibility including original sources, author names, and bibliographies. Although not a focus of this review, it is important to note that 1 limitation of this review relates to the limitations of the tools currently available to evaluate the strength of health-related apps. The development of novel tools appears to be warranted.

Secondary Objective

Relative to our secondary objective, item 19 of the MARS tool determines to what degree the app has been trialed and is considered to be evidence based. At the time of evaluation, only 1 app had a peer-reviewed publication associated with it (MyPreemie App); however, the identified study provided a description of the app's content and features with no indication of empirical testing [42]. There is a published abstract discussing MyChildren's app; however, there is no mention of the NICU specific content that was evaluated in this review [43]. Although identified after this review had been completed, the IFDC app was included in a recent publication regarding the developer's larger parent education project. However again, it appears the app was not tested; thus, although meeting our secondary objective, it would not impact the MARS score [39]. The low quantity of peer-reviewed literature is not unique to our review. Creber et al in 2016 found only 3 publications out of 34 identified apps related to cardiovascular disease management [29], whereas Owens et al's review (2018) of apps regarding chronic lung disease [33]and Tinschert et al on asthma-related apps found no publications [28]. The lack of existing peer-reviewed literature may put the apps identified in this review at a disadvantage as health care providers are less likely to recommend use of apps to patients and families if there is no evidence supporting their benefit. Thus, it is important for mHealth apps to provide an assessment of app quality through effectiveness studies and peer-reviewed publications [29].

Limitations

Despite following a rigorous systematic approach with expert reviewers, synthesizing and evaluating mobile apps is still a relatively new concept and thus, there are some limitations to address. We did not include reviewers from outside the country and because of this, only apps accessible in Canada were included. Manually gathering apps during the initial search significantly prolonged this stage of the review. However, this has been resolved for future reviews as 1 author (MS) has since developed a method to retrieve all app data relevant for the search stage. Reviewers with different devices and operating system versions experienced differences in functionality and app store description information; however, this was resolved through discussion and briefly reviewing the apps again with each device to come to a consensus. This could be avoided in future reviews by using devices that are strictly for research purposes (ie, not the reviewer's personal mobile phone). Due to the fact that we wanted to get an overview of all apps currently available to NICU parents, we not only included apps that were predominantly for monitoring and tracking data but also included other components such as a diary or a section to prompt questions to ask health care professionals. However, it was difficult to assess these apps using the PEMAT-AV as there was little health information within these apps. Moreover, we found that the PEMAT-AV may not be adequate to evaluate health literacy of mobile apps as the evaluation criteria have not yet been adapted to the unique variances within a mobile app platform. In addition, although the MARS tool was developed using rigorous methods, we would argue that most items in the core subscales elicit subjective responses, which could limit the objectivity and replicability of the measurement. The possibility for complete replicability of this review is limited because of the nature of mobile apps and various fluid components, including app ranking within stores, differences in devices, and recent updates of apps. Similarly, the findings of this review should be interpreted with some caution as the evaluation is based on the app version at the time of assessment, and thus, all future versions may differ in overall quality, usability, and credibility.

Conclusions

Despite the number of available apps for parents of infants in the NICU, this systematic review revealed that current mobile app resources vary in quality, usability, and credibility, with generally low scores. Additionally, peer-reviewed literature or empirical studies related to the identified apps are nearly nonexistent. Parents should be aware of the issues of quality and credibility identified in this review and be cautious when using an app for health information. This expert review was beneficial to provide a preliminary evaluation on the resources easily accessible by a parent's mobile phone. Building on this, usability testing or content analysis of the top 5 apps could be warranted to further explore how parents interact with the apps and provide a thorough evaluation examining the impact of these resources on parent learning needs, parent engagement, and neonatal outcomes. In addition, further attention to the development of high quality, credible resources targeted to NICU parents is needed.

Acknowledgments

The authors gratefully acknowledge the Mechanisms, Outcomes, Mobilization of maternally Led Interventions in Newborn Care research lab at the IWK Health Centre with their support for this publication through the ACOA Atlantic Innovation Fund, Project Number: 207518 "Chez NICU Home: Transforming health care delivery in neonatal care through greater parental presence and involvement."

BR receives support by the Dalhousie University School of Nursing Doctoral Scholarship and Helen Watson Memorial Scholarship, the Nova Scotia Health Research Foundation Scotia Scholars Award, BRIC NS Student Research Award, and Chez NICU Home Doctoral Scholarship. MC-Y is supported by a Canadian Institutes of Health Research New Investigator Award. AO is supported through an IWK Health Centre Summer Studentship Award. JD is funded through a Canadian Institute of Health Research Doctoral Award (FRN154341) as a PhD in Health trainee at Dalhousie University.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Complete version of Table 1. Description of apps and Table 5. Mobile App Rating Scale average scores per app.

[PDF File (Adobe PDF File), 57KB - mhealth_v7i4e11620_app1.pdf]

References

1. World Health Organization. 2018. Preterm birth URL: http://www.who.int/en/news-room/fact-sheets/detail/preterm-birth [WebCite Cache ID 6zxd97cLP]
2. Ray JG, Urquia ML, Berger H, Vermeulen MJ. Maternal and neonatal separation and mortality associated with concurrent admissions to intensive care units. CMAJ 2012 Dec 11;184(18):E956-E962 [FREE Full text] [doi: 10.1503/cmaj.121283] [Medline: 23091180]
3. Grosik C, Snyder D, Cleary GM, Breckenridge DM, Tidwell B. Identification of internal and external stressors in parents of newborns in intensive care. Perm J 2013;17(3):36-41 [FREE Full text] [doi: 10.7812/TPP/12-105] [Medline: 24355889]
4. Leigh B, Milgrom J. Risk factors for antenatal depression, postnatal depression and parenting stress. BMC Psychiatry 2008 Apr 16;8:24 [FREE Full text] [doi: 10.1186/1471-244X-8-24] [Medline: 18412979]
5. Carter BS, Carter A, Bennett S. Families' views upon experiencing change in the neonatal intensive care unit environment: from the 'baby barn' to the private room. J Perinatol 2008 Dec;28(12):827-829. [doi: 10.1038/jp.2008.102] [Medline: 18596717]
6. Obeidat HM, Bond EA, Callister LC. The parental experience of having an infant in the newborn intensive care unit. J Perinat Educ 2009;18(3):23-29 [FREE Full text] [doi: 10.1624/105812409X461199] [Medline: 20514124]
7. Wigert H, Johansson R, Berg M, Hellström AL. Mothers' experiences of having their newborn child in a neonatal intensive care unit. Scand J Caring Sci 2006 Mar;20(1):35-41. [doi: 10.1111/j.1471-6712.2006.00377.x] [Medline: 16489958]
8. De Rouck S, Leys M. Information needs of parents of children admitted to a neonatal intensive care unit: a review of the literature (1990-2008). Patient Educ Couns 2009 Aug;76(2):159-173. [doi: 10.1016/j.pec.2009.01.014] [Medline: 19321288]
9. Sundstrom B. Mothers "Google It Up:" Extending Communication Channel Behavior in Diffusion of Innovations Theory. Health Commun 2016;31(1):91-101. [doi: 10.1080/10410236.2014.936339] [Medline: 26075413]
10. Logsdon MC, Bennett G, Crutzen R, Martin L, Eckert D, Robertson A, et al. Preferred health resources and use of social media to obtain health and depression information by adolescent mothers. J Child Adolesc Psychiatr Nurs 2014 Nov;27(4):163-168. [doi: 10.1111/jcap.12083] [Medline: 25100157]
11. Logsdon MC, Mittelberg M, Myers J. Use of social media and internet to obtain health information by rural adolescent mothers. Appl Nurs Res 2015 Feb;28(1):55-56. [doi: 10.1016/j.apnr.2014.04.006] [Medline: 25015899]
12. Orr T, Campbell-Yeo M, Benoit B, Hewitt B, Stinson J, McGrath P. Smartphone and internet preferences of parents: information needs and desired involvement in infant care and pain management in the NICU. Adv Neonatal Care 2017 Apr;17(2):131-138. [doi: 10.1097/ANC.0000000000000349] [Medline: 27759567]
13. Gabbert TI, Metze B, Bührer C, Garten L. Use of social networking sites by parents of very low birth weight infants: experiences and the potential of a dedicated site. Eur J Pediatr 2013 Dec;172(12):1671-1677. [doi: 10.1007/s00431-013-2067-7] [Medline: 23913310]
14. Boulos MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014;5(3):229 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]
15. International Data Corporation. 2017. Smartphone OS market share, 2017 Q1 URL: https://www.idc.com/promo/smartphone-market-share/os [accessed 2018-06-06] [WebCite Cache ID 6zxfjmGTE]
16. Cisco. 2016. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017–2022 White Paper URL: https://tinyurl.com/y2lwempd [accessed 2018-06-06] [WebCite Cache ID 6zxg0KpFk]
17. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
18. Waring ME, Moore ST, Xiao RS, Lombardini LM, Allison JJ, Rosal MC, et al. Pregnant women's interest in a website or mobile application for healthy gestational weight gain. Sex Reprod Healthc 2014 Dec;5(4):182-184. [doi: 10.1016/j.srhc.2014.05.002] [Medline: 25433828]
19. Goetz M, Müller M, Matthies LM, Hansen J, Doster A, Szabo A, et al. Perceptions of patient engagement applications during pregnancy: a qualitative assessment of the patient's perspective. JMIR Mhealth Uhealth 2017 May 26;5(5):e73 [FREE Full text] [doi: 10.2196/mhealth.7040] [Medline: 28550005]
20. Lupton D, Pedersen S. An Australian survey of women's use of pregnancy and parenting apps. Women Birth 2016 Aug;29(4):368-375. [doi: 10.1016/j.wombi.2016.01.008] [Medline: 26874938]

21. Guerra-Reyes L, Christie VM, Prabhakar A, Harris AL, Siek KA. Postpartum health information seeking using mobile phones: experiences of low-income mothers. Matern Child Health J 2016 Dec;20(Suppl 1):13-21 [FREE Full text] [doi: 10.1007/s10995-016-2185-8] [Medline: 27639571]
22. Powell AC, Torous J, Chan S, Raynor GS, Shwarts E, Shanahan M, et al. Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. JMIR Mhealth Uhealth 2016;4(1):e15 [FREE Full text] [doi: 10.2196/mhealth.5176] [Medline: 26863986]
23. Eysenbach G, Powell J, Kuss O, Sa E. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. J Am Med Assoc 2002;287(20):2691-2700. [doi: 10.1001/jama.287.20.2691] [Medline: 12020305]
24. Dol J, Richardson B, Boates T, Campbell-Yeo M. Learning to parent from Google? Evaluation of available online health evidence for parents of preterm infants requiring neonatal intensive care. Health Informatics J 2018 Jan 1:1460458217752564. [doi: 10.1177/1460458217752564] [Medline: 29361867]
25. Davis DW, Logsdon MC, Vogt K, Rushton J, Myers J, Lauf A, et al. Parent education is changing: a review of smartphone apps. MCN Am J Matern Child Nurs 2017;42(5):248-256. [doi: 10.1097/NMC.0000000000000353] [Medline: 28639997]
26. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
27. Richardson B, Dol J, Rutledge K, Monaghan J, Orovec A, Howie K, et al. PROSPERO International prospective register of systematic reviews. 2017 Jul 26. Evaluation of smart-device applications targeted to parents of infants in the neonatal intensive care unit: a systematic review URL: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017072548 [accessed 2019-03-06] [WebCite Cache ID 76fodzTKt]
28. Tinschert P, Jakob R, Barata F, Kramer J, Kowatsch T. The potential of mobile apps for improving asthma self-management: a review of publicly available and well-adopted asthma apps. JMIR Mhealth Uhealth 2017 Aug 2;5(8):e113 [FREE Full text] [doi: 10.2196/mhealth.7177] [Medline: 28768606]
29. Masterson CR, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2016 Jun 14;4(2):e74. [doi: 10.2196/mhealth.5882] [Medline: 27302310]
30. 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 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]
31. Giunti G, Guisado FE, Dorronzoro ZE, Rivera RO. Supply and demand in mHealth apps for persons with multiple sclerosis: systematic search in app stores and scoping literature review. JMIR Mhealth Uhealth 2018 May 23;6(5):e10512 [FREE Full text] [doi: 10.2196/10512] [Medline: 29792295]
32. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
33. Owens OL, Beer JM, Reyes LI, Gallerani DG, Myhren-Bennett AR, McDonnell KK. Mindfulness-based symptom and stress management apps for adults with chronic lung disease: systematic search in app stores. JMIR Mhealth Uhealth 2018 May 15;6(5):e124 [FREE Full text] [doi: 10.2196/mhealth.9831] [Medline: 29764800]
34. Furlan A. Business of Apps. 2017. The Ultimate App Store List URL: http://www.businessofapps.com/guide/app-stores-list/ [accessed 2018-06-06] [WebCite Cache ID 6zxi1kEl3]
35. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: a systematic review of prevalent and innovative methods. Am J Prev Med 2016 Dec;51(6):1051-1059. [doi: 10.1016/j.amepre.2016.07.009] [Medline: 27659122]
36. Apple. Search Optimization - App Store - Apple Developer URL: https://developer.apple.com/app-store/search/ [accessed 2018-07-09] [WebCite Cache ID 70micnH7W]
37. Zolotareva K. TheTool. 2018. Google Play Store ASO Guide for: How to Optimize Android Apps URL: https://thetool.io/2017/how-to-optimize-google-play-aso-guide [accessed 2018-07-09] [WebCite Cache ID 70miSiYXm]
38. Knight E, Stuckey MI, Prapavessis H, Petrella RJ. Public health guidelines for physical activity: is there an app for that? A review of android and apple app stores. JMIR Mhealth Uhealth 2015;3(2):e43 [FREE Full text] [doi: 10.2196/mhealth.4003] [Medline: 25998158]
39. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
40. Shoemaker S, Wolf M, Brach C. Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. Patient Educ Couns 2014 Sep;96(3):395-403 [FREE Full text] [doi: 10.1016/j.pec.2014.05.027] [Medline: 24973195]
41. Trust It or Trash It?. URL: http://www.trustortrash.org/TrustorTrash.pdf [accessed 2018-06-06] [WebCite Cache ID 6zyOt5EDt]
42. Doron MW, Trenti-Paroli E, Linden DW. Supporting parents in the NICU: a new app from the US, 'MyPreemie'. J Neonatal Nurs 2013 Dec;19(6):303-307. [doi: 10.1016/j.jnn.2013.08.005]

43. Bourgeois FC, Mandl KD, Shaw D, Flemming D, Nigrin DJ. Mychildren's: integration of a personally controlled health record with a tethered patient portal for a pediatric and adolescent population. AMIA Annu Symp Proc 2009 Nov 14;2009:65-69 [FREE Full text] [Medline: 20351824]
44. Mäkelä H, Axelin A, Feeley N, Niela-Vilén H. Clinging to closeness: the parental view on developing a close bond with their infants in a NICU. Midwifery 2018 Jul;62:183-188. [doi: 10.1016/j.midw.2018.04.003] [Medline: 29684798]
45. Franck LS, Oulton K, Bruce E. Parental involvement in neonatal pain management: an empirical and conceptual update. J Nurs Scholarsh 2012 Mar;44(1):45-54. [doi: 10.1111/j.1547-5069.2011.01434.x] [Medline: 22339845]
46. Currie ER, Christian BJ, Hinds PS, Perna SJ, Robinson C, Day S, et al. Parent perspectives of neonatal intensive care at the end-of-life. J Pediatr Nurs 2016;31(5):478-489. [doi: 10.1016/j.pedn.2016.03.023] [Medline: 27261370]
47. Branchett K, Stretton J. Neonatal palliative and end of life care: what parents want from professionals. J Neonatal Nurs 2012 Apr;18(2):40-44. [doi: 10.1016/j.jnn.2012.01.009]
48. De Rouck S, Leys M. Illness trajectory and Internet as a health information and communication channel used by parents of infants admitted to a neonatal intensive care unit. J Adv Nurs 2013 Jul;69(7):1489-1499. [doi: 10.1111/jan.12007] [Medline: 22989266]
49. Ottosson C, Lantz B. Parental participation in neonatal care. J Neonatal Nurs 2017 Jun;23(3):112-118. [doi: 10.1016/j.jnn.2016.11.001]
50. Norman D. Emotional Design. New York: Basic Books; 2004.
51. Walter A. Designing for Emotion. New York: A Book Apart; 2011.

Abbreviations

IFDC: Integrated Family Delivered Care
MARS: Mobile Application Rating Scale
mHealth: mobile health
NICU: neonatal intensive care unit
PEMAT-AV: Patient Education Materials Assessment Tool for Audiovisual Materials
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach; submitted 31.07.18; peer-reviewed by A Powell, E Dorronzoro, M Stuckey; comments to author 21.09.18; revised version received 10.10.18; accepted 29.10.18; published 15.04.19.

Please cite as:
Richardson B, Dol J, Rutledge K, Monaghan J, Orovec A, Howie K, Boates T, Smit M, Campbell-Yeo M
Evaluation of Mobile Apps Targeted to Parents of Infants in the Neonatal Intensive Care Unit: Systematic App Review
JMIR Mhealth Uhealth 2019;7(4):e11620
URL: http://mhealth.jmir.org/2019/4/e11620/
doi:10.2196/11620
PMID:30985282

Original Paper

Smartphone Apps Targeting Alcohol and Illicit Substance Use: Systematic Search in in Commercial App Stores and Critical Content Analysis

Babak Tofighi[1], MD, MSc; Chemi Chemi[1], MPH; Jose Ruiz-Valcarcel[2], MPH, MSGIS; Paul Hein[1], MD; Lu Hu[1], PhD

[1]Department of Population Health, New York University School of Medicine, New York, NY, United States

[2]University of Puerto Rico School of Medicine, San Juan, Puerto Rico

Corresponding Author:
Babak Tofighi, MD, MSc
Department of Population Health
New York University School of Medicine
227 East 30th Street, Room 713
New York, NY, 10016
United States
Phone: 1 6468017345
Email: babak.tofighi@nyumc.org

Abstract

Background: Smartphone apps promise to enhance the reach of evidence-based interventions (cognitive behavior therapy, contingency management and therapeutic education system) for populations with substance use disorders, with minimal disruption to health systems. However, further studies are needed to systematically evaluate smartphone apps targeting alcohol and illicit substances.

Objective: The aim of this study was to evaluate the functionality, aesthetics, and quality of information of free or low-cost apps claiming to target alcohol, benzodiazepine, cocaine, crack/cocaine, crystal methamphetamine, and heroin use using the validated Mobile App Rating Scale (MARS) and critical content analysis.

Methods: A systematic search of iTunes and Google Play app stores for free or low-cost apps facilitating recovery was conducted in March 2018 and yielded 904 apps using the keywords described in previous studies (eg, recovery, sobriety, sober, alcohol, and heroin). An interdisciplinary team of clinicians, behavioral informatics, and public health reviewers trained in substance use disorders conducted a descriptive analysis of 74 apps categorized as reducing use. In addition to the MARS scale, a descriptive analysis of relevant apps was conducted by the study team to assess for quality indicators emphasized by expert guidelines and review articles.

Results: Most apps (n=74) claimed to reduce use or promote abstinence and yielded an overall low median MARS score of 2.82 (0.55) and a wide range of scores (1.64, 4.20). Ratings were also low for engagement (2.75 (0.72)), functionality (3.64 (0.78)), aesthetics (3.03 (0.87)), information (2.82 (0.62)), and satisfaction (1.76 (0.67)) subdomains. Innovative design and content features elicited in the review included initial assessments of substance use following app download, tracking substance use, and related consequences (eg, cost or calorie intake), remote and proximate peer support per geospatial positioning, and allowing users and family members of individuals with substance use disorders to locate 12-step meetings, treatment programs, and mental health services. Few apps integrated evidence-based psychotherapeutic (eg, cognitive behavioral therapy [CBT] or motivational interviewing) and pharmacologic interventions (eg, naloxone or buprenorphine).

Conclusions: Few commercially available apps yielded in our search integrated evidence-based interventions (eg, extended-release naltrexone, buprenorphine, naloxone, Self-Management and Recovery Training recovery, or CBT), and a concerning number of apps promoted harmful drinking and illicit substance use.

(JMIR Mhealth Uhealth 2019;7(4):e11831) doi:10.2196/11831

KEYWORDS

mHealth; substance use disorder; mobile health; alcohol abuse

Introduction

Mobile phone–based health (mHealth) interventions offer a ubiquitous and low-cost approach to improving health outcomes. Smartphone apps, short message service (SMS) text messaging, and interactive voice response are effective approaches to reducing the burden of substance use disorders (SUDs) [1-4]. Most Americans now own smartphones (77%), and smartphones are especially popular among younger adults aged 18 to 29 years (92%) [5]. Smartphone apps promise to enhance the reach of evidence-based interventions (cognitive behavior therapy, contingency management, and therapeutic education system) for populations with SUDs with minimal disruption to health systems [2,6-8]. Reports have estimated the availability of over 318,000 mHealth apps in 2017 and the use of health-related apps by approximately half of all smartphone users in 2018 [9]. Considering existing barriers to formal treatment for SUDs, including perceived stigma, cost, and limited treatment slots [10], smartphone apps are increasingly utilized by individuals excluded from care to reduce alcohol and illicit substance use [11-13].

Concerns regarding the quality, efficacy, and privacy of mHealth apps persist. App descriptions routinely include unsubstantiated claims of medical expertise and intervention efficacy, while failing to disclose the sale of health information and personal data gathered from the user to third party vendors for commercial use [14]. Studies assessing mHealth apps targeting SUDs in smartphone app stores (ie, Google Play and/or iTunes) mostly described commercially driven apps that failed to offer evidence-based psychosocial interventions or to link users to addiction treatment providers [11,12,15]. In a descriptive analysis of apps addressing alcohol use, Weaver et al also found that apps claiming to inform users of their possible blood alcohol concentration actually promoted risky drinking behavior via games and other entertaining features [12]. Searches also yielded many recreational apps that promoted drug cultivation, trafficking, and simulated use [16].

Despite these findings, further studies are needed to systematically evaluate smartphone apps targeting alcohol and illicit substances. Studies assessing app content are generally descriptive and limited to one-dimensional outcome quality measures [17]. Previous searches of apps targeting SUDs were often limited to a single app provider (iTunes or Google Play), did not use keywords described by users to search for apps targeting SUDs (ie, *recovery*, *sobriety*, *abstinence*, and *detox*), and did not incorporate validated methods of assessing smartphone apps [13,14,18]. In addition, in the last 3 years, as most of these searches were conducted, the number of mHealth apps has doubled from 165,000 apps in 2015 to 325,000 mHealth apps in 2017 [19].

Our study utilized the validated Mobile App Rating Scale (MARS)[18] and critical content analysis [14], which offered a standardized approach to evaluate the functionality, aesthetics, and quality of information of apps claiming to target alcohol, benzodiazepine, cocaine, crack/cocaine, crystal methamphetamine, and heroin use. The MARS is the first mHealth app–quality indicator of engagement, functionality, aesthetics, and informational content delivered by a multidisciplinary team of clinicians and technology experts [18]. Our research also utilized a critical content analysis of apps listed in the smartphone stores claiming to reduce substance use to assess actual clinical impact (ie, review of the literature and app developer website) and linkage to patient-centered care models for addiction treatment (ie, self-efficacy, education, linkage with addiction specialty care, primary care, self-help groups, and/or individual counseling) prioritized by public health experts [20,21].

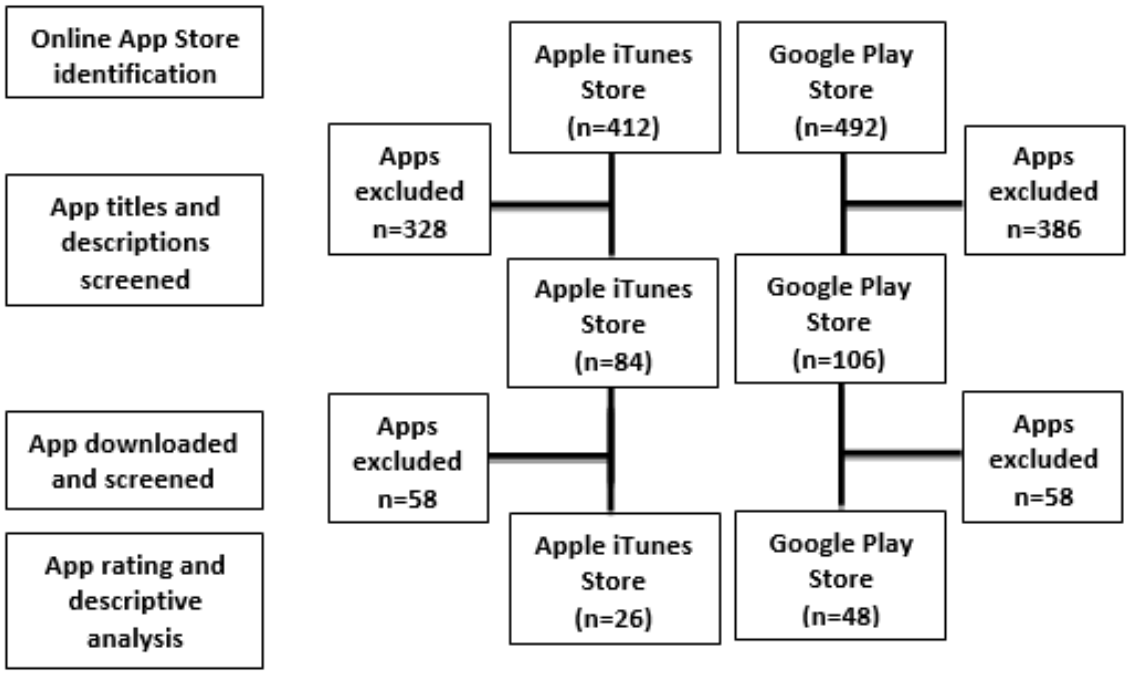
Methods

Smartphone App Selection

In March 2018, a systematic search of smartphone apps facilitating recovery from alcohol and illicit substances was conducted on the iTunes App Store and Google Play (see Figure 1) because of their popularity in the United States for app users. Additional apps targeting alcohol and illicit substance use described in the literature (eg, PubMed, Google Scholar, and PsycInfo) were also included in the search. The keyword search was in English, and apps that met the inclusion criteria were downloaded on an American mobile service network on the iPhone 6, iPhone 7, iPhone 7 plus, Samsung Galaxy J7, and Android LG G6 models.

The search terms were based on relevant terms described in the literature and in previous surveys among participants who utilized smartphone apps to reduce substance use (eg, *sober, sobriety, recovery, crystal methamphetamine*, *opioid*, *alcohol*, *cocaine, crack cocaine, and benzodiazepine*) [22]. The initial selection of apps excluded apps found to be irrelevant per the app title and online app store description (eg, music/relaxation, games, clocks, and religiosity), apps not in English, apps that lacked accessibility or functionality, apps designed for health care professionals, apps that cost more than $1 US dollars (on the grounds that they were unlikely to be purchased by a large number of users), and harmful apps promoting substance use. Although 12-step groups are an essential approach in reducing the burden of SUDs across diverse populations with SUDs [23], apps based solely on the 12-step model were not reviewed because of the large quantity of apps and limited study resources and their exclusion of content pertaining to medication-assisted treatments and effective behavior change models outlined by federal and expert guidelines [24]. The total number of apps yielded from our search was similar to previous studies on recovery apps [11,12,15]. Owing to the large number of apps retrieved in our initial search, apps without any user star ratings or reviews were also excluded in the preliminary screening. Apps that met preliminary inclusion criteria were then downloaded to the coauthors' smartphones and assessed for accessibility, functionality, and relevancy of app content to reducing substance use (see Figure 1).

Figure 1. Flow diagram for smartphone app selection.



App Assessment

App quality was assessed using the MARS [18]. The MARS domains (ie, engagement, functionality, aesthetics, information quality, and subjective quality) are constituted by 23 items based on a 5-point scale (1-Inadequate, 2-Poor, 3-Acceptable, 4-Good, and 5-Excellent). The classification domain is a descriptive survey of app price, platform, rating, and technical features (eg, password protection and log-in protocols). The excellent internal consistency (alpha=.92) and interrater reliability (intraclass correlation, ICC=.85) of the MARS scale make it ideal for conducting initial assessments of emerging apps [18]. Apps are then scored by calculating the mean scores of each respective subscale and the total mean score. The MARS scale has been successfully utilized to assess apps targeting a range of health conditions, including asthma, heart failure, and cancer [25-28]. However, the MARS scale has not been used to assess the quality of smartphone apps targeting alcohol and illicit substance use.

To ensure shared understanding of review criteria and MARS subscales, the study team reviewed the MARS literature and discussed each domain and subscale. The interdisciplinary study team included clinicians with experience in SUDs (BT and PH) and mHealth design (BT, PH, JRV, and LH). The reviewers convened and used 3 separate meetings to pilot apps and compare ratings per the MARS scale, assess the quality of their ratings, and resolve any discrepancies and ambiguities in the scale items. Each app was tested for at least 15 min. The raters then independently assessed 5 apps, and these apps were used to test the MARS tool, compare our judgments, and resolve any discrepancies. Interrater reliability among iTunes app raters was high (ICC=.80), but the overall app MARS score had fair interrater reliability (ICC=.58) requiring additional individual and group meetings to address discrepancies that emerged during the initial review. The study team compared ratings using iTunes for pragmatic reasons as most of the raters were Apple smartphone owners.

In the final round, we reviewed a total of 74 apps (26 Apple and 48 Google Play apps). Apps that met the inclusion criteria were assessed individually by the study authors on their smartphones in April 2018. In addition to the MARS scale, a descriptive analysis of relevant apps was conducted by the study team to assess for quality indicators emphasized by expert guidelines [8,29], as well as review articles [4,17,30], as evidence-based design practices in mHealth interventions targeting SUDs. This descriptive analysis elicited critical findings in app design and delivery features including (1) information or access to medication-assisted treatments via primary care and specialty addiction treatment programs (ie, residential treatment and intensive outpatient programs); (2) risk-reduction content (eg, safe sex practices, syringe exchange programs, naloxone, HIV, and hepatitis C virus [HCV] prevention education); (3) integration of behavior-change content in the intervention design or content (eg, cognitive behavioral therapy [CBT], motivational enhancement therapy, and contingency management); (4) empirical evidence demonstrating smartphone app efficacy by conducting searches in online databases (eg, PubMed and Google Scholar); and (5) privacy measures (eg, password protection, email or text verification, and user de-identification). Negative features of the reviewed apps assessed for the use of disruptive or distractive ads, religious texts, content exacerbating substance use (ie, access to discount liquor and bars with drink specials), and unverified claims of professional and clinical legitimacy.

Finally, we identified apps that were aligned with the National Institute on Drug Abuse's strategic objective of enhancing chronic disease management and personalized treatment that *matches an addicted person's changing needs* over time based on the Medical Management outline (ie, patient-provider communication, medication adherence, self-management, goal of opioid abstinence, and counseling participation) to improve linkage of office-based opioid treatment with effective pharmacotherapies (ie, buprenorphine-naloxone and extended-release naltrexone) [31-34]. During the final round of reviews of apps that met the eligibility criteria, the primary author reviewed 30% of the apps individually assessed by the secondary authors to ensure consistency with the MARS subscales and descriptive analysis guidelines.

Results

Summary of Findings

The Google and iTunes searches identified a total of 904 apps of which 412 were on iTunes and 492 from Google Play (see Figure 1). The initial review excluded apps that were irrelevant (n=322), required payment for use (n=184), were linked with harmful use (eg, gamification of binge drinking or simulating illicit substance use; n=118), were duplicates (n=70), and were not in English (n=20).

Our secondary review then assessed a total 84 Apple apps and 106 Google apps that met the inclusion criteria. Apps were assessed individually by the study team, and 116 apps were excluded owing to the following factors: (1) unavailability in the app stores 3 weeks between the initial and secondary reviews (n=32); (2) duplicates (n=18); (3) educational content for health care providers (n=16); (4) required payments for full use of the app (n=9); (5) focused exclusively on the 12-step approach (n=9); (6) irrelevant to recovery (n=8); (7) not functioning (n=5); (8) required invitation from a treatment program (n=4); (9) required a device purchase (eg, breathalyzer; n=4); (10) only displayed a clock without any recovery content (n=4); (11) were not functional (n=3); (12) only offered religious content (n=2); (13) not in English (n=1); and (14) exacerbated harmful use (n=1).

Of the remaining apps, 74 met the inclusion criteria and underwent further analysis.

Most apps targeted alcohol use (n=40). Fewer apps emerged from this review exclusively addressing opioid use (n=6) and none focused on cocaine, crack/cocaine, or methamphetamine use. The overall median score of apps included in this study was poor (2.82 (0.55)) based on the 5-point MARS scale and demonstrated a wide range of ratings (1.64, 4.20; see Table 1). Scores for engagement (2.75 (0.72)), functionality (3.64 (0.78)), aesthetics (3.03 (0.87)), information (2.82 (0.62)), and satisfaction (1.76 (0.67)). Apps that had the highest average MARS score included SoberWorx (4.20), Recovery Today Magazine (3.77), Sober Grid (3.75), and Addicaid: Addiction Recovery and Support (3.64). The lowest-ranking apps included Sick Not Stupid (1.64), Sober Day Recovery App (1.71), and Stop Drinking Alcohol Now (1.75).

Innovative Features of High-Quality Apps Targeting Substance Use Disorders

Use of the MARS scale and descriptive analysis allowed the study team to evaluate higher-rated apps (see Table 1) for design features that may be attributed to increased user engagement and potential clinical impact. Innovative design and content features elicited in the review included initial assessments, tracking substance use and related consequences (eg, cost and calorie intake), remote and proximate peer support per geospatial positioning, and allowing users and family members of individuals with SUDs to locate 12-step group meetings, treatment programs, and mental health services. Apps commonly elicited initial assessments of substance use patterns among users following app download. However, the only 2 apps that utilized an evidence-based approach to initial assessment were The Saying When and Alcohol Tracker apps.

The *Saying When* app (3.33) was developed by The Canadian Centre for Addiction and Mental Health to facilitate abstinence or reductions in the drinking quantity. The app features include an initial baseline assessment of drinking patterns, setting personalized goals, tracking drinks and urges, offering *tips for success*, and linkage to community treatment services. The self-help approach of the *Saying When* app was adapted from a manualized version based on *cognitive and behavioral strategies* that were not elaborated in the app store description and accompanying Web page. No scientific publications related to the app were available. However, app components appeared to be a cross-over of interventions described by the authors in earlier studies as having clinical impact in reducing drinking, including (1) exposure to initial assessments on drinking patterns [35] and (2) receipt of self-help books outlining initial assessments of drinking combined with the manualized version [36].

Alcohol Tracker (2.97) utilized the Alcohol Use Disorders Identification Test and Functional Analysis of Addictive Behaviors to notify users if they have surpassed alcohol consumption based on the National Institute for Health and Care Excellence UK Guidelines and link users to treatment resources. However, only users based in Singapore can access hotlines. In addition, reliance on self-reported drinking by apps tracking use is at risk of recall bias, particularly during binge-drinking episodes.

Table 1. Mobile app rating scale results. App availability is subject to removal per app developers, Google Play, and the Apple iTunes Store. Apps included in this table were rated as >3 or higher.

Smartphone app name	Engagement	Functionality	Aesthetics	Information	Satisfaction	Overall Score
SoberWorx	4.4	4.5	5	3.83	3.25	4.2
Recovery Today Magazine	3.8	4.5	4.3	3.5	2.75	3.77
Sober Grid	3.8	3.25	4.3	3.67	3.75	3.75
Addicaid: Addiction Recovery Support	3.8	4	4.67	3.5	2.25	3.64
BoozeFit	3.2	4.5	4	3.5	2	3.44
CleanTime Counter	2.2	4	4.25	3.5	3	3.39
Wise Drinking	3.2	3.5	4.667	2.83	2.75	3.39
SoberApp-Alcohol Calculator	3.4	4	3.33	3.67	2.5	3.38
Alcohol Check - BAC Calculator	3.4	3.5	4.3	3.4	2.25	3.37
FlexDek: Anglestrong Edition	3.8	4	3.3	3.16	2.5	3.35
Saying When	3.4	3.5	4	3.5	2.25	3.33
OARS Experience	2.8	3.75	2.5	4	3	3.21
Drinks Meter	2.75	4.25	2.67	3.5	2.75	3.18
Drive Sober	3.8	4.75	2.3	3.3	1.75	3.18
Best Alcohol Test	3.6	3.75	3.5	3	2	3.17
Stop OD NYC	2	4.25	3.3	3.28	3	3.17
Drug Addiction Recovery	2.6	3.75	4.3	3.6	1.5	3.15
Blood Alcohol Content Calculator+Timer	2.6	4.75	3.3	3	2	3.13
Alcohol Tracker	2.6	4.25	3.6	3.4	1.75	3.12
My Drink Control	3	4.25	3	3.6	1.75	3.12
AlcDroid Alcohol Tester	3	4.25	2.67	3.16	2.5	3.12
BACTrack	3.4	4.5	3.66	2.5	1.5	3.11
Sober Grid	3.75	3.3	3	3	2.5	3.11
Clean & Sober Time	3	3.5	4	3	2	3.1
Alcohol Calorie Counter	2.6	3.75	4.67	2.75	1.5	3.05
Intoxication Calculator	3.4	3.4	3.3	3.4	1.75	3.05
Addiction Quotes	3.2	4	3.3	3	1.75	3.05
Habit Tracker	2.6	4	3.3	3	2.25	3.03
Wbi.today	3.6	3.25	3.3	3.16	1.75	3.01

Several apps claimed to provide users access to sober peers via intra-app messaging, *help* icons, or forums. The most intriguing app (MARS score 3.75) facilitating peer support was *Sober Grid* and was developed by a team of academics and developers to offer a *global newsfeed* of shared posts on experiences in and insights into recovery and an instant help feature to link users to available peers online and in-person. The app then encourages patients to refer actively using peers to treatment. The app also encourages adoption among clinicians and health systems by offering an administrator dashboard, the option to launch mass notifications and messages to patients, and onboarding support. In addition, health systems may use the app to track substance use among patients during and post treatment and allow patients to meet and provide online support.

Addicaid Recovery Support (3.64) also offers linkage to peers in recovery and self-help group meetings. Some of the topics include *mothers in recovery, adult children of alcoholics, narcotics, methamphetamine, and heroin*. However, once a user joins the group, the posts are all from over a year ago. There are also a series of sessions such as *starting recovering*, *commitment and community*, *introducing a new routine*, and *recovery maintenance* and under each of these goals are comments including positive words of encouragement from peers. Once enabled, the app also provides users with a list of groups and meetings nearby.

Pocket Rehab facilitates text, telephone call, and video conferencing calls with other peers in recovery. Users are able to enter their location to search for nearby 12-step group meetings, a photo motivating sobriety, and an anticipated quit

date. However, more interestingly, the app offers a *Community chat* forum with features similar to a Facebook wall and an immediate helpline linking individuals to other sober peers who are using the app and willing to communicate with the user to provide support. Users may select individual or community support and communicate via SMS text messaging, voice, or video. If this contact was not helpful, the app then links users to *chat with an experienced peer* or *write a journal entry*. Although the app promises to link individuals to other users within 1 to 2 min, our study team was unable to establish contact with any app user. The app requested access to users' locations and assured that their information would not be available to third-party vendors. The study team gave a lower rating to the app (MARS score 2.66) owing to the lack of peer contact, linkage to behavioral health specialists in addition to peers in recovery, and integration of any evidence based-content.

Additional apps utilized fellow users to provide support. SoberWorx received a rating of 4.20 and was established by individuals in recovery to provide peer support, link to treatment resources, and also allow family members to locate treatment resources for loved ones with SUDs. Although the app claims to offer SMS text messaging–based contact with treatment programs via the app platform, we were unable to communicate with any providers or program staff. In addition to linking users to *treatment centers, addiction counseling, and sober living homes*, the app also allows for peer support, posting testimonials, and access to educational recovery content via YouTube videos. However, the peer support feature was also not interactive and was also difficult to identify online users and initiate contact.

The only app that clearly demonstrated application of an evidence-based psychotherapeutic approach was the Self-Management and Recovery Training (SMART)Recovery Cost Benefit Analysis (MARS rating 2.85). This app is based on the SMART recovery model, which integrates CBT to offer coping strategies for individuals in recovery to reduce the risk of relapse. The app allows users to enter the costs of ongoing use and the financial benefits of abstinence. However, the poor design (2) and lack of satisfaction (2) by the study team following app use resulted in a lower overall score. Despite the increasing popularity and clinical benefit of engaging with SMART recovery groups, counselors, and educational content based on the SMART platform [37], the app fails to effectively translate these resources into accessible and user-tailored features.

Government-Sponsored Smartphone Apps

Our search yielded few government-initiated apps and apps that offered risk reduction measures for individuals with SUDs. The most intriguing app was STOP OD NYC (3.17), developed by the New York City Department of Health and Mental Hygiene, which provides detailed information on opioids (eg, heroin and fentanyl) and instructions on naloxone administration in the event of an overdose. The app provides several risk-reduction content, including (1) a *find naloxone* option that links users to mapped pharmacies, harm reduction programs, and health care centers providing free naloxone; (2) naloxone administration instructions for intramuscular, intranasal, and auto-injector formulations of naloxone; (3) recognizing individuals suspected of an overdose; and (4) information on legal protection for individuals administering naloxone. The app utilizes SMS text messaging, cartoon, and YouTube-based videos to offer users multimedia educational instructions. Finally, users can click on the *NYC Health* icon to access other health resources within the Department of Health and Mental Hygiene platform (eg, cardiovascular health, reducing glucose intake, and smoking cessation).

My Drink Control (3.12) is an app and Web-based tool developed by the Public Health Department in Zurich, Switzerland, to facilitate tracking of alcohol use, financial costs of use, and calories gained with each drink. Although the app's *psychoeducational* content was based on CBT, our review of the app failed to elicit any theory-based content or design features. The app's overall functionality was limited to tracking drinks, reminders, and linkage to treatment and counseling services for alcohol use. However, compared with other government-sponsored apps, *My Drink Control* had a more appealing and user-friendly design.

Furthermore, 1 app designed for the Substance Abuse and Mental Health Services Administration (SAMHSA) app challenge, *FlexDek MAT* (2.90), described itself as linking participants to the methadone maintenance program but again failed to offer accurate and updated contact information regarding the Office based opioid treatment (OBOT) programs after this feature was utilized. Information pertaining to MAT, including *buprenorphine*, *methadone*, *MAT*, *naltrexone*, and *after naltrexone*, was limited to PDF files, links to SAMHSA's website, and an external website [38]. Other links to the 12-step and SMART recovery groups were not functioning after linking to an external website. The *forum* icon opened an error page and the *Coaches* option linked users to only 5 recovery coaches across the nation with a nonfunctioning link icon. The *rewards* option for using the app claimed to offer free hours of recovery coaching, but this was not evident in our review. In addition to linking to nonfunctioning pages, the app also exposed users to irrelevant ads in the lower segment of the screen.

Features of Low-Quality Apps Targeting Substance Use Disorders

Lower quality apps often claimed to support recovery through complex design features but were limited to 1 to 2 basic functions such as supportive quotes, timers, blood alcohol calculators, or logging daily substance use without providing tailored feedback and evidence-based interventions. Some apps only copied quotes from religious texts without specifying content in the title or description (eg, OARS experience). Furthermore, 1 app entitled *Alcoholism Treatment* would only play ambient electronic music that app developers claimed would stimulate desires to quit.

Many of these apps used deceptive descriptions of complex recovery resources but would inundate users with pop-up advertisements (eg, Stop Drinking Alcohol Now and Sobriety Clock), require users to log-in via Facebook or Google with access to their social networks, and request access to user location. Among blood alcohol calculators, most lacked useful information on evidence-based treatment approaches and were often rated poorly by our study team and app store users. In

addition, 1 app, BACtrack, developed by the San Francisco–based BACtrack, enables users to measure blood alcohol levels. However, to utilize the app, users were required to purchase a breathalyzer ($99.00 US dollars) to fully utilize the platform. Users were instructed to breathe into the breathalyzer and promised to log results wirelessly to their mobile device. The only free and functioning feature on the app was a link to Uber to access a ride. Additional apps also claimed to have initiated supportive peer networks, but forums and peer-messaging functions would typically be inactive. Others would link users to 12-step group meeting schedules or 12-step–based online forums rather than the app's own support networks. Motivational content would typically appear in the form of quotes without any integration of behavior-change principles (eg, Recovery Quotes, Clean & Sober Recovery, and Addiction Quotes).

Other features of low-quality apps included the use of basic functions such as motivational quotes, timers, or informational content to ultimately expose users to pop-up advertisements for a single private residential treatment program or clinician (eg, SoberBud, Sobriety Clock, and Stop Drinking Now). Other apps would solicit users to pay for the *full-version* app to receive more comprehensive recovery content (eg, Hypnosis for Alcoholism, Addiction and Recovery, and Sober Tree).

Some apps were concerning for the possibility of exacerbating substance use. The *Drugs* app offered informational content that was not easily accessible, and the forum included threads with individuals offering to sell drugs. The *Drive After: Alcohol calculator* app offered users information on strategies to *feel more sober* or conceal the odor of alcohol. *Best Alcohol Test* is intended to offer a blood alcohol level calculator but also offers users games to test their reflexes following binge-drinking episodes.

Another app described as *Best Home Cure for Alcoholism* (2.38) claimed to offer alternative remedies for alcohol use and withdrawal symptoms with topics such as *how to make someone stop drinking alcohol forever*; *how to quit alcohol ayurvedic*; and *how to stop alcohol drinking of my husband*. However, content was limited to increasing intake of fruits, vegetables, juices, water, and coffee without offering any empirical evidence on these approaches or elaborating on the credibility of the article authors and app designers.

Discussion

Summary

The initial search for apps targeting substance use yielded 74 apps; however, only 7 apps offered any evidence-based content, such as information on effective pharmacotherapies for SUDs (n=3), harm-reduction content (n=1), or behavior-change principles within app content or design features (n=3). None of the apps facilitated linkage to primary care–based treatment for SUDs, methadone maintenance treatment programs for Opioid use disorder (OUD), or clarified insurance requirements or costs related to available primary care or specialty addiction treatment programs.

Although none of the apps cited any empirical evidence suggesting potential clinical benefit or sustained engagement among users, app quality assessment via the MARS score offers a useful approach before evaluating for efficacy. The apps in this study had a low overall median quality MARS score of 2.81. However, the overall low information (2.81), engagement (2.75), and satisfaction (1.75) subscale scores highlight the lack of evidence-based content and gap in intervention design.

Although numerous apps emerged from the review targeting alcohol use, only STOP OD NYC and FlexDek MAT specifically targeted opioid use with evidence-based content (eg, effective pharmacotherapies for OUD). However, STOP OD NYC was limited to naloxone and overdose prevention and FlexDek MAT lacked basic functionality, was not aesthetically engaging, and the informational content was limited to PDF files and external links to the SAMHSA Web page. None of the commercially developed apps offered any informational content or access to online or clinical resources addressing harm-reduction practices and HIV and HCV prevention or screening content.

Few apps integrated evidence-based behavior change content (eg, SMART Recovery Cost Benefit) and mostly centered on basic informational content summarizing how many calories or money would be saved with alcohol abstinence, tracking time of abstinence with timers, using graphs to chart quantities of consumed alcoholic beverages, and basic information about addiction (eg, as a chronic disease), improvised *tips* on recovery, or quotes from the Bible that were not aligned with evidence-based psychotherapeutic approaches.

Empirical Evidence Demonstrating Smartphone App Efficacy

Although smartphone software apps are technologically capable to enhance care for SUDs with complex and multifaceted interventions, our review found that even apps with higher MARS ratings were typically limited to singular functions (eg, validated assessments such as Alcohol Use Disorders Identification Test (AUDIT) or Alcohol, Smoking and Substance Involvement Screening Test (ASSIST), active peer support systems, access to 12-step or SMART recovery group meetings, and/or linkage to specialty addiction treatment providers).

In a similar review of alcohol-related smartphone apps in 2012 by Weaver et al, only 44 of the 500 apps that met the initial inclusion criteria actually targeted reductions in alcohol use, and none presented evidence of efficacy [12]. Despite a doubling in apps targeting alcohol use in a subsequent review in 2014, there was no evidence of improved integration of evidence-based approaches (eg, effective psychotherapeutic interventions) and most promoted alcohol use [39]. Our review of commercially available alcohol-reduction apps is aligned with previous findings limiting their use in real-world clinical settings and is concerning for exacerbating relapse or worsening alcohol use [12,32].

MARS ratings of commercially available apps in this study also parallel the critical analysis of apps targeting heart failure and mindfulness and their lack of evidence-guided content [28,40]. These findings contrast with randomized controlled trials

demonstrating effectiveness for university-developed smartphone apps targeting alcohol use [2,41] and cravings [7,42]. However, our review found that these apps remain unavailable for individual use via existing app stores and without invitation by a licensed addiction treatment provider.

Although the apps lacked any clinical evidence of efficacy, there was also no evidence of sustained use. For instance, apps claiming to offer peer support via forums or SMS text messaging contact with other users were not active or unresponsive. In addition, forums lacked moderators or clinicians who could offer evidence-based responses to forum threads. Long-term engagement with technology-based interventions is critical to ensure behavior change and meaningful clinical outcomes. Challenges to larger-scale adoption of mHealth interventions include the lack of open mHealth frameworks that clarify underlying mechanisms linking intervention design features, effective psychotherapeutic approaches, and clinical outcomes. For instance, even basic process measures, such as the duration or intensity of app utilization, are not assessed or disclosed. Instead, users and researchers alike must rely on other user ratings and comments to gauge an app's potential benefit. Thus, preliminary studies are needed among participants in the community or addiction treatment settings assessing the impact of self-reported app usage targeting substance use and treatment utilization. Future smartphone app research requires elucidation of how app design features, content, and behavior-change principles impact targeted clinical outcomes. In July 2017, the Food and Drug Administration (FDA) launched the Digital Health Innovation Action Plan to offer additional oversight over clinical and patient decision support software. Further oversight by the FDA may facilitate transparency in app design, promotion of clinical studies assessing app safety and effectiveness, and increase the confidence of users and health systems for broader adoption [43].

Harm Reduction or Exacerbating Harmful Use?

Apps addressing alcohol use linked participants to app-based transportation companies or sober peers to facilitate rides if they were intoxicated. Some of the apps offering blood alcohol calculators for alcohol encouraged users to avoid reaching intoxication or hazardous drinking levels. However, other blood alcohol calculator apps would use cartoon imagery to gamify drinking and even allow users to compare their blood alcohol levels with other drinking peers. Another blood alcohol calculator app would offer users information about how to hide odors of alcohol in their breath or *sober up* if they exceeded certain levels of drinking. Our findings are aligned with reviews of smartphone apps in the last decade, emphasizing the availability of commercially developed apps that mostly exacerbate rather than mitigate harmful substance use [12,16,39]. Studies have reported the dramatic rise of apps promoting cigarette smoking [44], cannabis [15], and alcohol use [12,39]. In 2012, Bindhim et al searched for *cannabis, weed, marijuana, cocaine, heroin, and ecstasy* and reported an increase in harmful apps from 238 apps in February 2012 to 410 apps in May 2012 that encouraged contact with actively using peers, role-playing as cartel bosses or cannabis farmers, or simulating substance use [16]. Not surprisingly, in 2017, Google blocked approximately 700,000 of the nearly 3.5 million Android apps purged for promoting violence, hate, adult material, illicit activities, and substance use [45]. However, with 200 new apps entering the marketplace daily [19], app stores must develop more stringent restrictions considering the consistent growth in apps encouraging harmful substance use.

Linkage to Treatment

Although numerous descriptions attracted potential users to the app's capacity to link them to nearby treatment programs, nearly all lacked updated information about available programs, were not tailored to uninsured or Medicaid-insured patients, or would direct users to a single private practice therapist or residential treatment program even if the investigators were attempting to request an office-based opioid treatment program. The deceptive referral of all requests for treatment to a commercial advertisement or to a single private practice was common. Among apps developed by government agencies, users were able to eventually locate treatment programs after being redirected to government Web pages that listed available clinics but were not adapted for mobile phone–based internet browsers (eg, STOP OD NYC and Drive Sober Alabama). However, the STOP OD NYC app's overall design is the most ideally suited platform to integrate harm-reduction resources (ie, syringe exchange programs and naloxone) as well as linking users to low-cost office-based opioid treatment programs for buprenorphine and extended-release naltrexone treatment and the high density of 12-step and SMART recovery groups within New York City.

Integration of Behavior-Change Content

CBT and relapse prevention strategies offer a collaborative, individualized, psychological treatment recognized as effective approaches to generating behavioral, cognitive, and emotional adaption to a wide range of common psychological problems [46]. The efficacy of CBT and relapse prevention strategies has been supported by a comprehensive review of 106 meta-analyses across different clinical groups that also extends to SUDs. Despite their widespread adoption in academically developed smartphone apps (eg, A-CHESS), only the SMART cost-benefit app utilized behavior change models in this review.

SMART Recovery is based on both the Rational-Emotive-Behavior therapy and CBT approaches and reinforces learning skills to cope with (rather than avoid) emotional disturbances that exacerbate substance use [37,46]. The SMART Recovery website offers extensive resources, including articles, podcasts, videos, and self-help assignments that deepen user engagement with this approach. In addition, its online forum is active and offers unique discussion threads, including *Building and Maintaining Motivation*, *Coping with Urges*, *Managing Thoughts, Feelings, and Behaviors*, *Living a Balanced Life*, and specialized peer support group forums based on specific substances and a *family and friends* forum. The Smart Cost Benefit app has tremendous potential to integrate an already vibrant online forum and evidence-based psychotherapeutic approach.

Participation in self-help support groups, online forums, and even smartphone app communities can help motivate users to engage in healthy activities. A supportive app community can

help users share and discuss their recovery experiences and the challenges of regular practice. This could potentially complement or substitute for the support provided in face-to-face recovery treatment modalities. Although nearly many of the reviewed apps provided social network–sharing, few offered moderators or clinicians to guide discussions. Further research is needed to assess the impact of app-based forums and peer SMS text messaging to enhance engagement with the app and clinical outcomes.

Privacy Measures

Apps generally lacked the use of privacy measures to protect health information: few required password protection, elucidated the use of security certifications from cellphone providers, utilized encryption technologies for transmitted content, or 2-step verification during registration. Numerous apps asked participants for access to their Facebook profile and/or Google profile during registration and reassured users in the Terms and Conditions that their personal information would be safeguarded. Several apps were also used for research purposes but did not specify details of the research study, rights as a study subject, and how to terminate one's participation in the study and remove their data usage information. The Health Insurance Portability and Accountability Act and expert guidelines have outlined several measures to ensure the privacy of patient-physician communication in emerging health information technologies, including (1) the use of simple message content that refrains from disclosing patient name, diagnosis, or enrollment in treatment; (2) encouraging 2-step verification, password protection, and finger Touch identification; (3) obtaining security certifications from cellphone providers; (4) using encryption technologies for user responses; and (5) regularly deleting or setting expiration periods for communication content [47].

Limitations

This is one of the first comprehensive studies to review apps targeting illicit substances using the MARS scoring criteria and provide a reliable measure of engagement, functionality, visual appeal, and informational quality. Furthermore, this is the only review to assess for the integration of evidence-based psychotherapeutic or pharmacological approaches to SUDs and their applicability to office-based management of SUDs based on the medical management model. However, findings emerging from our descriptive analysis are not based on validated usability and/or efficacy study methods and require more rigorous study methods to assess the clinical impact. The review was limited to Google Play and Apple iOS and did not include apps available in F-Droid, Amazon Appstore, and GetJar, among other smaller scale app platforms. Our assessment of user privacy did not incorporate open-source developer codes for malicious purposes.

Conclusions

Online app stores offer unprecedented opportunities to expand access to effective harm reduction and treatment approaches for individuals with SUDs. However, the overall low MARS scale ratings and findings emerging from our descriptive analysis highlight the lack of evidence-based apps for individuals seeking additional support. Further studies are needed to assess the impact of existing evidence-based apps described in this review (eg, STOP OD NYC and FlexDek MAT). Investigators should leverage online app stores to assess the acceptability and clinical impact of effective apps targeting SUDs that are not yet available within online app stores for individual use (eg, A-CHESS). Finally, public health experts should utilize the popularity of online app stores to offer user-friendly and evidence-based apps that facilitate access to effective pharmacotherapies for SUDs, harm-reduction resources (eg, naloxone and syringe exchange programs), specialty addiction treatment programs (eg, intensive outpatient programs and methadone maintenance programs), and linkage to primary care–based treatment for SUDs.

Acknowledgments

BT is supported by an NIH Mentored Patient-Oriented Research Career Development Award (NIDA K23DA042140-01A1) and Clinical Translational Science Award (UL1 TR001445).

Authors' Contributions

BT, CC, LH, JRV, and PH made substantial contributions to conception, design, and writing of the manuscript.

Conflicts of Interest

None declared.

References

1. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
2. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
3. Brendryen H, Kraft P. Happy ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. Addiction 2008 Mar;103(3):478-84; discussion 485. [doi: 10.1111/j.1360-0443.2007.02119.x] [Medline: 18269367]

4. Tofighi B, Nicholson JM, McNeely J, Muench F, Lee JD. Mobile phone messaging for illicit drug and alcohol dependence: a systematic review of the literature. Drug Alcohol Rev 2017 Dec;36(4):477-491 [FREE Full text] [doi: 10.1111/dar.12535] [Medline: 28474374]
5. Smith A. Pew Research Center. 2017. Record Shares of Americans Now Own Smartphones, Have Home Broadband URL: http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/ [accessed 2019-03-15] [WebCite Cache ID 76suJpXmK]
6. Quanbeck AR, Gustafson DH, Marsch LA, McTavish F, Brown RT, Mares M, et al. Integrating addiction treatment into primary care using mobile health technology: protocol for an implementation research study. Implement Sci 2014 May 29;9:65 [FREE Full text] [doi: 10.1186/1748-5908-9-65] [Medline: 24884976]
7. Fowler LA, Holt SL, Joshi D. Mobile technology-based interventions for adult users of alcohol: a systematic review of the literature. Addict Behav 2016 Nov;62:25-34. [doi: 10.1016/j.addbeh.2016.06.008] [Medline: 27310031]
8. Marsch LA. Leveraging technology to enhance addiction treatment and recovery. J Addict Dis 2012;31(3):313-318 [FREE Full text] [doi: 10.1080/10550887.2012.694606] [Medline: 22873192]
9. IQVIA. 2017 Nov 7. The Growing Value of Digital Health URL: https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health [accessed 2019-03-13] [WebCite Cache ID 76qc4ZE52]
10. Duncan LG, Mendoza S, Hansen H. Buprenorphine maintenance for opioid dependence in public sector healthcare: benefits and barriers. J Addict Med Ther Sci 2015;1(2):31-36 [FREE Full text] [doi: 10.17352/2455-3484.000008] [Medline: 27088135]
11. Penzenstadler L, Chatton A, Van Singer M, Khazaal Y. Quality of smartphone apps related to alcohol use disorder. Eur Addict Res 2016;22(6):329-338. [doi: 10.1159/000449097] [Medline: 27598779]
12. Weaver ER, Horyniak DR, Jenkinson R, Dietze P, Lim MS. "Let's get wasted!" and other apps: characteristics, acceptability, and use of alcohol-related smartphone applications. JMIR Mhealth Uhealth 2013;1(1):e9 [FREE Full text] [doi: 10.2196/mhealth.2709] [Medline: 25100681]
13. Tofighi B, Leonard N, Greco P, Hadavand A, Acosta MC, Lee JD. Technology use patterns among patients enrolled in inpatient detoxification treatment. J Addict Med 2018 Dec 20 [FREE Full text] [doi: 10.1097/ADM.0000000000000494] [Medline: 30589653]
14. Lupton D, Jutel A. Soc Sci Med. 2015. 'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps URL: https://linkinghub.elsevier.com/retrieve/pii/S0277953615002245 [accessed 2019-03-13] [WebCite Cache ID 76qcHLZwk]
15. Ramo DE, Popova L, Grana R, Zhao S, Chavez K. Cannabis mobile apps: a content analysis. JMIR Mhealth Uhealth 2015 Aug 12;3(3):e81 [FREE Full text] [doi: 10.2196/mhealth.4405] [Medline: 26268634]
16. Bindhim NF, Naicker S, Freeman B, Mcgeechan K, Trevena L. Apps promoting illicit drugs—a need for tighter regulation? J Con Health Internet 2014 Feb 24;18(1):31-43 [FREE Full text] [doi: 10.1080/15398285.2014.869166]
17. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: a systematic review of prevalent and innovative methods. Am J Prev Med 2016 Dec;51(6):1051-1059. [doi: 10.1016/j.amepre.2016.07.009] [Medline: 27659122]
18. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
19. Jahns RG. Research2Guidance. 2015 Jan 1. 500m people will be using healthcare mobile applications in 2015 URL: https://research2guidance.com/500m-people-will-be-using-healthcare-mobile-applications-in-2015-2/ [accessed 2019-03-13] [WebCite Cache ID 76qczbyrL]
20. Saitz R, Cheng DM, Winter M, Kim TW, Meli SM, Allensworth-Davies D, et al. Chronic care management for dependence on alcohol and other drugs: the AHEAD randomized trial. J Am Med Assoc 2013 Sep 18;310(11):1156-1167 [FREE Full text] [doi: 10.1001/jama.2013.277609] [Medline: 24045740]
21. Weiss RD, Potter JS, Provost SE, Huang Z, Jacobs P, Hasson A, et al. A multi-site, two-phase, Prescription Opioid Addiction Treatment Study (POATS): rationale, design, and methodology. Contemp Clin Trials 2010 Mar;31(2):189-199 [FREE Full text] [doi: 10.1016/j.cct.2010.01.003] [Medline: 20116457]
22. Tofighi B, Grossman E, Bereket S. Text message content preferences to improve buprenorphine maintenance treatment in primary care. J Addict Dis 2016;35(2):92-100. [doi: 10.1080/10550887.2015.1127716] [Medline: 26670868]
23. Galanter M. Spirituality and recovery in 12-step programs: an empirical model. J Subst Abuse Treat 2007 Oct;33(3):265-272 [FREE Full text] [doi: 10.1016/j.jsat.2007.04.016] [Medline: 17889297]
24. Volkow ND, Collins FS. The role of science in addressing the opioid crisis. N Engl J Med 2017 Jul 27;377(4):391-394 [FREE Full text] [doi: 10.1056/NEJMsr1706626] [Medline: 28564549]
25. Tinschert P, Jakob R, Barata F, Kramer J, Kowatsch T. The potential of mobile apps for improving asthma self-management: a review of publicly available and well-adopted asthma apps. JMIR Mhealth Uhealth 2017 Aug 2;5(8):e113 [FREE Full text] [doi: 10.2196/mhealth.7177] [Medline: 28768606]
26. Athilingam P, Jenkins B. Mobile phone apps to support heart failure self-care management: integrative review. JMIR Cardio 2018 May 2;2(1):e10057 [FREE Full text] [doi: 10.2196/10057]

27. Adam A, Hellig J, Perera M, Bolton D, Lawrentschuk N. 'Prostate Cancer Risk Calculator' mobile applications (apps): a systematic review and scoring using the validated user version of the Mobile Application Rating Scale (uMARS). World J Urol 2018 Apr;36(4):565-573. [doi: 10.1007/s00345-017-2150-1] [Medline: 29222595]
28. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
29. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]
30. Dallery J, Jarvis B, Marsch L, Xie H. Mechanisms of change associated with technology-based interventions for substance use. Drug Alcohol Depend 2015 May 1;150:14-23 [FREE Full text] [doi: 10.1016/j.drugalcdep.2015.02.036] [Medline: 25813268]
31. National Institute on Drug Abuse. Principles Of Drug Addiction Treatment: A Research-based Guide: Third Edition. Bethesda, Maryland: Createspace Independent Publishing Platform; 2018.
32. Volkow ND. Medications for opioid use disorder: bridging the gap in care. Lancet 2018 Dec 27;391(10118):285-287 [FREE Full text] [doi: 10.1016/S0140-6736(17)32893-3] [Medline: 29150199]
33. Lee JD, Vocci F, Fiellin DA. Unobserved. J Addict Med 2014;8(5):299-308. [doi: 10.1097/ADM.0000000000000059] [Medline: 25254667]
34. Lee JD, Grossman E, Huben L, Manseau M, McNeely J, Rotrosen J, et al. Extended-release naltrexone plus medical management alcohol treatment in primary care: findings at 15 months. J Subst Abuse Treat 2012 Dec;43(4):458-462 [FREE Full text] [doi: 10.1016/j.jsat.2012.08.012] [Medline: 22985676]
35. Anderson P, Scott E. The effect of general practitioners' advice to heavy drinking men. Br J Addict 1992 Jun;87(6):891-900 [FREE Full text] [doi: 10.1111/j.1360-0443.1992.tb01984.x] [Medline: 1525531]
36. Spivak K, Sanchez-Craig M, Davila R. Assisting problem drinkers to change on their own: effect of specific and non-specific advice. Addiction 1994 Sep;89(9):1135-1142 [FREE Full text] [doi: 10.1111/j.1360-0443.1994.tb02790.x] [Medline: 7987190]
37. Kelly PJ, Deane FP, Baker AL. Group cohesion and between session homework activities predict self-reported cognitive-behavioral skill use amongst participants of SMART Recovery groups. J Subst Abuse Treat 2015 Apr;51:53-58 [FREE Full text] [doi: 10.1016/j.jsat.2014.10.008] [Medline: 25535099]
38. Methadone.US. URL: http://www.methadone.us/ [accessed 2019-03-15] [WebCite Cache ID 76tD6vzuy]
39. Crane D, Garnett C, Brown J, West R, Michie S. Behavior change techniques in popular alcohol reduction apps: content analysis. J Med Internet Res 2015;17(5):e118 [FREE Full text] [doi: 10.2196/jmir.4060] [Medline: 25977135]
40. Masterson Creber RM, Maurer M, Reading M, Hiraldo G, Hickey K, Iribarren S. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2016 Jun 14;4(2):e74 [FREE Full text] [doi: 10.2196/mhealth.5882] [Medline: 27302310]
41. Gonzalez VM, Dulin PL. Comparison of a smartphone app for alcohol use disorders with an internet-based intervention plus bibliotherapy: a pilot study. J Consult Clin Psychol 2015 Apr;83(2):335-345. [doi: 10.1037/a0038620] [Medline: 25622202]
42. Dulin PL, Gonzalez VM. Smartphone-based, momentary intervention for alcohol cravings amongst individuals with an alcohol use disorder. Psychol Addict Behav 2017 Dec;31(5):601-607 [FREE Full text] [doi: 10.1037/adb0000292] [Medline: 28703611]
43. Shuren J, Patel B, Gottlieb S. FDA regulation of mobile medical apps. J Am Med Assoc 2018 Jul 24;320(4):337-338. [doi: 10.1001/jama.2018.8832] [Medline: 29971339]
44. BinDhim NF, Freeman B, Trevena L. Pro-smoking apps for smartphones: the latest vehicle for the tobacco industry? Tob Control 2014 Jan;23(1):e4. [doi: 10.1136/tobaccocontrol-2012-050598] [Medline: 23091161]
45. Steinberg J. Inc Magazine. 2018 Feb 1. Google removed 700,000 problematic apps From its android play store in 2017 URL: https://www.inc.com/joseph-steinberg/google-removed-700000-problematic-apps-from-its-android-play-store-in-2017.html [accessed 2019-03-13] [WebCite Cache ID 76qdn7jVN]
46. Kellogg S, Kreek MJ. On blending practice and research: the search for commonalities in substance abuse treatment. Subst Abus 2006 Jun;27(1-2):9-24 [FREE Full text] [doi: 10.1300/J465v27n01_03] [Medline: 17062541]
47. Tofighi B, Grossman E, Sherman S, Nunes EV, Lee JD. Mobile Phone Messaging During Unobserved. J Addict Med 2016;10(5):309-313. [doi: 10.1097/ADM.0000000000000198] [Medline: 26933874]

Abbreviations

ASSIST: Alcohol, Smoking and Substance Involvement Screening Test
AUDIT: Alcohol Use Disorders Identification Test
CBT: Cognitive Behavioral Therapy
FDA: Food and Drug Administration
HCV: hepatitis C virus

ICC: intraclass correlation
MARS: Mobile App Rating Scale
mHealth: mobile health
OBOT: Office based opioid treatment
OUD: Opioid use disorder
SAMHSA: Substance Abuse and Mental Health Services Administration
SMART: Self-Management and Recovery Training
SMS: short message service
SUD: substance use disorder

Edited by G Eysenbach; submitted 03.08.18; peer-reviewed by P Dulin, J Glass, D Gustafson, A Carr; comments to author 07.10.18; revised version received 24.10.18; accepted 12.12.18; published 22.04.19.

Please cite as:
Tofighi B, Chemi C, Ruiz-Valcarcel J, Hein P, Hu L
Smartphone Apps Targeting Alcohol and Illicit Substance Use: Systematic Search in in Commercial App Stores and Critical Content Analysis
JMIR Mhealth Uhealth 2019;7(4):e11831
URL: https://mhealth.jmir.org/2019/4/e11831/
doi:10.2196/11831
PMID:31008713

Original Paper

To Text or Not to Text: Electronic Message Intervention to Improve Treatment Adherence Versus Matched Historical Controls

Marily A Oppezzo[1], PhD, MS, RD; Michael V Stanton[2], PhD; Ariadna Garcia[3], MS; Joseph Rigdon[3], PhD; Jae R Berman[1], MSc; Christopher D Gardner[1], PhD

[1]Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, United States

[2]Department of Health Sciences, California State University, East Bay, Hayward, CA, United States

[3]Quantitative Sciences Unit, Stanford University School of Medicine, Stanford, CA, United States

Corresponding Author:
Marily A Oppezzo, PhD, MS, RD
Stanford Prevention Research Center
Stanford University School of Medicine
1265 Welch Road
Stanford, CA, 94305-5411
United States
Phone: 1 650 723 6254
Email: moppezzo@stanford.edu

Abstract

Background: Ensuring treatment adherence is important for the internal validity of clinical trials. In intervention studies where touch points decrease over time, there is even more of an adherence challenge. Trials with multiple cohorts offer an opportunity to innovate on ways to increase treatment adherence without compromising the integrity of the study design, and previous cohorts can serve as historical controls. Electronically delivered nudges offer low-cost opportunities to increase treatment adherence.

Objective: This study aimed to evaluate the effectiveness of electronic messages (e-messages) on treatment adherence to the last cohort of a parent weight loss intervention during the second half of a year-long trial, when intervention checkpoint frequency decreases. Treatment adherence is measured by intervention class attendance and adherence to the intervention diet.

Methods: All participants in the last cohort (cohort 5, n=128) of a large randomized weight loss study were offered an e-message intervention to improve participant adherence during the last 6 months of a 1-year weight loss program. Overall, 3 to 4 electronic weekly messages asked participants about intervention diet adherence. A propensity score model was estimated using 97 participants who opted to receive e-messages and 31 who declined in cohort 5 and used to pair match cohort 5 e-message participants to a historical control group from cohorts 1 to 4. Moreover, 88 participants had complete data, yielding 176 participants in the final analyses. After matching, intervention and matched control groups were compared on (1) proportion of class attendance between the 6 and 12 month study endpoints, (2) diet adherence, as measured by total carbohydrate grams for low-carbohydrate (LC) and total fat grams for low-fat (LF) diets at 12 months, and (3) weight change from 6 to 12 months. The dose-response relationship between the proportion of text messages responded to and the 3 outcomes was also investigated.

Results: Compared with matched controls, receiving e-messages had no effect on (1) treatment adherence; class attendance after 6 months +4.6% (95% CI −4.43 to 13.68, *P*=.31), (2) adherence; LC −2.5 g carbohydrate, 95% CI −29.9 to 24.8, *P*=.85; LF +6.2 g fat, 95% CI −4.1 to 17.0, *P*=.26); or on (3) the secondary outcome of weight change in the last 6 months; +0.3 kg (95% CI −1.0 to 1.5, *P*=.68). There was a positive significant response correlation between the percentage of messages to which participants responded and class attendance (*r*=.45, *P*<.001).

Conclusions: Although this e-message intervention did not improve treatment adherence, future studies can learn from this pilot and may incorporate more variety in the prompts and more interaction to promote more effective user engagement. Uniquely, this study demonstrated the potential for innovating within a multicohort trial using propensity score–matched historical control subjects.

Trial Registration: ClinicalTrials.gov NCT01826591; https://clinicaltrials.gov/ct2/show/NCT01826591

International Registered Report Identifier (IRRID): RR2-10.1016/j.cct.2016.12.021

(JMIR Mhealth Uhealth 2019;7(4):e11720) doi:10.2196/11720

KEYWORDS

treatment adherence; intervention; short message service; mobile health; propensity score

Introduction

Background

An important challenge of behavioral interventions is treatment adherence. Treatment adherence strategies increase participants' enactment of the intervention delivered, thereby increasing the internal validity of the results and accuracy of conclusions drawn from the study [1,2]. Weight loss diet interventions often show high recidivism, with adherence to the treatment an inherent challenge, and often built into the intervention itself [3-6]. With adherence being highly variable in diet studies, it becomes difficult to evaluate the effectiveness of diet types. Applying treatment adherence strategies to all treatment arms, or diet groups, would extend participant contact and enhance the strength of study conclusions. Furthermore, in large trials with multiple cohorts or staggered enrollment, there is an opportunity to utilize lessons learned from early enrollees to design protocol modifications to increase treatment adherence in later enrollees. This would be the case for the Diet Intervention Examining The Factors Interacting with Treatment Success (DIETFITS) study, which contrasted healthy low-carbohydrate (LC) versus healthy low-fat (LF) diets for a sample of 609 generally healthy, free-living adults with body mass index 28 to 40 kg/m^2 [7].

Objectives

This paper highlights the deployment of a treatment adherence enhancement for the fifth cohort of a large 5-cohort, weight loss trial testing the differential effects of adhering to randomly assigned, high-quality, LC or LF diets. The novelty of this paper is three-fold: first, our intervention demonstrates dynamic trial-design potential. It was developed to increase treatment adherence and was a response to feedback from the previous 4 cohorts from months 7 to 12 of the 12-month protocol. Our intervention was then deployed to *both* diet treatment arms of the parent study in cohort 5. Second, the evaluation of the effectiveness of our intervention capitalized on the preceding 4 cohorts to provide historical, matched control participants. Third, our intervention targeted *adherence to* the treatment (ie, in-person class attendance and adherence to the diet of either LC or LF content), *not* the intended outcome of the treatment (ie, weight loss). We explored weight loss only as a secondary outcome.

The National Institutes of Health's Behavior Change Consortium's Treatment Fidelity Workgroup outlined goals for treatment fidelity strategies across 5 research process phases [1]. Our intervention was designed to increase the *treatment skills enactment* phase of the larger DIETFITS trial. We define the treatment skills as attending the classes for their diet treatment (class attendance) and adhering to their assigned diet (diet adherence). Although our intervention is under the NIH Behavior Change Consortium's umbrella of treatment fidelity, we will refer to it as treatment adherence in this paper.

The DIETFITS intervention gave both LC and LF groups weekly (8 classes), biweekly (5 classes), every third week (4 classes), and monthly (5 classes over months 7-12) education classes designed to improve participant adherence to the highest-quality version of their assigned diet. Postintervention focus groups and surveys from the first 4 cohorts of the DIETFITS trial indicated that participants desired more accountability during months 7 to 12, when face-to-face study contact substantially decreased. In addition, objective data from the first 4 cohorts showed a decrease in class attendance during months 7 to 12, averaging 46.8% compared with 73.8% in months 1 to 6. Diet adherence in the DIETFITS study had the broad guideline of reaching as low as the participants could go in carbohydrate or fat grams during the first 8 weeks and then titrating slowly back up to find the lowest amount they could maintain for the study duration. Adherence data, in the form of 24-hour recalls, were collected via phone at 4 major data collection time points during the DIETFITS trial. Similar to other studies, recidivism in diet studies tends to start after participants have lost significant amounts of weight and are having difficulty in maintaining behavioral changes [8,9], commonly 6 months. For cohorts 1 to 4 in DIETFITS, at 6 months, the healthy LC diet averaged 111.6 g of carbohydrates versus 128.0 g at 12 months, and the healthy LF diet averaged 50.0 g of fat at 6 months versus 56.1 g at 12 months. This average increase in grams of carbohydrates or fat during the second 6 months shows slight recidivism.

In response to participants' requests and for the purpose of minimizing recidivism and maximizing adherence, we sought to increase treatment skills enactment during months 7 to 12 for cohort 5, specifically increasing class attendance and adherence to their diet. Extended contact after the initial intervention contact points have decreased has been shown to support long-term diet behavior change [10]. However, extended contact via face to face is costly and time-consuming. Our treatment adherence enhancement intervention (henceforth referred to as the e-message intervention) utilized short message service (SMS) text messages or email to enhance adherence. Electronically delivered messages are convenient, cost-effective, asynchronous (ie, can be read by participants at times that suits the individual) delivery channels for behavior change intervention touches and do not require labor-intensive face-to-face contact [11]. Using the internet to deliver periodic prompts to a behavior change method has been effective in 11 of the 19 reviewed studies [12], and another review article found that 13 of the 14 studies reviewed SMS-delivered interventions with positive, short-term effects on behavior change [11]. One weight loss study delivered 4 times daily SMS text messages with tips, self-monitoring reminders, and motivational messages and found that although message delivery was not a sufficient stand-alone intervention compared with a control group for significant weight loss, participants who responded to a greater proportion of SMS text messages tended to have the greatest weight loss at 6 and 12 months [13]. Another 4-month trial found significant weight loss in a 2- to 5- times daily personalized SMS text message intervention compared with a control group [14]. Although there is some evidence of success when using mobile technology to deliver interventions, less

research has been conducted on using SMS text messages to extend intervention contacts or enhance treatment adherence [2,15]. One recent trial showed an average of 8 SMS text messages every 2 weeks as an intervention contact extension–supported attenuation of weight gain compared with those without the contact extension [16]. Our work builds on this success, using SMS text messages or email as a low-cost, adjunct strategy to extend contact when face-to-face contact substantially decreases. Distinctly, our e-message contacts target enhancing treatment adherence.

This study is an innovative approach to modifying an existing randomized controlled trial (RCT), after it had begun, to increase treatment adherence. In response to the first 4 RCT cohorts' request for more accountability in the latter half of their treatment when experimenter interaction decreased, we created an e-message intervention and applied it to both arms of the RCT for cohort 5. Our e-message intervention was designed to increase all participants' adherence to the treatment assigned, in this case, a diet plan. Treatment adherence was measured by class attendance and adherence to the diet assigned. Weight loss was a secondary outcome. We used propensity score matching to historical controls in the previous 4 cohorts.

Methods

Study Participants

Parent Trial

The DIETFITS trial is an institutional review board-approved large, randomized, weight loss study in which 609 participants were randomly assigned to either a (LC) or (LF) diet, with 16 classes spread throughout months 1 to 6 and 6 monthly booster classes during months 7 to 12. The primary dietary goal for both diet groups was to "go as low as you can go, and maintain" for carbohydrates in the LC group and for fat in the LF group; there was no set target number of grams or percent of calories to reach; therefore, the lower the grams of carbohydrate or fat achieved, the more adherent this was considered [17]. The trial was split into 5 cohorts over a 3-year period, with the first cohort (cohort 1) beginning in April 2013 and the final measurements in the last cohort (cohort 5) taken in March 2016.

This Study

Cohort 5 had 128 participants offered the e-message intervention during their 6-month class; 97 participants provided consent. Designed as an SMS text message study, to maximize inclusivity and allow participant preference, we offered an email-delivery option.

Electronic Message Development

Verbal and written feedback from surveys and focus groups with participants in the first 4 cohorts of the study voiced concern that during months 7 to 12, when the class meeting frequency dropped to monthly, there was a feeling of decreased experimenter support and inadequate participant accountability. Decreased 7- to 12-month class attendance and diet adherence supported these subjective accounts. This e-message protocol enhancement was designed in response to this to increase class attendance and diet adherence for both LC and LF groups in cohort 5.

The e-message intervention was grounded in the nudge framework, whereby small touches can elicit significant behavior changes [18]. At a high level, receiving the SMS text or email message provides a cue to the treatment, a reminder that the participant is still in active treatment even though study interaction had decreased in frequency. More specifically, the questions elicited awareness of (1) the discrepancy between participants' current behavior and their goal behavior and (2) their emotional response and subsequent coping behavior.

The first question, "how adherent have you been to your eating plan since your last survey?", was based on the cybernetic model of self-control, which suggests that monitoring for discrepancies between the goal and current behavior can trigger behavioral corrections to mitigate the gap [19]. The subsequent 3 questions were based on feedback from the health coaches who taught the diet classes to participants in cohorts 1 to 4. They noted that many of the participants struggled with emotional eating and reported that discrepancy between goal and state increased the likelihood of the *what the hell* effect [20]. Research indicates that increased attention to emotional state is associated with a decreased emotional eating response [21]. Therefore, Question 2 showed a 7-point Likert scale depicting 7 faces from extreme negative to extreme positive emotion, with the text: "Based on the images below, how are you feeling about your adherence to your eating plan? Please click on the appropriate image." Question 3 asked "What words would you use to describe your current feelings about your eating plan?" The possible responses were selected to fit in a 3×2 organizational structure of emotional responses to one's goal actions: temporality (prospective, current, and retrospective) by valence (positive and negative) . Finally, Question 4 asked the participant's behavioral response to the current emotion: "Given your current feelings and current eating plan adherence, which of the following actions are you motivated to do, if anything?" (see Multimedia Appendix 1). The responses to the e-messages are beyond the scope of this efficacy-only analysis. Importantly, though back-and-forth interactions were not a feature, soliciting responses to the questions from the participant made the SMS text messages 2-way, rather than 1-way pushes.

Electronic Message Procedure

All consenting cohort 5 study participants who opted to receive the e-message intervention received 1 text or email with a link to a REDCap survey of 4 questions sent 3 to 4 times per week on randomly selected days each week [22]. Message frequency was based on Spark et al's average contacts, with the goal of having regular, but not overwhelming, contact [16]. If the participant did not respond after 6 to 8 hours, 1 reminder was sent. Participants were encouraged to reply to each question, and each response was recorded within REDCap.

Measures

Attendance was recorded for all 22 assigned classes. Weight was collected at baseline, 3 months, 6 months, and 12 months. Demographics such as age, gender, and race were collected at baseline. E-message responses were collected after the

introduction at the 6-month mark. In this paper, we focus on 3 key outcomes: (1) proportion of class attendance at months 7 to 12; (2) diet adherence measured by 3 unannounced 24-hour recalls, as average grams of either total carbohydrate (LC) or total fat (LF) at 12 months, target being lower than matched controls total grams; and (3) weight change from 6 to 12 months.

As the assigned diets were to healthfully "go as low as you can sustain" in grams of either carbohydrates or fat, dietary adherence was defined by reduction of grams of the target diet. Dietary intake was assessed by using 3 unannounced 24-hour dietary recalls at each of the 4 major data collection time points throughout the yearlong study. Data were collected using the Nutrition Database System for Research, a computer-based software application developed at the University of Minnesota Nutrition Coordinating Center. Data were collected using a standardized multiple-pass interview approach to increase accuracy. Average daily grams of carbohydrates and fats were calculated from the 3 24-hour diet recalls. Even though there is bias in 24-hour recalls, because food records potentially have reactivity bias, the 24-hour recall is considered the least biased of the self-reported instruments and the best single dietary assessment instrument for many purposes [23]. Using 2 weekdays and 1 weekend day and not announcing the assessment days are methods of minimizing some error [23]. For more details, consult the separate methods paper for this study [17].

In addition, within the cohort 5 group that consented to participate, the relationship between the percentage of e-messages responded to (engagement) and all 3 outcomes mentioned in the previous paragraph was investigated.

Two questions were given at the end of the intervention for participants to evaluate the perceived usefulness of the e-messages and their preferred frequency of e-message receipt. Perceived usefulness was measured on a 5-point Likert scale with labels: 1, *not at all useful*; 2, *slightly useful*; 3, *somewhat useful*; 4, *moderately useful*; and 5, *extremely useful*. Preferred e-message frequency had 6 response choices. Of them, 4 response choices were in decreasing frequency of messages per week: *6-7*, *4-5*, *2-3*, or *1*. *Not sure* and *text messages did not help* were the final response choices. Open-text space was provided for intervention comments or improvement suggestions.

Study Design

This was a longitudinal observational study of those who opted to participate in the e-message intervention in cohort 5, with a comparison with their matched historical controls from cohorts 1 to 4.

Statistical Analysis

Propensity score matching can control for observed potential confounding covariates [24,25]. A propensity score model was used to generate a matched control group for the e-message participants in cohort 5 [26]. The propensity score model was developed using only cohort 5 data and capitalizes on cohort 5 having both participants who selected to be in the e-message study (n=97) and participants who were offered but chose not to participate (n=31). The propensity score model included the following covariates: age at 6 months, sex, race, median-centered weight change before 6 months, median-centered proportion of attendance before 6 months, and text consent. The propensity score model was estimated using data from cohort 5 only as cohorts 1 to 4 did not have the opportunity to refuse consent to the e-message intervention.

Participants in cohort 5 who selected to participate in the e-message intervention were matched 1:1 to individuals in cohorts 1 to 4 using the following variables: propensity score (P-score: propensity to participate or not, calculation described above), age, sex, race, proportion of classes attended before 6 months, and weight change before 6 months [27]. Absolute standardized differences between the e-message group and the controls were checked both pre- and postmatch to ensure balance was achieved by the matching procedure. The matching procedure generated a control group comparable with the treatment group on potential confounding variables, allowing us to attribute differences in our outcomes to treatment (e-message or control), assuming no unmeasured confounders. Matching was executed using the R package *optmatch* [27]. After generating the control group, a 2-sample *t* test was used to assess the differences between the e-message group and the control group in the 2 primary study outcomes (attendance and diet adherence) and 1 secondary outcome (weight loss maintenance). Diet adherence was measured by grams of fat for LF and grams of carbohydrates for LC, and consequently, adherence was analyzed within each diet as the ranges of fat and carbohydrate could not easily be compared (eg, fat on the order of 60 g and carbohydrate on the order of 130 g).

Scatterplots and Spearman correlations were used to assess the relationship between the proportion of e-messages responded to and the 3 outcomes described in a previous section: (1) proportion of class attendance after 6 months, (2) diet adherence measured as average total carbohydrate grams for LC and average total fat grams for LF at 12 months, (3) 12-month minus 6-month weight change. All statistical analyses were performed using R version 3.3.3 [28].

Results

Method of Message Delivery

Of the 88 participants with complete data for the analyses, 46 chose SMS delivery, 36 chose email, and 6 chose both. No differences in the primary outcomes were found by the method of participation (see Multimedia Appendix 2). Therefore, we collapsed the groups for the analyses.

Baseline Characteristics

Baseline demographics such as age, gender, race, weight change before 6 months, and proportion of class attendance before 6 months are displayed by intervention and historical control group in Table 1, along with estimated propensity scores (P-scores), both pre- and postmatch. Prematch and postmatch adjusted means, standardized differences, and *P* values were also included to compare the balance between the intervention and control groups before and after the match. Control group before the matching refers to all of cohorts 1 to 4 with sufficient data for the match (n=376); control group postmatch was 1:1

nearest neighbor matched to cohort 5 e-message participants with no missing weight data at 6 months (n=88).

The mean age for the intervention group was 39.5 years, whereas the mean age for the control group was 41.2 and 39.6 years pre- and postmatch, respectively. Similarly, the mean proportion of attendance before 6 months in the intervention group was 79.4, whereas the mean proportion of attendance before 6 months was 81.8 and 80.3 pre- and postmatch, respectively. The standardized differences postmatch for the intervention and control groups were closer to 0 than those for the prematch group, and all passed the rule of thumb of 0.2 for small effect sizes [29].

Standardized differences postmatch were 0 for gender and race and −0.02, −0.01, −0.03, and −0.07 for P-score, age, weight change before 6 months, and proportion of class attendance before 6 months, respectively. Figure 1 shows the balance between intervention and control pre- and postmatch. P-score, age, and Hispanic ethnicity were found to be significant prematch, but all variables were nonsignificant postmatch. As these results indicated that postmatch intervention and control groups were better balanced, the study hypotheses were only tested using postmatch groups.

Effectiveness of Electronic Message Intervention

No statistically significant differences were seen in proportion of class attendance after 6 months between participants who received the e-message intervention (mean 57.0, SD 29.6) and those who did not (mean 52.4, SD 31.2); mean difference of intervention minus control was 4.6 (95% CI −4.4 to 13.7, *P*=.31; see Table 2). No significant differences were seen in diet adherence (total grams of carbohydrate per day) between participants on the LC diet in the intervention group (mean 133.7, SD 49.2, n=43) and those on the LC diet in the control group (mean 136.2, SD 63.3, n=39); mean difference was −2.5 (95% CI −29.9 to 24.8, *P*=.85). Similar results were observed for participants in the LF diet (grams of fat per day) between the intervention (mean 63.6, SD 23.4, n=45) and control groups (mean 57.4, SD 26.8, n=49); mean difference was +6.2 (95% CI −4.6 to 17.0, *P*=.26). No significant differences were seen in weight change between participants who received the intervention (mean 2.0, SD 4.3) and the control group (mean 2.0, SD 3.4); mean difference was 0.3 (95% CI −1.0 to 1.5, *P*=.68; see Table 2).

Table 1. Descriptive statistics for study variables in the electronic message group versus the control group pre- and postmatch. Means reported for age, weight change, class attendance; proportions reported for sex and race. SD=standardized difference (unitless measure of similarity, where closer to 0 is more similar). Variables with blank SD were not included in the match.

Variable	Prematch				Postmatch			
	Control (n=376)	Intervention (n=88)	SD	*P* value	Control (n=88)	Intervention (n=88)	SD	*P* value
Age (years)	41.2	39.5	−0.27	.03	39.6	39.5	−0.01	.70
Female (prop[a])	0.57	0.66	0.18	.13	0.66	0.66	0	>.99
Hispanic (prop)	0.2	0.3	0.25	.03	0.3	0.3	0	>.99
Black (prop)	0.03	0.02	−0.07	.57	0.02	0.02	0	>.99
Asian/Pacific Islander (prop)	0.12	0.08	−0.13	.28	0.08	0.08	0	>.99
Other (prop)	0.05	0.08	0.14	.24	0.08	0.08	0	>.99
Weight (kg) change 6 months	−7.5	−6.2	0.21	.07	−6.0	−6.2	−0.03	.50
Class attendance 6 months (%)	81.8	79.4	−0.17	.16	80.4	79.4	−0.07	.15
Propensity score	0.84	0.86	0.33	.01	0.87	0.86	−0.02	.26
Adherence at 6 months								
Low carb (g)	112.8[b]	115.9[c]	—[d]	—	114.9[e]	115.9[c]	—	—
Low fat (g)	217.2[f]	206.7[g]	—	—	217.1[h]	206.7[g]	—	—

[a]prop: proportions.
[b]n=195.
[c]n=43.
[d]Not applicable.
[e]n=39.
[f]n=181.
[g]n=45.
[h]n=49.

Figure 1. Standardized differences in study variables in electronic message group vs control group pre- and postmatch.

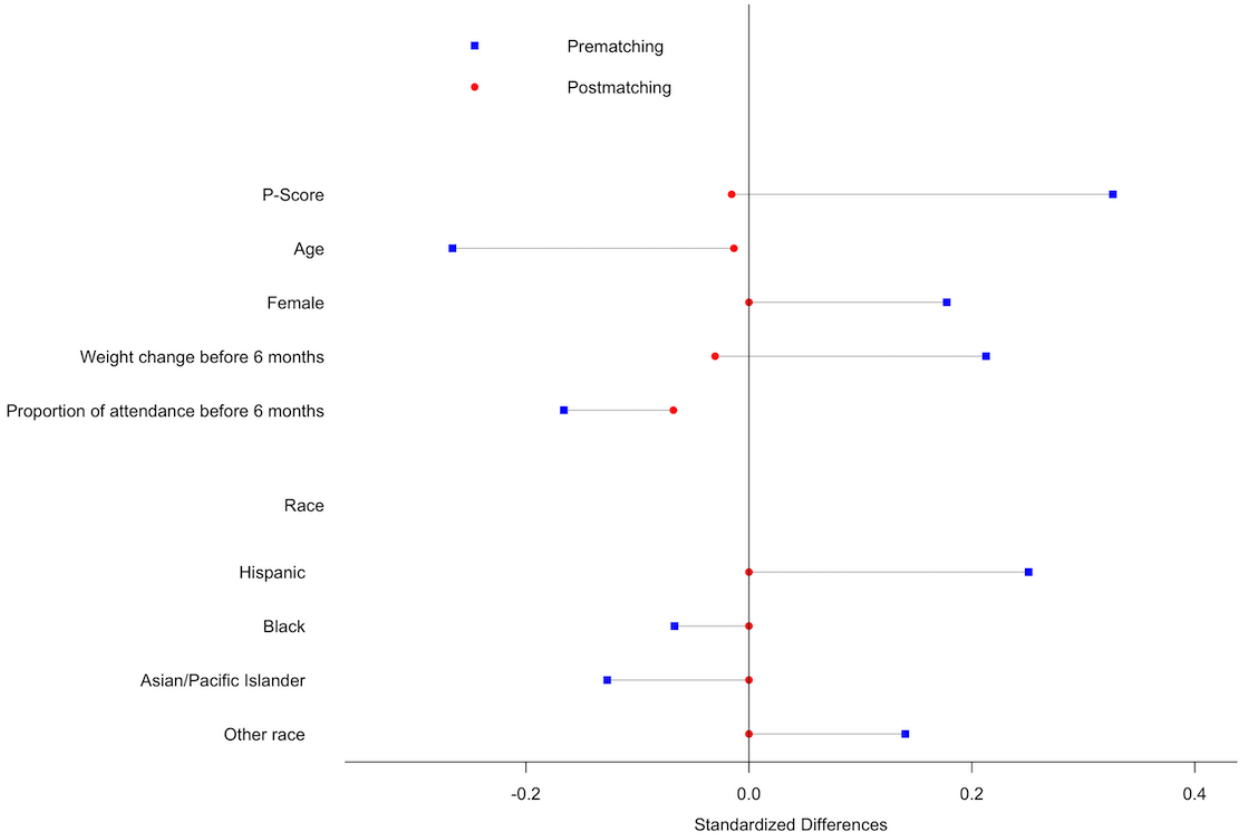


Table 2. Estimated means and mean difference per outcome in the electronic message group versus the control group.

Outcome	Electronic message (n=88), mean (SD)	Control (n=88), mean (SD)	Mean difference	95% CI	*P* value
Class attendance	57.0 (29.6)	52.4 (31.2)	4.6	−4.4 to 13.7	.31
Diet adherence (average g/day at 12 months)					
Low carb (n=82)	133.7 (49.2)[a]	136.2 (63.3)[b]	−2.5	−29.9 to 24.8	.85
Low fat (n=94)	63.6 (23.44)[c]	57.43 (26.8)[d]	6.2	−4.1 to 17.0	.26
Weight change (kg)	1.9 (4.3)	1.6 (3.4)	0.3	−1.0 to 1.5	.68

[a]n=43.
[b]n=39.
[c]n=45.
[d]n=49.

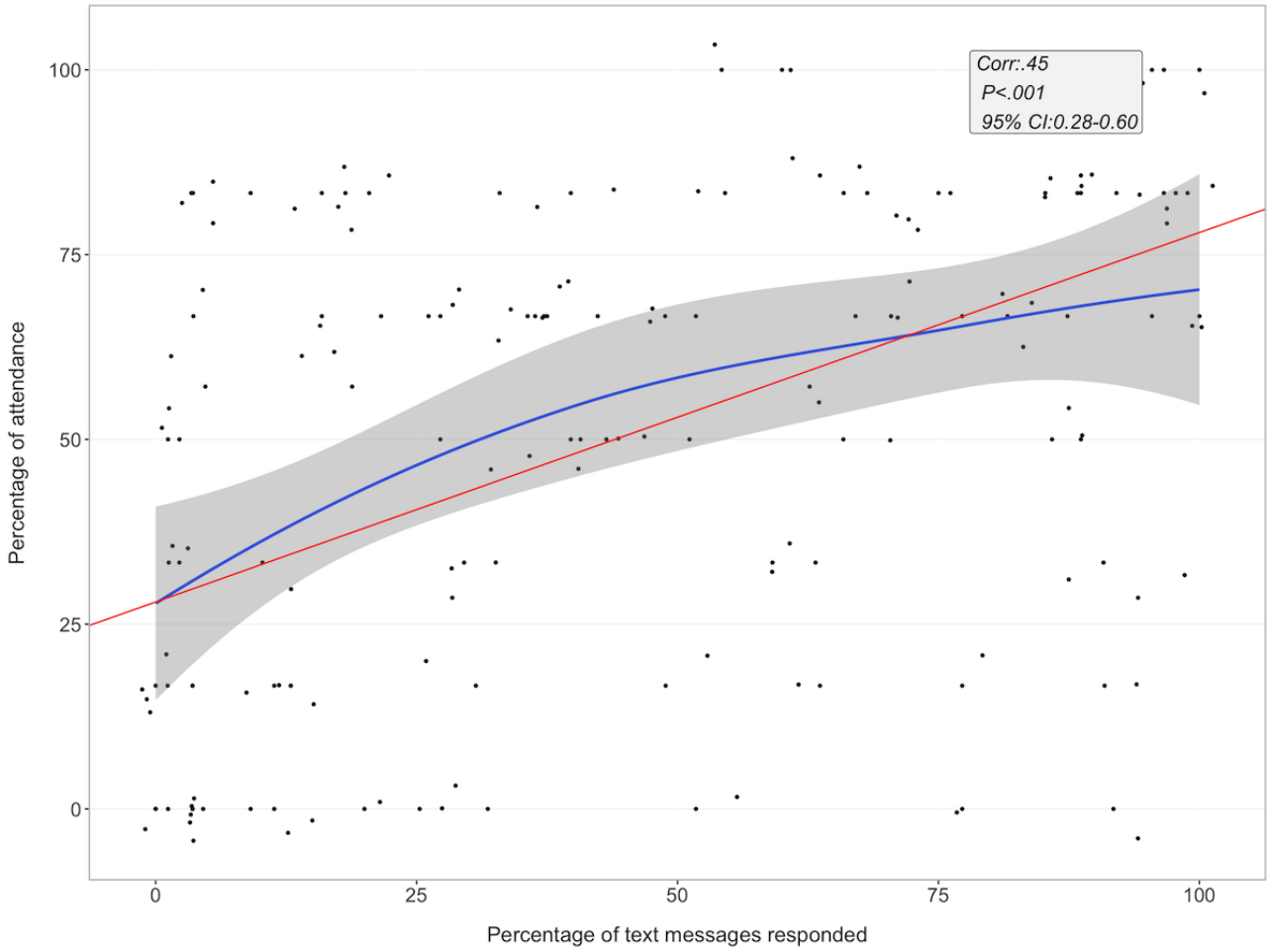
Electronic Message Engagement Within Intervention Participants

E-message response rate changed over time. Overall, 64 out of the 97 participants who received the intervention responded to e-messages during the first 2 weeks, whereas only 29 responded by the last week of the intervention. The average proportion of e-messages responded to out of those received by each participant (6-8 messages per 2-week period) started at 77% and declined to 25% by the end of the 6-month period (see Multimedia Appendix 3).

The percentage of people who responded to at least one message per week began at 94% (63 out of 67 participants) and ended at 38% (37 out of 67 participants; see Multimedia Appendix 4). There was a positive relationship between e-messages and percentage of classes attended (r=.45, P<.001; see Figure 2).

There were no significant relationships between e-message response rate and diet adherence (r=−.03, P=.87) or e-message response rate and weight change (r=−.09, P=.41).

Figure 2. Electronic message response rate and class attendance after 6 months. Blue line denotes loess line fit and red line denotes linear regression line fit. E-message: electronic message; Corr: correlation.



Participant Perceptions of Electronic Messages

Perceived usefulness of the e-message intervention was answered by 58 of the 88 participants. On a 5-point Likert scale of perceived usefulness, 18 of them reported *not at all*; 28, *slightly useful*; 11, *somewhat useful*; 1, *moderately useful*; and 0, *extremely useful*. When asked about the preferred frequency of e-messages, 3 selected 6 to 7 messages/week, 5 chose 4 to 5 messages/week, 15 chose 2 to 3 messages/week, 9 chose 1 text/week, 5 chose not sure, and 21 said the e-messages did not help.

Discussion

Principal Findings

In response to a request from previous participants desiring more accountability for adhering to their diet, a low-cost e-message intervention was offered to the final cohort. Receiving frequent e-message prompts during the second 6-month period of the diet intervention to monitor goal adherence, emotional responses, and behavioral responses to goal discrepancies did not significantly improve class attendance, diet adherence, or weight loss retention compared with matched controls. Within those who did receive the e-messages, there was a positive relationship between overall e-messages response rate and classes attended during the second 6-month period; however, no relationship to diet adherence or weight change was found. One explanation for this association is that response to the e-messages is a proxy of program engagement, reflecting the action of participants already engaged.

One possibility for the null result is insufficient engagement in the e-message intervention. There was a significant reduction in e-message responses over time, with 75% of participants responding to any e-messages by week 5 and less than half of the participants responding to 1 e-message per week by the end of the intervention. Moreover, 18 of the 58 respondents who answered the postintervention questionnaires rated the messages as not useful at all, and when asked how frequently they thought the messages should be sent, 21 said the e-messages did not help. Almost half the participants (40/88) made open-ended suggestions and comments for improvement of the e-messages. Coauthors categorized these responses into broad themes of suggestions. Some of these themes were more variability of questions and possible responses (n=8), reducing the frequency of messages (n=6), more adaptable and accountable e-messages (n=5), messages were “annoying” (n=4), and response choices for “what words would you use to describe your current feelings about your eating plan?” were too ambiguous (n=4). Only 5 of the 40 who gave comments reported that it worked for them. Future e-message interventions may increase engagement by varying the message questions or content, thus decreasing possible message fatigue.

Limitations

The study had several limitations. Although our e-messages elicited a participant response designed to increase the participant's awareness of their diet adherence and internal reaction to it, there was no feedback, tracking of progress, or even automatic reply sent when the participant did respond. To increase participant engagement in the e-message intervention, a 2-way interaction could be provided. A recent review of self-directed weight loss interventions suggested that individualized feedback, email counseling, and online social support seemed to enhance effectiveness [30]. Indeed, studies that have found significant effects on weight loss behaviors used several of these elements as well as customized content around behavioral tips [2,14,31,32]. Our e-message intervention had none of these features and consisted instead of 1-way pushed questionnaires that were designed to make participants more aware of their dietary progress and internal emotional states. It is possible that some of these features from past studies are critical to motivate behavior change. Indeed, Spark et al's study had more *human* elements with experimenter-automated replies using the participant's name and signing the assigned health coach's name [16]. Even if automated, there is evidence that the *mere belief* of social presence enhances arousal and engagement [33]. Another limitation is that although propensity score adjustment for treatment selection bias strengthens the causal interpretations of our findings, this adjustment cannot balance across all possible confounders, only the ones we included. A final limitation is the low power because of our small sample size. The historically matched controls allowed for us to fully use cohort 5 for the intervention population; however, a larger sample size would have had more power to detect effects and would have allowed for more stratified subgroup analyses.

A unique strength of this study was using propensity score–matched historical controls to compare the e-message intervention's effect on outcomes. In addition, the study provides data on potential limitations and what did not work as well as open-ended participant feedback on why. This material can inform the development of other treatment adherence interventions, especially those involving SMS text or email messages. Finally, optimizing content and frequency of *nudge* mobile interventions to promote accountability with minimal experimenter cost is an iterative process that requires multiple studies to identify the optimal method and frequency of participant contact. Using SMS text message offers a broader potential impact on public health across the socioeconomic and geographic spectrum [11].

Conclusions

This e-message intervention did not have a main effect on treatment adherence measured by class attendance or diet adherence; however, it did indicate a measure of engagement, with a relationship between e-message response and class attendance. Despite the limited effect of e-messages, this work significantly contributes to the space of mobile health interventions and hopefully inspires adaptive trial design. Feedback from earlier cohorts requested more accountability, and these messages were a low-cost, automated tool to support this request. By focusing on minimal *nudge-like* interventions that are able to produce the most change with the least effort, it may be possible to generate powerful behavioral public health tools that can be extended to individuals across the socioeconomic spectrum. Those most in need of public health interventions often have the least capability to access resources. Continued efforts to develop these types of innovative interventions may help bridge the gap.

Acknowledgments

Research reported in this publication was supported by the National Heart, Lung, And Blood Institute of the National Institutes of Health under Award Number K01HL136702. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Eating plan adherence questionnaire.

[PNG File, 225KB - mhealth_v7i4e11720_app1.png]

Multimedia Appendix 2

Outcomes by method of participation.

[PNG File, 369KB - mhealth_v7i4e11720_app2.png]

Multimedia Appendix 3

Mean proportion electronic message response over time.

[PNG File, 49KB - mhealth_v7i4e11720_app3.png]

Multimedia Appendix 4

Percentage participants with greater than 1 response/week.

[PNG File, 41KB - mhealth_v7i4e11720_app4.png]

Multimedia Appendix 5

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 5MB - mhealth_v7i4e11720_app5.pdf]

References

1. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol 2004 Sep;23(5):443-451. [doi: 10.1037/0278-6133.23.5.443] [Medline: 15367063]
2. Shaw RJ, Bosworth HB, Silva SS, Lipkus IM, Davis LL, Sha RS, et al. Mobile health messages help sustain recent weight loss. Am J Med 2013 Nov;126(11):1002-1009 [FREE Full text] [doi: 10.1016/j.amjmed.2013.07.001] [Medline: 24050486]
3. Dansinger ML, Gleason JA, Griffith JL, Selker HP, Schaefer EJ. Comparison of the Atkins, Ornish, Weight Watchers, and Zone diets for weight loss and heart disease risk reduction: a randomized trial. J Am Med Assoc 2005 Jan 5;293(1):43-53. [doi: 10.1001/jama.293.1.43] [Medline: 15632335]
4. Wadden TA, Berkowitz RI, Womble LG, Sarwer DB, Phelan S, Cato RK, et al. Randomized trial of lifestyle modification and pharmacotherapy for obesity. N Engl J Med 2005 Nov 17;353(20):2111-2120. [doi: 10.1056/NEJMoa050156] [Medline: 16291981]
5. Burke LE, Sereika SM, Music E, Warziski M, Styn MA, Stone A. Using instrumented paper diaries to document self-monitoring patterns in weight loss. Contemp Clin Trials 2008 Mar;29(2):182-193 [FREE Full text] [doi: 10.1016/j.cct.2007.07.004] [Medline: 17702667]
6. Warziski MT, Sereika SM, Styn MA, Music E, Burke LE. Changes in self-efficacy and dietary adherence: the impact on weight loss in the PREFER study. J Behav Med 2008 Feb;31(1):81-92. [doi: 10.1007/s10865-007-9135-2] [Medline: 17963038]
7. Gardner CD, Trepanowski JF, Del Gobbo LC, Hauser ME, Rigdon J, Ioannidis JPA, et al. Effect of low-fat vs low-carbohydrate diet on 12-month weight loss in overweight adults and the association with genotype pattern or insulin secretion: the DIETFITS randomized clinical trial. J Am Med Assoc 2018 Dec 20;319(7):667-679. [doi: 10.1001/jama.2018.0245] [Medline: 29466592]
8. Greenway FL. Physiological adaptations to weight loss and factors favouring weight regain. Int J Obes (Lond) 2015 Aug;39(8):1188-1196 [FREE Full text] [doi: 10.1038/ijo.2015.59] [Medline: 25896063]
9. MacLean PS, Higgins JA, Giles ED, Sherk VD, Jackman MR. The role for adipose tissue in weight regain after weight loss. Obes Rev 2015 Feb;16(Suppl 1):45-54 [FREE Full text] [doi: 10.1111/obr.12255] [Medline: 25614203]
10. Artinian NT, Fletcher GF, Mozaffarian D, Kris-Etherton P, Van Horn L, Lichtenstein AH, American Heart Association Prevention Committee of the Council on Cardiovascular Nursing. Interventions to promote physical activity and dietary lifestyle changes for cardiovascular risk factor reduction in adults: a scientific statement from the American Heart Association. Circulation 2010 Jul 27;122(4):406-441 [FREE Full text] [doi: 10.1161/CIR.0b013e3181e8edf1] [Medline: 20625115]
11. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
12. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009 May 14;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
13. Shapiro JR, Koro T, Doran N, Thompson S, Sallis JF, Calfas K, et al. Text4Diet: a randomized controlled study using text messaging for weight loss behaviors. Prev Med 2012 Nov;55(5):412-417. [doi: 10.1016/j.ypmed.2012.08.011] [Medline: 22944150]
14. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009;11(1):e1 [FREE Full text] [doi: 10.2196/jmir.1100] [Medline: 19141433]
15. Donaldson EL, Fallows S, Morris M. A text message based weight management intervention for overweight adults. J Hum Nutr Diet 2014 Apr;27(Suppl 2):90-97. [doi: 10.1111/jhn.12096] [Medline: 23738786]
16. Spark LC, Fjeldsoe BS, Eakin EG, Reeves MM. Efficacy of a text message-delivered extended contact intervention on maintenance of weight loss, physical activity, and dietary behavior change. JMIR Mhealth Uhealth 2015;3(3):e88 [FREE Full text] [doi: 10.2196/mhealth.4114] [Medline: 26373696]
17. Stanton MV, Robinson JL, Kirkpatrick SM, Farzinkhou S, Avery EC, Rigdon J, et al. DIETFITS study (diet intervention examining the factors interacting with treatment success) - study design and methods. Contemp Clin Trials 2017 Dec;53:151-161 [FREE Full text] [doi: 10.1016/j.cct.2016.12.021] [Medline: 28027950]

18. Thaler R, Sunstein C. Nudge: Improving Decisions about Health, Wealth, and Happiness. Cambridge: Yale University Press; 2008.
19. Inzlicht M, Legault L, Teper R. Exploring the mechanisms of self-control improvement. Curr Dir Psychol Sci 2014 Aug 4;23(4):302-307. [doi: 10.1177/0963721414534256] [Medline: 22973420]
20. Polivy J, Herman CP, Deo R. Getting a bigger slice of the pie. Effects on eating and emotion in restrained and unrestrained eaters. Appetite 2010 Dec;55(3):426-430. [doi: 10.1016/j.appet.2010.07.015] [Medline: 20691231]
21. Moon A, Berenbaum H. Emotional awareness and emotional eating. Cogn Emot 2009 Apr;23(3):417-429. [doi: 10.1080/02699930801961798]
22. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
23. Thompson FE, Kirkpatrick SI, Subar AF, Reedy J, Schap TE, Wilson MM, et al. The National Cancer Institute's dietary assessment primer: a resource for diet research. J Acad Nutr Diet 2015 Dec;115(12):1986-1995. [doi: 10.1016/j.jand.2015.08.016]
24. Baumel A, Tinkelman A, Mathur N, Kane JM. Digital peer-support platform (7Cups) as an adjunct treatment for women with postpartum depression: feasibility, acceptability, and preliminary efficacy study. JMIR Mhealth Uhealth 2018 Feb 13;6(2):e38 [FREE Full text] [doi: 10.2196/mhealth.9482] [Medline: 29439944]
25. Sohn M, Jung M. Effect of viewing smoking scenes in motion pictures on subsequent smoking desire in audiences in south Korea. JMIR Public Health Surveill 2017 Jul 17;3(3):e46 [FREE Full text] [doi: 10.2196/publichealth.7093] [Medline: 28716768]
26. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983 Apr;70(1):41. [doi: 10.2307/2335942]
27. Hansen B, Klopfer S. Optimal full matching and related designs via network flows. J Comput Graph Stat 2006 Sep;15(3):609-627. [doi: 10.1198/106186006X137047]
28. The R Project for Statistical Computing. URL: https://www.r-project.org/ [accessed 2019-03-05] [WebCite Cache ID 76dfFVzWy]
29. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Lawrence Erlbaum Associates; 1988.
30. Tang JC, Abraham C, Greaves CJ, Yates T. Self-directed interventions to promote weight loss: a systematic review of reviews. J Med Internet Res 2014 Feb 19;16(2):e58 [FREE Full text] [doi: 10.2196/jmir.2857] [Medline: 24554464]
31. Fischer HH, Fischer IP, Pereira RI, Furniss AL, Rozwadowski JM, Moore SL, et al. Text message support for weight loss in patients with prediabetes: a randomized clinical trial. Diabetes Care 2016 Dec;39(8):1364-1370. [doi: 10.2337/dc15-2137] [Medline: 26861922]
32. McCoy P, Leggett S, Bhuiyan A, Brown D, Frye P, Williams B. Text messaging: An intervention to increase physical activity among african american participants in a faith-based, competitive weight loss program. Int J Environ Res Public Health 2017 Dec 29;14(4) [FREE Full text] [doi: 10.3390/ijerph14040326] [Medline: 28353650]
33. Okita S, Bailenson J, Schwartz D. Mere Belief in Social Action Improves Complex Learning. In: Proceedings of the 8th International Conference on International Conference for the Learning Sciences. 2008 Presented at: ICLS'08; June 24-28, 2008; Utrecht, The Netherlands p. 132-139.

Abbreviations

e-message: electronic message
DIETFITS: Diet Intervention Examining The Factors Interacting with Treatment Success
LC: low-carbohydrate
LF: low-fat
RCT: randomized controlled trial
SMS: short message service

Edited by G Eysenbach; submitted 30.07.18; peer-reviewed by G Papandonatos, A McKenzie, M Amato; comments to author 26.09.18; revised version received 19.11.18; accepted 22.11.18; published 09.04.19.

Please cite as:
Oppezzo MA, Stanton MV, Garcia A, Rigdon J, Berman JR, Gardner CD
To Text or Not to Text: Electronic Message Intervention to Improve Treatment Adherence Versus Matched Historical Controls
JMIR Mhealth Uhealth 2019;7(4):e11720
URL: https://mhealth.jmir.org/2019/4/e11720/
doi:10.2196/11720
PMID:30964436

Original Paper

Use of the Smartphone App WhatsApp as an E-Learning Method for Medical Residents: Multicenter Controlled Randomized Trial

Thomas Clavier[1*], MD, MSc; Julie Ramen[1*], MD; Bertrand Dureuil[1], MD, PhD; Benoit Veber[1], MD, PhD; Jean-Luc Hanouz[2], MD, PhD; Hervé Dupont[3], MD, PhD; Gilles Lebuffe[4], MD, PhD; Emmanuel Besnier[1], MD, MSc; Vincent Compere[1], MD, PhD

[1]Department of Anesthesiology and Critical Care, Rouen University Hospital, Rouen, France

[2]Department of Anesthesiology and Critical Care, Caen University Hospital, Normandie Université, Caen, France

[3]Department of Anesthesiology and Critical Care, Amiens University Hospital, Amiens, France

[4]Pôle d'Anesthésie-Réanimation, Lille University Hospital, Lille, France

[*]these authors contributed equally

Corresponding Author:
Thomas Clavier, MD, MSc
Department of Anesthesiology and Critical Care
Rouen University Hospital
1 rue de Germont
Rouen, 76000
France
Phone: 33 232888990
Email: thomasclavier76@gmail.com

Abstract

Background: The WhatsApp smartphone app is the most widely used instant messaging app in the world. Recent studies reported the use of WhatsApp for educational purposes, but there is no prospective study comparing WhatsApp's pedagogical effectiveness to that of any other teaching modality.

Objective: The main objective of this study was to measure the impact of a learning program via WhatsApp on clinical reasoning in medical residents.

Methods: This prospective, randomized, multicenter study was conducted among first- and second-year anesthesiology residents (offline recruitment) from four university hospitals in France. Residents were randomized in two groups of online teaching (WhatsApp and control). The WhatsApp group benefited from daily delivery of teaching documents on the WhatsApp app and a weekly clinical case supervised by a senior physician. In the control group, residents had access to the same documents via a traditional computer electronic learning (e-learning) platform. Medical reasoning was self-assessed online by a script concordance test (SCT; primary parameter), and medical knowledge was assessed using multiple-choice questions (MCQs). The residents also completed an online satisfaction questionnaire.

Results: In this study, 62 residents were randomized (32 to the WhatsApp group and 30 to the control group) and 22 residents in each group answered the online final evaluation. We found a difference between the WhatsApp and control groups for SCTs (60% [SD 9%] vs 68% [SD 11%]; *P*=.006) but no difference for MCQs (18/30 [SD 4] vs 16/30 [SD 4]; *P*=.22). Concerning satisfaction, there was a better global satisfaction rate in the WhatsApp group than in the control group (8/10 [interquartile range 8-9] vs 8/10 [interquartile range 8-8]; *P*=.049).

Conclusions: Compared to traditional e-learning, the use of WhatsApp for teaching residents was associated with worse clinical reasoning despite better global appreciation. The use of WhatsApp probably contributes to the dispersion of attention linked to the use of the smartphone. The impact of smartphones on clinical reasoning should be studied further.

(JMIR Mhealth Uhealth 2019;7(4):e12825) doi:10.2196/12825

KEYWORDS

education, medical, graduate/methods; educational measurement; anesthesiology; internship and residency; trauma; hemorrhage; mobile applications; WhatsApp; smartphone; teaching materials; mobile phone

Introduction

Many computer-based teaching materials have been developed in recent years, and electronic learning (e-learning) is becoming increasingly popular in medical schools, with the appearance of guides on e-learning deployment [1]. E-learning has many organizational advantages over face-to-face teaching: temporal and spatial flexibility for learners, live updates, and easy and uniform dissemination of teaching resources for teachers. Moreover, the emergence of social networks facilitates personal and professional communication and exchange [2,3]. The time spent on mobile phone screens per day ("screen time") has increased exponentially since the introduction of the latest-generation phones known as smartphones, in particular, among young people, leading to a growing interest of mobile learning (m-learning) among teachers [1,4-6].

The WhatsApp smartphone app, developed by WhatsApp Inc (owned by Facebook Inc, Menlo Park, CA), is the most widely used instant messaging app in the world, with more than one billion active users per month and more than 40 billion WhatsApp messages exchanged each day in 2016 [7]. It allows communication between group participants without the need for unity in place or time. Participants are free to choose when they want to access the information posted and can view and interact with other group members regarding the information delivered at any time. In view of its popularity with medical students, it seems interesting to envisage a new use of WhatsApp, by orienting it toward an educational objective (with the opportunity to recover screen time from students) [8,9]. The first reports of the use of this app for educational purposes date to the early 2017, for teaching medical students or training pathology residents [10,11]. Both of these observational studies showed satisfaction among WhatsApp participants and highlighted the ease of use and the quick access to lessons through the app. However, there is no prospective study comparing WhatsApp to any other teaching modality.

Residents involved in tutored practice exchange groups have better medical reasoning, as evaluated by the script concordance test (SCT), which is a well-validated medical reasoning–assessment tool for residents [12-14]. Similar to practice exchange groups, WhatsApp allows direct communication between teachers and students with the possibility of discussing real clinical cases and commenting on residents' management of the case. Thus, we hypothesized that WhatsApp could have the same effect as practice exchange groups on clinical reasoning. The main objective of this study was to measure the impact of WhatsApp on clinical reasoning by using the SCT. As severe trauma is one of the leading causes of death in the world, with more than 5 million deaths, and posttraumatic hemorrhage is the leading cause of mortality, we selected posttraumatic hemorrhage management as the topic for our teaching and evaluation [15].

Methods

Population Selection

This prospective, randomized, unblinded, multicenter study was conducted among first- and second-year anesthesiology residents from four French university hospitals with trauma center (Amiens, Caen, Lille, and Rouen) comparing WhatsApp to a control group. Since computer-based e-learning did not show any noninferiority compared to traditional teaching, it was chosen as a control teaching platform for this study [16,17]. The Ethics and Evaluation Committee for Non-Interventional Research of Rouen University Hospital approved the study (E2017-37). All participants received information before any study procedures were undertaken, and residents were invited to participate as subjects in the study. All information about and during the trial was sent by email, and participants knew that the WhatsApp group was the "intervention" group. Agreement to participate was provided online by email or telephone by each resident who could stop participating at any time.

The inclusion criteria were ongoing medical residency, possession of a mobile phone that could download the WhatsApp app, and attestation for agreement to use WhatsApp for this study. Noninclusion criteria were refusal to participate, noncompatibility of a mobile phone to download WhatsApp, and failure to download WhatsApp. This study was carried out in addition to the official teaching program of the residents and was not integrated into usual teaching nor did it replace previous teaching.

The primary measure was medical reasoning evaluated by the SCT. The secondary parameters were medical knowledge measured by multiple-choice questions (MCQ); feasibility and acceptability of using WhatsApp, assessed by collecting resident testimonials; the Cronbach coefficient alpha, calculated after optimizing the test for SCT [18]; and self-assessment (by quantitative and semiquantitative numerical scales using a satisfaction questionnaire) of time spent working, quality of teaching, global satisfaction, teachers' availability, impression that the teaching met the learning objectives, relevance of clinical cases, and volume of teaching documents used.

Study Procedures

After inclusion, residents were randomized with their last names in two groups (WhatsApp and control) by author TC, using an online open-access app for stratification according to the student's hospital and year of residency [19]. Concerning intervention or evaluation, this was a purely app- or Web-based trial without face-to-face components between resident and teachers. After randomization and before the beginning of the course, all residents were emailed a short evaluation with 10 SCTs and 10 MCQs on basic knowledge of anesthesiology (intensive care, regional anesthesia, obstetrics, etc) to check the initial comparability of the groups and to familiarize first-year residents with the SCT. Students returned the SCTs and MCQs by email after completing them. After this short evaluation, the WhatsApp group benefited from daily delivery of teaching documents specially prepared for easy readability on a smartphone (from Monday to Thursday, morning and afternoon; 2-4 documents/day; Multimedia Appendix 1) through the WhatsApp app. These documents were inspired by the most recent guidelines on the management of traumatic hemorrhagic shock and were validated by anesthesiology teachers (VC, BV, and BD) [20,21]. It is strongly suggested that resolution of

clinical cases have a significant role in the acquisition of medical reasoning [22,23]. Thus, every Friday, residents were given a "step-by-step" clinical case on WhatsApp for 3-4 hours (Multimedia Appendix 2), supervised by a senior anesthesiologist (TC) who questioned the residents (to create an interest in the clinical cases) and provided them with feedback and validation or correction, if necessary, as described in the practice exchange groups [12]. Several screenshots of the use of WhatsApp for learning purposes during the protocol are presented in the Multimedia Appendix 3. The total duration of teaching was 3 weeks, and the choice of the length of the teaching period was based on both the availability of teachers and the estimated acceptability of students. In the control group, residents had access to the same documents via a computer e-learning platform, and the senior anesthesiologist teacher was available by email. They had access to the three clinical cases with their answers but had no live interaction with a teacher. The two groups had the same program and learning objectives. Participants did not receive any documents during the weekends and were free to stop the courses at any time. The characteristics of WhatsApp-assisted m-learning and traditional e-learning used in this study are summarized in Table 1. At the end of the teaching period, the two groups had the same formative evaluation by 29 SCTs and 30 MCQs sent by email and completed during the month following the end of the teaching period (Multimedia Appendices 4 and 5). Students returned the SCTs and MCQs by email after completing them. In case of nonresponse, residents were sent two reminders by email before being considered lost to follow-up. The residents of the two groups who responded to the final evaluation completed an online satisfaction questionnaire specifically created for this study (not previously validated in the literature; Multimedia Appendix 6).

Table 1. Characteristics of WhatsApp-assisted m-learning and control e-learning used in this study. The two groups had the same program, learning objectives, and educational documents.

Characteristics	WhatsApp group (m-learning[a])	Control group (traditional e-learning[b])
Length of teaching	3 weeks	3 weeks
Accessibility of educational documents	Sent daily on WhatsApp from Monday to Friday	Available on a computer e-learning platform
Teacher availability	Available and can be contacted by WhatsApp	Available and can be contacted by email
Conduct of clinical cases	Live on Friday on WhatsApp, with questions and answers from the teacher as the case progresses	Cases accessible on the platform with their answers. Teacher available if the student has any questions.

[a]m-learning: mobile learning.

[b]e-learning: electronic learning.

Design of the Script Concordance Test and Multiple-Choice Questions

The MCQs and SCTs were written by one of the teachers (TC). They were directly related to issues covered during teaching and were reviewed (and possibly modified, if needed) by two other teachers (JR and VC). The SCTs were designed as previously described [12,24]. The SCT confronted the residents with authentic uncertain clinical situations concerning traumatic hemorrhagic shock, which were described in vignettes, each corresponding to one of the previously set objectives. The clinical situations were problematic even for experienced clinicians, either because there were not enough data or the situations were ambiguous. There were several options for diagnosis, investigation, or treatment. The items (questions) were based on a panel of questions that an experienced clinician would consider relevant to this type of clinical setting. The item was consistent with the presentation of relevant options and new data (not described in the vignette). The task for the student was to determine the effect these new data on the status of the option. The resident's task was to assess, using a 5-point Likert scale, the influence of this new element on the diagnostic hypothesis, the plan for investigation, or the treatment. The different points on the scale corresponded to positive values (the option was enhanced by the new data), neutral values (the data did not change the status of the option), or negative values (this option was ruled out by the data). The scoring system was based on the principle that any answer given by one expert had an intrinsic value, even if that answer did not coincide with those of other experts. In the present study, a group of 13 anesthesiologist practitioners regularly involved in the management of traumatic hemorrhagic shock formed the expert panel. The principles of SCT are that for each item, the answer entitled the resident to a credit corresponding to the number of experts who had chosen it. All items had the same maximum credit, and raw scores were transformed proportionally to obtain a one-point credit for the answer that was chosen by most experts. Other choices received a partial credit. Thus, to calculate the scores, all results were divided by the number of individuals who had given answers chosen by the largest number of respondents. The total score for the test was the sum of all credits earned for each item. The total score was then transformed into a percentage score. An automatic correction software (freely accessible on the website of the University of Montreal) was used for scoring [25]. Each MCQ was worth one point, and it was possible for an MCQ to have several correct answers. To obtain a point for an MCQ, the resident had to tick all the correct answers and none of the incorrect ones. Otherwise, the student did not receive any points. The final rating was based on the total number of proposed MCQs.

Statistical Analysis

With regard to our previous publication on the use of SCT by anesthesiology residents, we assumed that a difference of 6% between the two groups would be clinically significant [12,24]. Based on these findings, assuming that the SD was the same

between the populations and using a power of 0.90 with a level of statistical significance at .05, it was estimated that 22 students should be analyzed in each group. A randomized study on e-learning showed that about a quarter of the students included do not participate or are lost to follow-up [26]. Based on these findings, it was estimated that a minimum of 28 students should be included in each group to be able to analyze 22 students.

The values are presented as number and percentage values for qualitative variables, as mean and SDs for quantitative variables with a normal distribution, and as median and interquartile range for quantitative variables with a non-normal distribution. Residents who did not respond to the final evaluation were excluded from the final analysis (lack of analyzable parameters). After performing a Shapiro-Wilk normality test, the quantitative variables were compared using a Student *t* test (if the distribution was normal) or a Mann-Whitney test (if the distribution was not normal). The qualitative variables were analyzed using a Fischer or a chi-square test. The significance threshold was set at .05. All statistics were analyzed using GraphPad PRISM software (v 5.0; GraphPad Software Inc, San Diego, CA).

Results

Residents' Characteristics

Among 142 eligible anesthesiology residents, 62 (44%) agreed to participate and were randomized as follows: 32 to the WhatsApp group and 30 to the control group. Their main characteristics are summarized in Table 2. Two students randomized to the WhatsApp group were excluded after randomization. The first withdrew from the study for personal reasons, and the second was excluded following failure to download WhatsApp.

Results of the Script Concordance Tests and Multiple-Choice Questions

The lessons took place from March 12 to 30, 2018. For final evaluation, SCTs including 12 scenarios for a total of 36 items were submitted to a panel of 13 experts. Thereafter, 7 items of the SCT were excluded (not enough variability in replies), leaving 29 items of SCT spread over 12 clinical situations. According to the recommendations of Lubarsky et al, we optimized SCT by performing a post-hoc analysis [18]. Items with high variability, low variability, or binomial responses were excluded. We obtained a final version with 10 scenarios and 24 items. After this optimization, Cronbach coefficient alpha was .55. In the WhatsApp group, 20 residents answered the preliminary evaluation, 1 resident who responded to the preliminary evaluation did not answer the final evaluation, and 3 residents who did not respond to the preliminary evaluation answered the final evaluation. In the control group, 22 residents answered the preliminary evaluation, 1 resident who responded to the preliminary evaluation did not answer the final evaluation, and 1 resident who did not answer the preliminary evaluation answered the final evaluation. There was no demographic disparity between the residents who answered and those who did not answer the final evaluation. Their main characteristics are summarized in Tables 3 and 4. The flow chart of the study is presented in Figure 1.

On the preliminary evaluation (before teaching), there was no significant difference between the WhatsApp and control groups for SCT (64% [SD 7%] vs 62% [SD 6%]; *P*=.41) or MCQ (8/10 [SD 1] vs 7/10 [SD 2]; *P*=.33), showing no difference in clinical reasoning or medical knowledge. For the final evaluation (after teaching), we found a significant difference between the WhatsApp and control groups for SCT (60% [SD 9%] vs 68% [SD 11%]; *P*=.006) but not for MCQs (18/30 [SD 4] vs 16/30 [SD 4]; *P*=.22). In the WhatsApp group, there was no difference in the SCT between the initial evaluation and the final evaluation (*P*=.14). In the control group, the SCT scores of the final evaluation were significantly higher than those of the initial evaluation (*P*=.02).

Table 2. Demographic characteristics of the residents.

Characteristic	Control group (n=30), n (%)	WhatsApp group (n=32), n (%)
Year of residency		
First	14 (47)	14 (44)
Second	16 (53)	18 (56)
Sex		
Male	17 (57)	23 (72)
Female	13 (43)	9 (28)
University hospital		
Rouen	10 (33)	11 (34)
Lille	10 (33)	11 (34)
Caen	6 (20)	5 (16)
Amiens	4 (14)	5 (16)

Table 3. Demographic characteristics of residents who answered the final evaluation.

Characteristic	Control group (n=22), n (%)	WhatsApp group (n=22), n (%)
Year of residency		
First	9 (41)	9 (41)
Second	13 (59)	13 (59)
Sex		
Male	14 (64)	16 (73)
Female	8 (36)	6 (27)
University hospital		
Rouen	10 (45)	10 (45)
Lille	6 (27)	5 (23)
Caen	5 (23)	4 (18)
Amiens	1 (5)	2 (10)

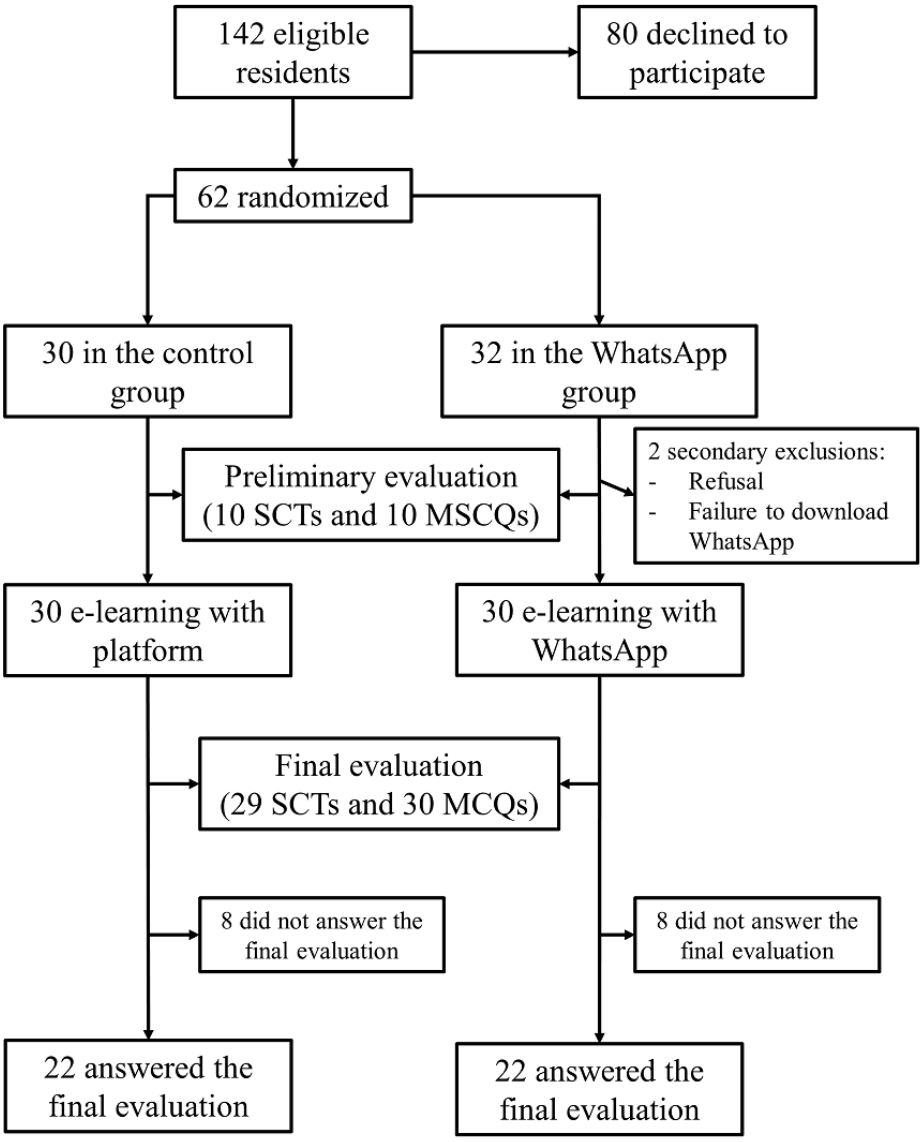
Table 4. Demographic characteristics of residents who did not answer the final evaluation.

Characteristic	Control group (n=8), n (%)	WhatsApp group (n=8), n (%)
Year of residency		
First	5 (63)	3 (37)
Second	3 (37)	5 (63)
Sex		
Male	3 (37)	5 (63)
Female	5 (63)	3 (37)
University hospital		
Rouen	0 (0)	0 (0)
Lille	4 (50)	4 (50)
Caen	1 (13)	1 (13)
Amiens	3 (37)	3 (37)

Use of WhatsApp and Residents' Satisfaction

The residents of the two groups who filled the final evaluation were asked to fill an online satisfaction questionnaire. Twenty (67%) residents in the WhatsApp group and 13 (43%) residents in the control group answered this questionnaire. All the scores from the satisfaction evaluation had a non-normal distribution. There was a difference between the WhatsApp and control groups, with the WhatsApp group showing a better global satisfaction rate (8/10 [interquartile range, 8-9] vs 8/10 [interquartile range 8-8]; *P*=.049), a better feeling that the lessons met the learning objectives (10/10 [interquartile range 8-10] vs 8/10 [interquartile range 7-10]; *P*=.03), and a feeling that the teachers were more available (10/10 [interquartile range 9-10] vs 9/10 [interquartile range 8-10]; *P*=.007). We found no differences between the WhatsApp and control groups in terms of the perceived quality of educational materials (9/10 [interquartile range 8-10] vs 8/10 [interquartile range 8-10]; *P*=.15), the usefulness and relevance of clinical cases (10/10 [interquartile range 8-10] vs 9/10 [interquartile range 7-10]; *P*=.40), the quantity of teaching documents used by the residents (in the WhatsApp group, 14 residents [70%] used more than 50% of the documents and 6 [30%] used less than 50% of the documents; in the control group, 10 residents [77%] used more than 50% of the documents and 3 [23%] used less than 50% of the documents; *P*=.66), or the time spent working on the program (in the WhatsApp group, 2 residents [10%] spent between 5 h and 10 h and 18 [90%] spent between 1 h and 5 h; in the control group, 4 residents [31%] spent between 5 h and 10 h and 9 [69%] spent between 1 h and 5 h; *P*=.18). Textbox 1 presents quotes from the free comments section of the satisfaction questionnaire of the WhatsApp group.

Figure 1. Flowchart of the study. e-learning: electronic learning; MCQ: multiple-choice question; SCT: script concordance test.



Textbox 1. Quotes from the free comments section of the satisfaction questionnaire (WhatsApp group residents).

"For participation to the Friday clinical case it depends on the availability of everyone. The fact that it is on whatsapp makes it easier to communicate and ask questions. Having notifications is more motivating to consult documents than on a platform."

"Having what's app notifications allows me to be more assiduous, the possibility to ask questions directly in the conversation is a big advantage, it sometimes allows small discussions, so very useful. Great classes, interesting cases, and not feeling evaluated is fun. Suggestion: a new session."

"The documents were very well done, difficult on Friday to answer all the questions of the clinical cases online according to our occupations in the ward."

"Very good idea to teach via WhatsApp, which allows to be informed quickly of the presence of new educational documents and to have regular reminder shots since the notifications are displayed. Doing clinical cases on the application during a day with the participation of several people is very instructive. The only problem is that the documents are difficult to consult on a small telephone, perhaps it would be necessary to adapt the documents in the form of slides format telephone. Otherwise it was great! High quality educational documents. Thank you!"

"Interesting to be able to consult documents via whatsapp. As far as Friday clinical cases are concerned, it is quite difficult to switch between ward presence or other obligations and whatsapp."

"Very nice project. I think it's useful to have cards per whatsapp but the flow was too high: 4 documents per day, we end up having too much delay in the readings."

"The idea is good but it's quite laborious to read lessons on a mobile phone, especially long pdf. The well ventilated and clear synthetic dcouments [sic] are on the other hand interesting. It is also interesting to be able to ask questions directly and get quick answers. But it can't replace classical education."

Discussion

Principal Results

This randomized, multicenter study is the first to focus on the impact of WhatsApp on clinical reasoning in medical students. We found that the use of WhatsApp, instead of a traditional e-learning platform, to teach a specific topic was associated with worse clinical reasoning despite better global appreciation.

Comparison With Prior Work

Several recent studies have reported the potential interest of specific smartphone apps in medical education, but our objective was to assess the interest of a very widely used nonmedical app (thus easily usable by all) for teaching [27-29]. Given that WhatsApp allows interaction between teachers and students, with the possibility of discussing clinical cases, we believed its use would improve medical reasoning, as previously described for face-to-face practice exchange groups [12]. We did not find any difference in the global amount of work or the number of educational documents consulted, which is consistent with similar personal work between the two groups. It has been shown that e-learning methods improve the medical knowledge of health care professionals [28]. The absence of a difference in the medical knowledge assessed by MCQs shows that the weakness of clinical reasoning related to WhatsApp is not related to less knowledge of the subject. We can therefore assume that this decrease in the quality of reasoning is directly related to WhatsApp or the use of a smartphone. It is likely that reading on WhatsApp between two other activities was less effective than time spent solely on an e-learning platform. A recent study showed that a smartphone app dedicated to teaching medical students Dermatology, in combination with traditional teaching, improved medical knowledge measured by MCQs [27]. Although we did not find any improvement in medical knowledge in our work, the smartphone was seen as an alternative to conventional e-learning and not as a complement. It is interesting to note that in the literature, most of the educational benefits reported with smartphone use stem from very “visual” specialties (Dermatology or Pathology) and that this tool, which allows easy communication of iconography, is probably more relevant in this context than in “less visual” medical specialties [10,27].

Residents pointed out two limitations: the difficulty of participating in clinical cases on Friday in parallel with their usual activities and the difficulty in referring to documents on small smartphone screens. Unlike for practice exchange groups, there was no time dedicated specifically to clinical case resolution on WhatsApp, and residents had to respond in addition to their usual activities [12]. This probably favored a multitasking activity with a difficulty to focus on the pedagogical content. However, it is interesting to note that the comments from WhatsApp residents were very positive, with a higher overall satisfaction rating. The novelty and originality of the concept probably contributed to this satisfaction, but it underlines the fact that the students were not aware of the possible negative impact of the use of WhatsApp. A recent randomized pedagogic study assessed the impact of learning modules using m-learning on knowledge gain, skill gain, and satisfaction for otorhinolaryngology and head and neck surgery disorders in undergraduate medical students [28]. Despite the absence of differences in knowledge gain in the mobile interactive multimedia group, satisfaction was higher in the mobile group (like in our cohort). Therefore, we can assume that our data are concordant with the literature on m-learning.

Finally, the daily use of WhatsApp for medical education probably contributes to the dispersion of attention linked to the use of the smartphone. In view of these results, it does not seem justified to continue to develop WhatsApp for teaching medical reasoning to medical residents. However, the targeted use of WhatsApp with other educational objectives (eg, medical imaging or video) remains to be evaluated and should be the subject of future randomized studies. It is known that blended learning can have a beneficial effect on knowledge acquisition in health professions [30]. Thus, it might also be interesting to study the use of WhatsApp as a complement to another form of teaching. Given the increasing use of smartphones by health workers, it also seems appropriate to consider future work to assess the quality of clinical reasoning between two populations of physicians with or without usual smartphone use in hospitals.

Limitations

Our study has several limitations. First, the Cronbach coefficient alpha in our SCT evaluation was low. The minimum coefficient usually retained for normative evaluations is 0.7, but in our work the evaluation was only formative and integrated into teaching. The limited number of SCTs probably explains this low coefficient. However, teaching in a specific and specialized area made it difficult to find at least 60 SCTs (as is usually recommended) without redundancy [18]. Second, residents’ participation in our work was limited: Only 62 of 142 residents participated. As previously observed, self-training with e-learning is impacted by a significant dropout rate [26]. In our work, only 22 of the 30 residents participated in the final evaluation. As this teaching was optional, participation in our study represented additional personal work for the residents. It is therefore possible that the majority of residents were discouraged by this prospect. In addition, residents without smartphones or those who did not wish to use WhatsApp logically refused to participate. Third, we could not prevent cross-communication among students while they answered the SCTs and MCQs, and the residents could have communicated with each other during the final evaluation. The fact that this evaluation was not sanctioned and had no value, as it was not integrated into usual teaching methods, probably limited this communication. Finally, we did not use a prevalidated questionnaire to measure satisfaction. As we wanted to evaluate specific points related to the use of WhatsApp in our population, we created a new dedicated questionnaire, but this choice made it more difficult to compare our satisfaction results to those of others.

Conclusions

Compared to traditional e-learning, the use of WhatsApp as an m-learning method for residents teaching is associated with worse clinical reasoning despite better global appreciation. The use of the WhatsApp app probably contributes to the dispersion of attention linked to the use of the smartphone.

Acknowledgments

The authors are grateful to Nikki Sabourin-Gibbs, Rouen University Hospital, for her help in editing the manuscript; to all the senior anesthesiologists from Caen and Rouen University Hospitals who agreed to be experts for SCT answers; and to all the anesthesiology residents who agreed to participate in this work. Funding support was provided solely from departmental sources.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Teaching documents especially prepared for easy readability on a smartphone (in French with English translation).

[PDF File (Adobe PDF File), 3MB - mhealth_v7i4e12825_app1.pdf]

Multimedia Appendix 2

"Step-by-step" clinical cases (in French with English translation).

[PDF File (Adobe PDF File), 428KB - mhealth_v7i4e12825_app2.pdf]

Multimedia Appendix 3

Several examples of the use of WhatsApp for learning purposes during the protocol (screenshots in French with English translation).

[PDF File (Adobe PDF File), 741KB - mhealth_v7i4e12825_app3.pdf]

Multimedia Appendix 4

Script concordance test used for the final evaluation (in French with English translation).

[PDF File (Adobe PDF File), 632KB - mhealth_v7i4e12825_app4.pdf]

Multimedia Appendix 5

Multiple-choice questions used for the final evaluation (in French with English translation).

[PDF File (Adobe PDF File), 720KB - mhealth_v7i4e12825_app5.pdf]

Multimedia Appendix 6

Online satisfaction questionnaire (in French with English translation).

[PDF File (Adobe PDF File), 550KB - mhealth_v7i4e12825_app6.pdf]

Multimedia Appendix 7

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i4e12825_app7.pdf]

References

1. Ellaway R, Masters K. AMEE Guide 32: e-Learning in medical education Part 1: Learning, teaching and assessment. Med Teach 2008 Jun;30(5):455-473. [doi: 10.1080/01421590802108331] [Medline: 18576185]
2. Alsobayel H. Use of Social Media for Professional Development by Health Care Professionals: A Cross-Sectional Web-Based Survey. JMIR Med Educ 2016 Sep 12;2(2):e15 [FREE Full text] [doi: 10.2196/mededu.6232] [Medline: 27731855]
3. Chan WS, Leung AY. Use of Social Network Sites for Communication Among Health Professionals: Systematic Review. J Med Internet Res 2018 Mar 28;20(3):e117. [doi: 10.2196/jmir.8382] [Medline: 29592845]
4. Christensen MA, Bettencourt L, Kaye L, Moturu ST, Nguyen KT, Olgin JE, et al. Direct Measurements of Smartphone Screen-Time: Relationships with Demographics and Sleep. PLoS One 2016;11(11):e0165331 [FREE Full text] [doi: 10.1371/journal.pone.0165331] [Medline: 27829040]
5. Eggermont S, Bloemendaal PM, van Baalen JM. E-learning any time any place anywhere on mobile devices. Perspect Med Educ 2013 Mar 30 [FREE Full text] [doi: 10.1007/s40037-013-0045-4] [Medline: 23670682]

6. Masters K, Ellaway RH, Topps D, Archibald D, Hogue RJ. Mobile technologies in medical education: AMEE Guide No. 105. Med Teach 2016 Jun;38(6):537-549. [doi: 10.3109/0142159X.2016.1141190] [Medline: 27010681]
7. Meola A. Business Insider French. WhatsApp is the most popular chat app in more than half the world Internet URL: http://www.businessinsider.fr/us/whatsapp-is-the-most-popular-chat-app-in-more-than-half-the-world-2016-5/ [accessed 2018-11-12] [WebCite Cache ID 73sJKGsIb]
8. AlFaris E, Irfan F, Ponnamperuma G, Jamal A, Van der Vleuten C, Al Maflehi N, et al. The pattern of social media use and its association with academic performance among medical students. Med Teach 2018 Dec;40(sup1):S77-S82. [doi: 10.1080/0142159X.2018.1465536] [Medline: 29732945]
9. Jamal A, Temsah M, Khan SA, Al-Eyadhy A, Koppel C, Chiang MF. Mobile Phone Use Among Medical Residents: A Cross-Sectional Multicenter Survey in Saudi Arabia. JMIR Mhealth Uhealth 2016 May 19;4(2):e61 [FREE Full text] [doi: 10.2196/mhealth.4904] [Medline: 27197618]
10. Goyal A, Tanveer N, Sharma P. WhatsApp for Teaching Pathology Postgraduates: A Pilot Study. J Pathol Inform 2017;8:6 [FREE Full text] [doi: 10.4103/2153-3539.201111] [Medline: 28400995]
11. Raiman L, Antbring R, Mahmood A. WhatsApp messenger as a tool to supplement medical education for medical students on clinical attachment. BMC Med Educ 2017 Jan 06;17(1):7 [FREE Full text] [doi: 10.1186/s12909-017-0855-x] [Medline: 28061777]
12. Compère V, Abily J, Moriceau J, Gouin A, Veber B, Dupont H, et al. Residents in tutored practice exchange groups have better medical reasoning as measured by script concordance test: a controlled, nonrandomized study. J Clin Anesth 2016 Aug;32:236-241. [doi: 10.1016/j.jclinane.2016.03.012] [Medline: 27290981]
13. Ducos G, Lejus C, Sztark F, Nathan N, Fourcade O, Tack I, et al. The Script Concordance Test in anesthesiology: Validation of a new tool for assessing clinical reasoning. Anaesth Crit Care Pain Med 2015 Feb;34(1):11-15. [doi: 10.1016/j.accpm.2014.11.001] [Medline: 25829309]
14. Bursztejn A, Cuny J, Adam J, Sido L, Schmutz J, de Korwin J, et al. Usefulness of the script concordance test in dermatology. J Eur Acad Dermatol Venereol 2011 Dec;25(12):1471-1475. [doi: 10.1111/j.1468-3083.2011.04008.x] [Medline: 22077872]
15. Dutton RP, Stansbury LG, Leone S, Kramer E, Hess JR, Scalea TM. Trauma mortality in mature trauma systems: are we doing better? An analysis of trauma mortality patterns, 1997-2008. J Trauma 2010 Sep;69(3):620-626. [doi: 10.1097/TA.0b013e3181bbfe2a] [Medline: 20093983]
16. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. JAMA 2008 Sep 10;300(10):1181-1196. [doi: 10.1001/jama.300.10.1181] [Medline: 18780847]
17. Kulier R, Coppus SFPJ, Zamora J, Hadley J, Malick S, Das K, et al. The effectiveness of a clinically integrated e-learning course in evidence-based medicine: a cluster randomised controlled trial. BMC Med Educ 2009 May 12;9:21 [FREE Full text] [doi: 10.1186/1472-6920-9-21] [Medline: 19435520]
18. Lubarsky S, Dory V, Duggan P, Gagnon R, Charlin B. Script concordance testing: from theory to practice: AMEE guide no. 75. Med Teach 2013;35(3):184-193. [doi: 10.3109/0142159X.2013.760036] [Medline: 23360487]
19. Dcode. Tirage au Sort URL: https://www.dcode.fr/tirage-au-sort-nombre-aleatoire [accessed 2018-11-14] [WebCite Cache ID 73vLExugV]
20. Rossaint R, Bouillon B, Cerny V, Coats TJ, Duranteau J, Fernández-Mondéjar E, et al. The European guideline on management of major bleeding and coagulopathy following trauma: fourth edition. Crit Care 2016 Apr 12;20:100 [FREE Full text] [doi: 10.1186/s13054-016-1265-x] [Medline: 27072503]
21. Duranteau J, Asehnoune K, Pierre S, Ozier Y, Leone M, Lefrant J. Recommandations sur la réanimation du choc hémorragique. Anesthésie & Réanimation 2015 Feb;1(1):62-74. [doi: 10.1016/j.anrea.2014.12.007]
22. Kassirer JP. Teaching clinical reasoning: case-based and coached. Acad Med 2010 Jul;85(7):1118-1124. [Medline: 20603909]
23. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. Med Educ 2015 Oct;49(10):961-973. [doi: 10.1111/medu.12775] [Medline: 26383068]
24. Enser M, Moriceau J, Abily J, Damm C, Occhiali E, Besnier E, et al. Background noise lowers the performance of anaesthesiology residents' clinical reasoning when measured by script concordance: A randomised crossover volunteer study. Eur J Anaesthesiol 2017 Dec;34(7):464-470. [doi: 10.1097/EJA.0000000000000624] [Medline: 28394819]
25. Faculté de Médecine - Université de Montréal Internet. Obtenir les scores de vos TCS par voie informatique - Centre de pédagogie appliquée aux sciences de la santé (CPASS) URL: https://www.cpass.umontreal.ca/recherche/groupe-de-recherche-cpass/axes-de-recherches/concordance/tcs/corriger_tcs/ [accessed 2018-11-14] [WebCite Cache ID 73vBHLU6Q]
26. Moreira IC, Ventura SR, Ramos I, Rodrigues PP. Development and assessment of an e-learning course on breast imaging for radiographers: a stratified randomized controlled trial. J Med Internet Res 2015;17(1):e3 [FREE Full text] [doi: 10.2196/jmir.3344] [Medline: 25560547]
27. Fransen F, Martens H, Nagtzaam I, Heeneman S. Use of e-learning in clinical clerkships: effects on acquisition of dermatological knowledge and learning processes. Int J Med Educ 2018 Jan 17;9:11-17 [FREE Full text] [doi: 10.5116/ijme.5a47.8ab0] [Medline: 29352748]

28. Lee L, Chao Y, Huang C, Fang J, Wang S, Chuang C, et al. Cognitive Style and Mobile E-Learning in Emergent Otorhinolaryngology-Head and Neck Surgery Disorders for Millennial Undergraduate Medical Students: Randomized Controlled Trial. J Med Internet Res 2018 Feb 13;20(2):e56 [FREE Full text] [doi: 10.2196/jmir.8987] [Medline: 29439943]
29. Shaw CM, Tan SA. Integration of mobile technology in educational materials improves participation: creation of a novel smartphone application for resident education. J Surg Educ 2015;72(4):670-673. [doi: 10.1016/j.jsurg.2015.01.015] [Medline: 25823746]
30. Liu Q, Peng W, Zhang F, Hu R, Li Y, Yan W. The Effectiveness of Blended Learning in Health Professions: Systematic Review and Meta-Analysis. J Med Internet Res 2016;18(1):e2 [FREE Full text] [doi: 10.2196/jmir.4807] [Medline: 26729058]

Abbreviations

e-learning: electronic learning
MCQ: multiple-choice questions
m-learning: mobile learning
SCT: script concordance test

Edited by G Eysenbach; submitted 15.11.18; peer-reviewed by LA Lee, VK Bandi, A Paglialonga; comments to author 10.01.19; revised version received 17.01.19; accepted 23.02.19; published 09.04.19.

Please cite as:
Clavier T, Ramen J, Dureuil B, Veber B, Hanouz JL, Dupont H, Lebuffe G, Besnier E, Compere V
Use of the Smartphone App WhatsApp as an E-Learning Method for Medical Residents: Multicenter Controlled Randomized Trial
JMIR Mhealth Uhealth 2019;7(4):e12825
URL: http://mhealth.jmir.org/2019/4/e12825/
doi:10.2196/12825
PMID:30964435

Original Paper

Language Translation Apps in Health Care Settings: Expert Opinion

Anita Panayiotou[1], BBNSci, PGDipArts, Dpsych, DipMgt, CertIV HealthAdmin; Anastasia Gardner[1], BSN, BMedSci, BSc; Sue Williams[1], BSci, BASc, MSci; Emiliano Zucchi[2], BA; Monita Mascitti-Meuter[3], BA, MA, MIB, Cert IVTrainAss; Anita MY Goh[4], BSc, DPsych; Emily You[5], B PubMedServMgt, MSocSciHSMgt, PhD; Terence WH Chong[3], MBBS, MP, MBA, FRANZP, CertOldAgePsych; Dina Logiudice[6], FRACP, PhD; Xiaoping Lin[1], BA, Grad Dip, PGDipArts, PhD; Betty Haralambous[1], BSW, MSW, GCBA; Frances Batchelor[1], BAppSc, GradDipEd, MHS, PhD

[1]National Ageing Research Institute, Parkville, Australia

[2]Northern Health, Epping, Australia

[3]St Vincent's Hospital, Fitzroy, Australia

[4]National Ageing Research Institute , Parkville, VIC, Academic Unit for Psychiatry of Old Age, University of Melbourne, Melbourne Health, Parkville VIC, Parkville, Australia

[5]Academic Unit for Psychiatry of Old Age, University of Melbourne, Parkville, Australia

[6]Melbourne Health, Parkville, Australia

Corresponding Author:
Frances Batchelor, BAppSc, GradDipEd, MHS, PhD
National Ageing Research Institute
PO Box 2127 The Royal Melbourne Hospital
Parkville, 3050
Australia
Phone: 61 383872383
Email: f.batchelor@nari.edu.au

Abstract

Background: Currently, over 300 languages are spoken in Australian homes. People without proficient English from non-English speaking countries may not receive equitable care if their health care workers do not speak their primary language. Use of professional interpreters is considered the gold standard; however, for a variety of reasons, it is often limited to key aspects of care such as diagnosis and consent. With the emergence of mobile technologies, health care workers are increasingly using digital translation tools to fill this gap. However, many of these technologies have not been developed for health care settings and their use has not been evaluated.

Objective: This study aimed to evaluate iPad-compatible language translation apps to determine their suitability for enabling everyday conversations in health care settings.

Methods: Translation apps were identified by searching the Apple iTunes Store and published and grey literature. Criteria for inclusion were that the apps were available at no cost, able to translate at least one of the top 10 languages spoken in Australia, and available for use on iPad. Apps that met inclusion criteria were reviewed in 2 stages. Stage 1 was the feature analysis conducted by 2 independent researchers, where apps were evaluated for offline use, input and output methods, and number of languages. Stage 2 was the analysis of suitability for everyday communication in the health care setting, conducted by 2 independent professionals with expertise in translation and cross-cultural communication. Apps that enabled key aspects of care normally within the realm of professional interpreters, such as assessment, treatment and discharge planning, and seeking consent for medical treatments, were considered unsuitable.

Results: In total, 15 apps were evaluated. Of these, 8 apps contained voice-to-voice and voice-to-text translation options. In addition, 6 apps were restricted to using preset health phrases, whereas 1 app used a combination of free input and preset phrases. However, 5 apps were excluded before stage 2. In addition, 6 of the 10 remaining apps reviewed in stage 2 were specifically designed for health care translation purposes. Of these, 2 apps were rated as suitable for everyday communication in the health care setting—culturally and linguistically diverse Assist and Talk To Me. Both apps contained simple and appropriate preset health phrases and did not contain conversations that are normally within the realm of professional interpreters.

Conclusions: All iPad-compatible translation apps require a degree of caution and consideration when used in health care settings, and none should replace professional interpreters. However, some apps may be suitable for everyday conversations, such as those that enable preset phrases to be translated on subject matters that do not require a professional interpreter. Further research into the use of translation technology for these types of conversations is needed.

(JMIR Mhealth Uhealth 2019;7(4):e11316) doi:10.2196/11316

KEYWORDS

health care; communication; language; technology

Introduction

Background

The widespread prevalence of telemedicine and telehealth has led to an increasing acceptance of technology in health care. Although there is limited evidence for the effective use of translation technology in medical and health care settings, clinicians anecdotally report the use of internet and mobile apps for language translation purposes. This raises potential concerns as most apps have not been specifically developed or validated for use in a medical or health care context. However, there is potential for technology to be used to improve everyday clinical communication between patients and staff in the health care setting when used as an adjunct to professional interpreters [1].

Australia is one of the most ethnically and culturally diverse countries in the world [2]. According to the Australian Bureau of Statistics [2], almost half of all Australians were either born overseas or had at least one parent who was born overseas. In 2015, Australia had the ninth largest population of people born overseas worldwide and a higher proportion of overseas-born people (26%) compared with other countries founded on migration such as New Zealand (23%), Canada (22%), and the United States (14%) [2]. Net overseas migration continues to increase in Australia and has shown periods of influx linked to major world events. For example, following the Second World War, Australia saw a high proportion of European migrants. In more recent times, migration has predominantly been from China, India, and the Middle East, with Asian countries now making up 8 out of the top 10 migration countries to Australia [3]. In 2016, there were over 300 different languages spoken in Australian homes and more than one-fifth of Australians spoke a language other than English at home [2]. After English, the 10 most common languages spoken at home in Australia are Mandarin, Arabic, Cantonese, Vietnamese, Italian, Greek, Hindi, Spanish, Punjabi, and Tagalog [2].

The ability to convey essential care needs (eg, addressing pain, help with hygiene), communicate simple safety messages, and provide orientation cues are essential in health care settings. People without proficient English from non-English speaking (NES) countries may not receive equitable care if their health care workers do not speak their primary language [4,5]. The use of professional interpreters is considered the gold standard [6]. However, because of issues related to cost, access, availability, and time constraints [7,8], use of professional interpreters in health care is often limited to specific aspects of care, such as comprehensive assessments, procedural consent, diagnosis, and the development of treatment plans. Everyday communication between health care workers and clients, when there is a language barrier, generally occurs without professional interpreters and has been described in the literature as *getting by*, where health care workers rely on gestures, facial expressions, and knowledge of minimal key words in the target language [9,10]. The *getting by* approach has the potential for miscommunication, which may lead to inappropriate or inadequate care provision and patients' needs being unmet. At worst, the *getting by* approach can result in the provision of inappropriate or nonbeneficial treatments and care as highlighted by Runci et al (2012) [11], finding a higher frequency of prescription of antipsychotic drugs for Italian speaking residents in English-speaking residential care facilities than their counterparts in language specific facilities.

Although using an interpreter remains the gold standard for complex medical and legal discussions in all settings, in some situations, it is not appropriate or feasible to use an interpreter, yet communication remains an issue. Through the widespread uptake of mobile devices, technology enabling language translation has been identified as a potential way to improve communication between patients and staff in health care settings when used as an adjunct to professional interpreters [1]. Very few studies have evaluated the use of translation apps in medical and health care settings and even fewer have compared multiple translation apps or examined the contexts in which their use may be suitable.

Although early studies of Web-based language tools, such as Google Translate, highlight high levels of user satisfaction [12], the risks relating to accuracy when used in the clinical setting have become more apparent [12-15]. One study evaluated text translation of 10 common questions relating to medical history and assessment from English into 10 languages and found a wide discrepancy in the accuracy depending on the target language [13]. Vietnamese and Polish translation had the lowest accuracy (10% correctly translated), whereas Spanish had the highest accuracy (80% correctly translated). Another study evaluated the accuracy of 10 medical phrases in 26 languages [14]. Of the total translations, approximately 58% of the translations were accurately translated from English. However, the accuracy among the different languages also showed variability, with African languages scoring the lowest accuracy (42%), followed by Asian languages (46%), then Eastern European languages (62%), and Western European languages (74%). The authors reported the presence of some phrases where the translation resulted in considerable changes to the intended meaning when using complex medical terminology in high-risk situations. For example, "Your child is fitting" was incorrectly translated to "Your child is dead" in Swahili and "Your husband has the opportunity to donate his organs" was incorrectly

translated to "Your husband can donate his tools" in Polish. As a result, the authors cautioned against the use of Google Translate when obtaining consent for surgery or other medical procedures and participation in research.

Beh and Canty [15] also reviewed the accuracy of Google Translate in a simulated preanesthetic consultation between an English-speaking anesthetist and a Mandarin-speaking patient. In total, 24 English phrases and 13 Mandarin phrases were tested and an independent anesthetist fluent in both English and Mandarin assessed the translation accuracy. The accuracy of translation from English to Mandarin was 72% and from Mandarin to English was 67%, improving with short or simple phrases that did not contain technical information or when speaking clearly and slowly. The authors concluded that Google Translate was not accurate enough to replace professional interpreters but might be useful if an interpreter was not available.

Albrecht et al [16] conducted a 6-week trial of a German translation app called *xprompt–multilingual assistance*, designed for use in health care settings. The app allowed for preset health phrases to be translated and was used for basic conversations. Nursing staff were then surveyed about their experiences. Although most staff reported that the translation tool was helpful for communicating with patients who spoke another language, was easy to use, and that there were no obvious problems with the usability of the device, some reported that the technology was not practical, was too time consuming, and did not integrate well into existing workflows. They also reported difficulties using the technology with older patients who were unfamiliar with technology or unable to use the app because of visual impairment or illiteracy. The staff also reported that the desired target language was not always available. Explaining the menu items in the app caused problems in some instances.

Objectives

Given the availability and widespread acceptance of language translation apps by the general public and anecdotal evidence of their use in health care settings, more research is required to evaluate their use, particularly in health care situations when professional interpreters are not normally used, such as everyday clinical conversations, and with particular cohorts, such as older people from NES backgrounds. To date, research evaluating language translation technology in health care settings has done so in complex situations, such as those that involve seeking consent [14], conducting medical assessments [13], or engaging in technical or complex medical conversations [15].

A previous study [17] evaluated mobile medical translation apps where the authors identified key features and scored each app on the basis of usability and functionality. This evaluation aims to provide an overview of iPad (Apple Inc) compatible language translation apps (at no cost) and considerations for use in real-world health care settings. This study uses experts in the field of health care translation and cross-cultural communication to evaluate the content of translation apps and provide expert opinion regarding their suitability for health care situations in which an interpreter would not be available, such as providing orientation cues and conveying essential care needs, including identifying pain or the need for toileting. This is the first study to evaluate translation apps on the basis of their suitability for everyday conversations in the health care setting.

Methods

Study Design

The study design involves 2 components:

1. A search for iPad-compatible language translation apps at no cost.
2. An evaluation of retrieved apps comprising 2 stages—feature analysis and analysis of suitability for everyday clinical conversations in health care settings.

Component 1: Search for Available Language Translation Apps

Searching for iPad-compatible translation apps was conducted by first searching the Apple iTunes Store (Apple Inc) on 22 August, 2017, using the search terms in Figure 1. Following this, grey literature (Google Search and Google Scholar) and published literature (PubMed) searches were conducted for published articles related to smartphone or tablet apps used in health care settings for translation purposes. Later, iMedicalApps, a website that reviews all medical apps, was searched using the terms in Figure 2. Finally, any apps that the authors were familiar with from professional experience, which were not discovered in the previous searches, were included.

Figure 1. Search Terms for Apple Store.

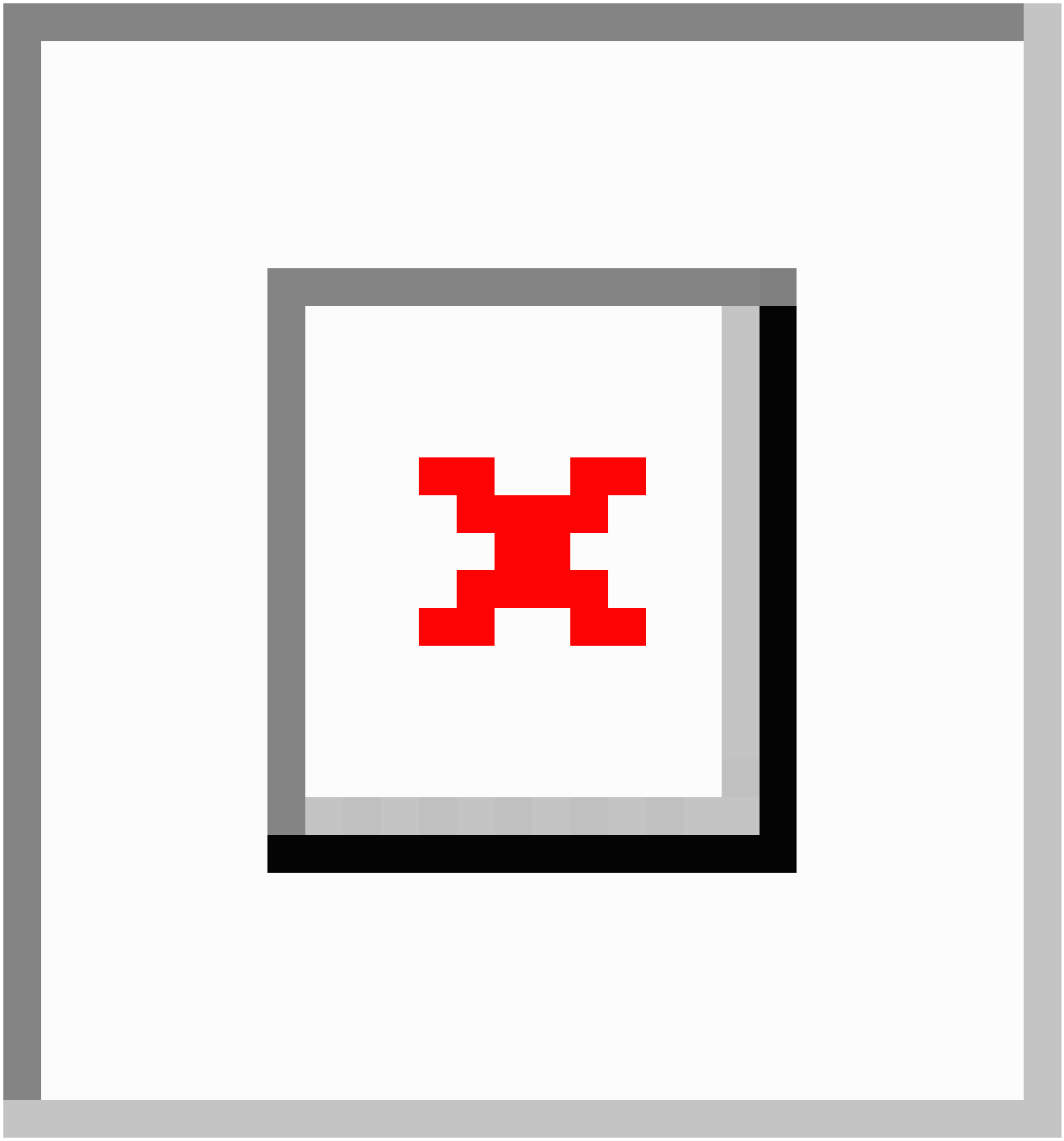


Figure 2. Search Terms for Literature, iMedicalApps and Google Search.

(multilingual or language*) and translat*
and
app or apps or application* or technolog*
and
healthcare or health care or hospital* or medical or health or clinical

Translation apps were included if they were developed for or used for language translation purposes, were available at no cost, were available on iPad (Apple Inc), and enabled translation to or from English. Each app had to include translation to at

least one of the top 10 languages spoken in Australian homes as of 2016 (Mandarin, Arabic, Cantonese, Vietnamese, Italian, Greek, Hindi, Spanish, Punjabi, and Tagalog, excluding English) [2]. The apps had to operate on iOS 10.3 that was available during August 2017 on the iPad. Apps that required a fee or were only available for use on Android devices were excluded as this project forms a part of a larger pragmatic evaluation of the use of iPad-compatible apps in health care settings, where budgetary limitations are known to prohibit health care providers use of paid apps.

Component 2: Evaluation of Retrieved Apps

Stage 1: Feature Analysis

Once identified, an analysis of all apps that met inclusion criteria was conducted by 2 independent clinical researchers (AP and RTJ). The researchers both had clinical backgrounds and extensive knowledge of technology. The researchers evaluated the apps according to the following key categories: offline use, input and output method, and languages available. Issues that may arise for use in the health care setting were also recorded, including ease of use on the basis of whether the app required a high level of user knowledge or required many steps to navigate through the app (Multimedia Appendix 1). A consensus approach was adopted by the 2 clinical researchers on all aspects of each app. Apps were excluded after the feature analysis stage if they required any in-app purchases or subscriptions, as this was considered a barrier to use in the health care setting as part of a larger study.

Stage 2: Analysis of Suitability for Everyday Communication

The apps included were then evaluated on the basis of their suitability for everyday communication in health care settings by 2 independent professionals (EZ and MM) with expertise in translation and cross-cultural communication in health care. Both experts are academically qualified in language and cultural studies, are polyglots, and have extensive experience as language and cultural diversity managers in large public hospitals. This stage of the evaluation focused on the suitability of apps for situations in which an interpreter would not be necessary, such as providing orientation cues and conveying essential care needs, including identifying pain or the need for toileting. Apps were considered less suitable when they contained content or allowed for conversations that were considered critical points in health care. These critical conversations were defined, in line with the state government Language Services Policy [18], as those involving clinical assessment, provision of diagnoses, conversations about treatment and care planning, discharge planning, and medicolegal information such as seeking consent for medical treatments. These types of conversations were considered beyond the scope of translation technology as they require a professional interpreter because of the need for a high degree of accuracy, the need to confirm understanding from patients, and the need to allow patients to ask questions. Other factors that could have an impact on suitability for everyday communication in health care settings were also identified. These included the type, content, and structure of phrases available for selection in the apps, taking into consideration the complexity and sensitivity of information and the ability to allow open-ended or 2-way conversation with appropriate responses available for selection by the patient. Recommendations for use in a health care setting were included in this stage (see Table 1). The evaluators were required to reach consensus about each aspect of each app.

Table 1. Analysis of suitability for everyday conversations in the health care setting.

App name	Can the app be used for critical points in health care requiring professional interpreters					Other factors or comments	Overall rating of suitability for everyday communication (low or high)
	Ax[a]	Dx[b]	Tx[c]	D/C[d]	Medicolegal[e]		
CALD[f] Assist	Y[g]	N[h]	Y	Y	N	Some phrases and questions are lengthy or complex. Only translates preset phrases. The phrases relating to assessment, treatment, and discharge are considered within the scope of everyday clinical conversations (eg, “Do you have pain?”; “I need to do a scan of your bladder”; “You are leaving hospital today”).	High
Canopy Speak	Y	Y	Y	Y	Y	Many questions are lengthy, highly detailed, complex, open-ended, cover a broad scope, and are highly sensitive (eg, “Do you have thoughts of killing others?”; “When you take that medicine, does it make you feel sleepy, give you a headache, or make you feel nauseated?”; “Do you use tobacco now? In the past? For how long? Type and amount daily?”).	Low
Google Translate	Y	Y	Y	Y	Y	Free input allows for any information input to be translated. Therefore, it is considered beyond the scope of everyday clinical conversation.	Low
MediBabble Translator	Y	N	N	N	Y	Many questions are lengthy, highly detailed, complex, sensitive, and/or cover a broad scope, which is beyond the scope of everyday clinical conversation (eg, “Are you allergic to any medication?”; “Do you think about harming yourself?”; ”I’d like to know what the pain feels like”; “Do you experience recurrent or persistent thoughts, impulses, or images that are inappropriate or upsetting?”; “Are you experiencing prolonged or excessive menstrual bleeding at irregular intervals or more frequently than your normal menstrual periods?”).	Low
Microsoft Translator	Y	Y	Y	Y	Y	Free input allows for any information input to be translated. Therefore, it is considered beyond the scope of everyday clinical conversation.	Low
Naver Papago Translate	Y	Y	Y	Y	Y	Free input allows for any information input to be translated. Therefore, it is considered beyond the scope of everyday clinical conversation.	Low
SayHi Translate	Y	Y	Y	Y	Y	Free input allows for any information input to be translated. Therefore, it is considered beyond the scope of everyday clinical conversation.	Low
Talk To Me	Y	N	Y	Y	N	Only translates preset phrases. The phrases relating to assessment, treatment, and discharge are considered within the scope of everyday clinical conversation (eg, “Are you sad?”; “I will take your blood pressure”).	High
TripLingo	Y	Y	Y	Y	Y	Free input allows for any information input to be translated. Therefore, it is considered beyond the scope of everyday clinical conversation.	Low
Universal Doctor Speaker	Y	Y	Y	Y	U[i]	Only allows for limited preset phrases and questions to be translated. In addition, includes open-ended questions, which are considered beyond the scope of everyday clinical conversation. Medical information about an individual can be saved and this poses a risk to confidentiality (eg, “Are you allergic to any medication?”; “You have the following illness or condition—depression / anxiety / chronic depression / obsessive-compulsive disorder, etc”).	Low

[a]Ax: assessment.

[b]Dx: diagnosis.

[c]Tx: treatment and care planning.

[d]D/C: discharge.

[e]medicolegal: medicolegal conversations including consent.

[f]CALD: culturally and linguistically diverse.

[g]Y: yes.

[h]N: no.

[i]U: unsure.

Results

In total, 15 apps met the inclusion criteria and were evaluated (Figure 3).

Stage 1: Feature Analysis

An initial agreement of 93% (70/75) was achieved among the researchers regarding the features available within each app. Differences identified on the remaining aspects of an app were resolved via discussion. Most apps enabled free voice or text input, and this feature usually required an internet connection even after the language package had been downloaded. In total, 53% (8/15) of the apps were capable of voice-to-voice translation, 53% (8/15) of the apps were capable of voice-to-text translation, 33% (5/15) of the apps were capable of text-to-voice translation, and 33% (5/15) of the apps were capable of text-to-text translation. In total, 7 out of 15 apps (47%) enabled the translation of preset phrases. However, 6 of these 7 apps did not allow for free input. Furthermore, 33% (5/15) of the apps could be used offline, but they required language packages to be downloaded. Although TripLingo was capable of multiple input and output functions, it was not specifically designed for the health care setting and contained very few preset phrases that were considered suitable.

In addition, 6 of the 15 apps (40%) were related specifically to health care translation. These were CALD Assist, Canopy Speak, Dr. Passport (Personal), MediBabble Translator, Talk To Me, and Universal Doctor Speaker. All of these apps were limited to the use of preset phrases and did not allow free voice, text, or image input. Only 2 of the 15 apps (13%) were capable of 2-way conversation between a patient and health care worker—CALD Assist and Dr. Passport. In addition to containing closed questions that required a simple *Yes* or *No* response, CALD Assist also enabled some open-ended questions with limited selections to be made by the patient and some follow-up questions. An example of a follow-up question was "Have you lost weight in the last six months?" then "How much weight have you lost?" Several options were available on the screen for the patient to select. Dr. Passport also allowed for 2-way conversation between a patient and health care worker. However, this was only possible by enabling patients to select preset phrases to translate to their health care worker. This app is divided into preset health phrases for the patient and a separate section for health care workers. This app appeared to be intended for patient-led conversations and not for conversations led by health care practitioners.

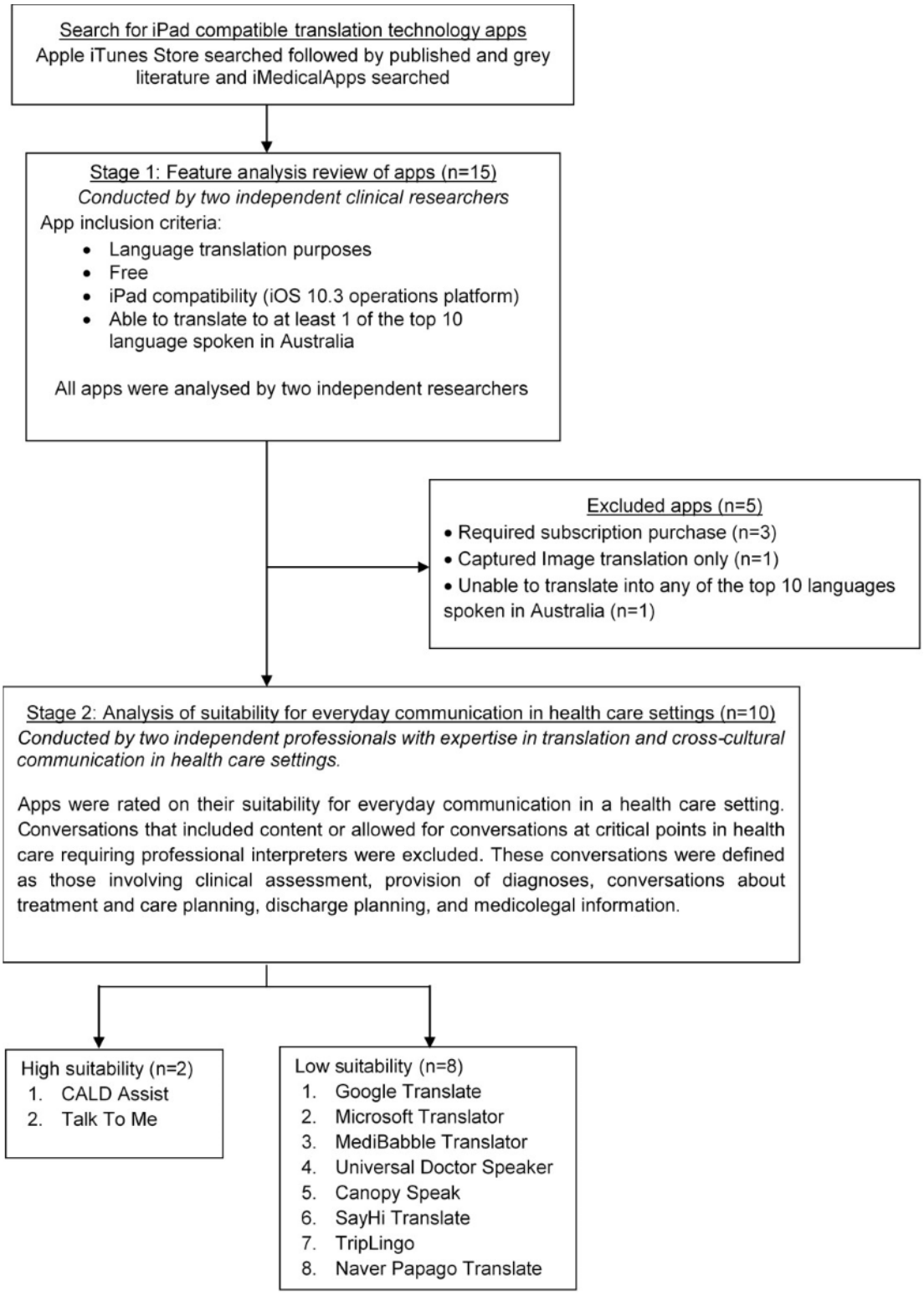
Of the 15 apps evaluated in stage 1, only 10 continued to stage 2 (Figure 3). iTranslate, iTranslate Voice, and Speak and Translate were excluded as they required monthly subscriptions once the free trial period had ended. Waygo was excluded as it only translated captured images (ie, text within images). Dr. Passport was also excluded as it was only available for free when using it to translate from English to Japanese. All other language translations required a fee.

Stage 2: Analysis of Suitability for Everyday Communication in Health Care Settings

An initial agreement of 73% (44/60) was achieved between the evaluators regarding the evaluation of the content of the translation apps. Differences identified on the remaining aspects of an app were resolved via discussion.

Of the 10 apps evaluated for their suitability, none were entirely suitable (refer to Table 1). All 10 apps enabled conversations about assessment and all apps, except for one (MediBabble Translator), enabled conversations about treatment or care planning and discharge. Furthermore, 3 of the 10 apps did not enable conversations about diagnosis and medicolegal information. The apps that enabled conversations in the least number of critical points in health care were MediBabble Translator, CALD Assist, and Talk To Me. This contributed to an overall suitability rating of either high or low suitability. The 2 apps, CALD Assist and Talk To Me, were rated as highly suitable on this basis.

Figure 3. Flow diagram of the process used to identify eligible apps for languages translation. CALD: culturally and linguistically diverse.



Discussion

Principal Findings

This study evaluated iPad-compatible language translation apps on the basis of their features and provided expert opinion on their suitability for everyday conversations in which an interpreter would not be available, such as providing orientation cues and conveying essential care needs, including identifying pain or the need for toileting, in a real-world health care setting. This is the first study to involve experts in health care translation and cross-cultural communication in the evaluation of translation technology suitability. This study's results show that only 2 apps could be considered suitable.

In total, 15 iPad-compatible language translation apps were identified from searches in the Apple iTunes Store and published and grey literature that met the inclusion criteria. These apps were evaluated in 2 stages to analyze features and suitability

for everyday communication in health care settings. The feature analysis stage identified that the apps enabled translation of over 100 languages, enabled different input and output modes, which determined whether they could be used offline, and had been developed for a range of different purposes, most commonly for health care, travel, and business purposes. Apps that had been designed for other purposes, such as travel or business, were considered to have limited applicability to the health care setting. Of these 15 apps, 5 were excluded from stage 1 and 10 apps were then evaluated for their suitability of enabling everyday conversations in a health care setting. All apps evaluated in stage 2 required a degree of caution and consideration before being used in health care settings. However, the degree of suitability varied across apps. For example, apps that allowed free input of information to be translated were deemed less suitable as there were no limits to the way the apps could be used, whereas apps that only enabled translation of preset phrases had the potential to limit the contexts in which the apps were used and were deemed more suitable. Despite this, all the apps enabled conversations about topics considered to be critical points in health care, which would normally require professional interpreters, such as clinical assessment and conversations about treatment and care planning. Translating these conversations requires a high degree of accuracy and the ability to confirm understanding from patients and allow patients to ask questions, which are not met by these translation apps. Discussing these topics with translation technology could result in miscommunication, which might lead to serious negative health outcomes for patients.

Apps considered to be most suitable for the health care setting were CALD Assist and Talk To Me. Both enabled conversations in the least number of critical phase topics that required professional interpreters and limited the translation to preset phrases that were led by health care practitioners. Although MediBabble Translator limited topics of discussion to 2 critical phase topics and restricted input to preset phrases, the phrases were deemed to be outside the scope of everyday clinical conversations as they included highly detailed, sensitive and/or personal open-ended questions, such as “Do you think about harming yourself or others?”

A recent study [17] evaluated mobile medical translation apps where the authors identified key features and scored each app on the basis of usability and functionality. Apps that were low cost, able to be used offline, and did not contain advertisements and in-app purchases were compatible with multiple platforms (iPhone, iPad, Android, and Nexus), were easy to navigate, and were well presented, and these apps scored highly. The study rated Canopy Medical Translator, Universal Doctor Speaker, and Vocre Translate favorably. However, these apps were either excluded at stage 1 of our evaluation on the basis of requiring payment or were rated as beyond the scope of everyday clinical conversations at stage 2. Given that this study focused on the clinical utility of translation apps for enabling everyday conversations in the health care setting and that eligibility for inclusion was determined on the basis of informing a larger pragmatic evaluation, which necessitated that the apps were available at no cost and were compatible with iPads, the ratings differed.

Considerations for Clinical Use

This study evaluated translation technology features and provided expert opinion regarding their suitability for everyday clinical conversations, not necessitating professional interpreters. These factors are important considerations for use in health care settings and for the development of new translation technologies. In addition to these considerations, other factors such as current policy regarding the use of translation technology, data security, and confidentiality should be considered carefully in the design and use of apps for health care. Those that require access to the internet or that keep a record of conversations may require additional precautions, for example, not disclosing the patient’s identity with the app, avoiding collection of personal or sensitive information with the app, avoiding the use of personal devices when using Web-based apps, and using a secure internet connection so that the individual device or location cannot be identified. Other considerations in design and development of apps may include avoiding use of a device with a small screen (eg, smartphone), which may pose difficulties for patients with visual impairments or reduced motor dexterity.

It is not possible to provide access to professional interpreters for a patient’s entire health care episode, and there is an urgent need for effective, accessible, and safe tools, such as translation technology, to facilitate everyday communication to improve health outcomes. However, there is a risk that translation technology may become the preferred means of communicating with patients with limited English proficiency because of the perceived simplicity, accessibility, and timeliness. Over dependence and over reliance on this technology may impact negatively on establishing rapport and providing high quality care. CALD Assist and Talk To Me were the only apps included in this evaluation that provided users with a disclaimer about their limitations and stressed the importance of using professional interpreters where possible. CALD Assist and Talk To Me were specifically designed for health care settings and both restricted communication to preset phrases that were considered within the scope of everyday clinical conversations. They did not include topics and situations that were considered to require professional interpreters. Although Canopy Speak and MediBabble Translate were also specifically designed for health care settings, the evaluators found that these apps were difficult to navigate and contained content that was beyond everyday clinical conversations, requiring a professional interpreter.

Limitations

It was beyond the scope of this study to examine translation accuracy and cultural suitability. These are important aspects of apps that would have an impact upon the effectiveness of use in health care settings and involve the suitability of translated words for the context, the syntax of the translated phrases (eg, order of words and grammar), and the ability to recognize different accents and dialects (when using free voice input). Previous studies have identified poorer accuracy for the translation of non-western languages in Google Translate [14-16]. Further research involving experts, health professionals, and consumers is required to evaluate the translation accuracy and cultural suitability of other apps and in other contexts.

Given how rapidly technology develops and changes, it was not possible to capture every available iPad-compatible app. Therefore, this study provides a snapshot of the available iPad-compatible translation apps and considerations for use in the health care setting. As more apps become available, further research will be required. Furthermore, as this study excluded apps that were available at a cost, only available on Android devices, or involved languages other than English as the primary language, further research evaluating these apps is warranted.

Conclusions

Overall, the findings of this evaluation have identified that iPad-compatible language translation technology requires careful consideration when used in health care settings, and it may be completely or partly prohibited by existing health care policies. The degree of suitability for health care settings varies on the basis of the content and features available within each app. Those that allow translation of free voice, text, or image information were deemed to be the least suitable for health care settings. Of the apps evaluated, only 2 were considered to be highly suitable for the health care setting, on the basis of their use of preset health related phrases; these were CALD Assist and Talk To Me. When the content was appropriate, preset health phrases were considered the most suitable because information was brief, simple, and contained. Although many apps featured preset phrases, the content was frequently considered unsuitable as phrases were overly complex, lengthy, contained sensitive information, and did not allow for an appropriate answer. When considering the use of translation technology in health care settings, clinicians are encouraged to consider the capabilities of the translation technology itself, as well as the particular situation, the patient, and any organizational policies. Translation technology is not an appropriate substitute for a professional interpreter and further research is required to evaluate its use for everyday conversations in real clinical settings. However, it is not logistically or financially possible to access a professional interpreter for every interaction in a patient's health care episode. Therefore, there is a pressing need to develop tools that facilitate this communication in a safe and effective manner. Translation technology can play an important role, but this research clearly shows the importance of considering the scope of its use.

Acknowledgments

The Melbourne Ageing Research Collaboration provided support and funding to conduct this project.

The authors thank Mr Ryan Townley-Jones for his assistance in the feature analysis stage of the research. The authors would also like to thank the Melbourne Ageing Research Collaboration.

Conflicts of Interest

MMM was involved in the initial development and testing of the Talk To Me app. All other authors declare no conflicts of interest.

Multimedia Appendix 1

Feature analysis stage.

[DOCX File, 16KB - mhealth_v7i4e11316_app1.docx]

References

1. Chang DT, Thyer IA, Hayne D, Katz DJ. Using mobile technology to overcome language barriers in medicine. Ann R Coll Surg Engl 2014 Sep;96(6):e23-e25 [FREE Full text] [doi: 10.1308/003588414X13946184903685] [Medline: 25198966]
2. Australian Bureau of Statistics Census. 2016. Cultural Diversity in Australia URL: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2071.0~2016~Main%20Features~Cultural%20Diversity%20Data%20Summary~30 [accessed 2018-04-20] [WebCite Cache ID 6yoNiMmwJ]
3. Phillips J, Simon-Davies J. Parliamentary of Australia. Canberra: Commonwealth of Australia; 2016. Migration to Australia: A quick guide to the statistics URL: https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/rp1617/Quick_Guides/MigrationStatistics [accessed 2019-03-05] [WebCite Cache ID 76eU775Aw]
4. Cowley S, Houston AM. A structured health needs assessment tool: acceptability and effectiveness for health visiting. J Adv Nurs 2003 Jul;43(1):82-92. [Medline: 12801399]
5. Rao DV, Warburton J, Bartlett H. Health and social needs of older Australians from culturally and linguistically diverse backgrounds: issues and implications. Australas J Ageing 2006 Dec;25(4):174-179 [FREE Full text] [doi: 10.1111/j.1741-6612.2006.00181.x]
6. Victorian State Government. Melbourne: Victorian Government Office of Multicultural Affairs and Citizenship; 2014. Using interpreter Services: Victorian Government guidelines on policy and procedures URL: https://www.multicultural.vic.gov.au/images/2017/UsingInterpretingServicesVictorianGovernmentGuidelines.pdf [accessed 2019-03-03] [WebCite Cache ID 76bztKoUa]

7. Vanstone R. The Victorian Foundation Foundation for Survivors of Torture. Melbourne: Foundation House; 2012. Exploring barriers and facilitators to the use of qualified interpreters in health URL: https://refugeehealthnetwork.org.au/wp-content/uploads/Interpreters+in+health+discussion+paper+FINAL.pdf [accessed 2019-03-03] [WebCite Cache ID 76bzNqO5s]
8. O'Hagan M. Massively open translation: unpacking the relationship between technology and translation in the 21st Century. Int J Commun 2016;10:929-946.
9. Diamond LC, Schenker Y, Curry L, Bradley EH, Fernandez A. Getting by: underuse of interpreters by resident physicians. J Gen Intern Med 2009 Feb;24(2):256-262 [FREE Full text] [doi: 10.1007/s11606-008-0875-7] [Medline: 19089503]
10. Villarruel AM, Portillo CJ, Kane P. Communicating with limited English proficiency persons: implications for nursing practice. Nurs Outlook 1999;47(6):262-270. [Medline: 10626284]
11. Runci SJ, Eppingstall BJ, O'Connor DW. A comparison of verbal communication and psychiatric medication use by Greek and Italian residents with dementia in Australian ethno-specific and mainstream aged care facilities. Int Psychogeriatr 2012 May;24(5):733-741. [doi: 10.1017/S1041610211002134] [Medline: 22137090]
12. Kaliyadan F, Gopinathan PS. The use of Google language tools as an interpretation aid in cross-cultural doctor-patient interaction: a pilot study. Inform Prim Care 2010;18(2):141-143 [FREE Full text] [Medline: 21078237]
13. Nguyen-Lu N, Reide P, Yentis SM. Do you have a stick in your mouth?-use of Google Translate as an aid to anaesthetic pre-assessment. Anaesthesia 2010;65(1):96-97. [doi: 10.1111/j.1365-2044.2009.06184_4.x]
14. Patil S, Davies P. Use of Google Translate in medical communication: evaluation of accuracy. Br Med J 2014 Dec 15;349:g7392 [FREE Full text] [Medline: 25512386]
15. Beh TH, Canty DJ. English and Mandarin translation using Google Translate software for pre-anaesthetic consultation. Anaesth Intensive Care 2015 Nov;43(6):792-793 [FREE Full text] [Medline: 26603812]
16. Albrecht UV, Behrends M, Schmeer R, Matthies HK, von Jan U. Usage of multilingual mobile translation applications in clinical settings. JMIR Mhealth Uhealth 2013 Apr 23;1(1):e4 [FREE Full text] [doi: 10.2196/mhealth.2268] [Medline: 25100677]
17. Khander A, Farag S, Chen KT. Identification and evaluation of medical translator mobile applications using an adapted APPLICATIONS scoring system. Telemed J E Health 2017 Dec 22;24(8):594-603. [doi: 10.1089/tmj.2017.0150] [Medline: 29271702]
18. State of Victoria, Department of Health and Human Services. Language Services Policy. Melbourne: Victorian Department of Health and Human Services; 2017:1-29.

Abbreviations

CALD: culturally and linguistically diverse
NES: non-English speaking

Edited by G Eysenbach; submitted 26.06.18; peer-reviewed by UV Albrecht, K Chen, H Guo; comments to author 06.10.18; revised version received 29.11.18; accepted 31.12.18; published 09.04.19.

Please cite as:
Panayiotou A, Gardner A, Williams S, Zucchi E, Mascitti-Meuter M, Goh AMY, You E, Chong TWH, Logiudice D, Lin X, Haralambous B, Batchelor F
Language Translation Apps in Health Care Settings: Expert Opinion
JMIR Mhealth Uhealth 2019;7(4):e11316
URL: https://mhealth.jmir.org/2019/4/e11316/
doi:10.2196/11316
PMID:30964446

Original Paper

The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study

Leming Zhou[1], DSc, PhD; Jie Bao[1], MSc; I Made Agus Setiawan[1,2], MSc; Andi Saptono[1], PhD; Bambang Parmanto[1], PhD

[1]Department of Health Information Management, University of Pittsburgh, Pittsburgh, PA, United States

[2]Department of Computer Science, Udayana University, Badung, Bali, Indonesia

Corresponding Author:
Leming Zhou, DSc, PhD
Department of Health Information Management
University of Pittsburgh
6021 Forbes Tower
3600 Forbes Ave at Meyran Ave
Pittsburgh, PA, 15260
United States
Phone: 1 412 383 6653
Fax: 1 412 383 6655
Email: lmzhou@gmail.com

Abstract

Background: After a mobile health (mHealth) app is created, an important step is to evaluate the usability of the app before it is released to the public. There are multiple ways of conducting a usability study, one of which is collecting target users' feedback with a usability questionnaire. Different groups have used different questionnaires for mHealth app usability evaluation: The commonly used questionnaires are the System Usability Scale (SUS) and Post-Study System Usability Questionnaire (PSSUQ). However, the SUS and PSSUQ were not designed to evaluate the usability of mHealth apps. Self-written questionnaires are also commonly used for evaluation of mHealth app usability but they have not been validated.

Objective: The goal of this project was to develop and validate a new mHealth app usability questionnaire.

Methods: An mHealth app usability questionnaire (MAUQ) was designed by the research team based on a number of existing questionnaires used in previous mobile app usability studies, especially the well-validated questionnaires. MAUQ, SUS, and PSSUQ were then used to evaluate the usability of two mHealth apps: an interactive mHealth app and a standalone mHealth app. The reliability and validity of the new questionnaire were evaluated. The correlation coefficients among MAUQ, SUS, and PSSUQ were calculated.

Results: In this study, 128 study participants provided responses to the questionnaire statements. Psychometric analysis indicated that the MAUQ has three subscales and their internal consistency reliability is high. The relevant subscales correlated well with the subscales of the PSSUQ. The overall scale also strongly correlated with the PSSUQ and SUS. Four versions of the MAUQ were created in relation to the type of app (interactive or standalone) and target user of the app (patient or provider). A website has been created to make it convenient for mHealth app developers to use this new questionnaire in order to assess the usability of their mHealth apps.

Conclusions: The newly created mHealth app usability questionnaire—MAUQ—has the reliability and validity required to assess mHealth app usability.

(JMIR Mhealth Uhealth 2019;7(4):e11500) doi:10.2196/11500

KEYWORDS

questionnaire design; reliability and validity; mobile apps

Introduction

Background

Mobile health (mHealth) apps can be used to perform tasks in areas such as wellness management, behavior change, health data collection, disease management, self-diagnosis, and rehabilitation as well as act as an electronic patient portal and medication reminder [1,2]. A number of research studies have been performed on mHealth apps, and the results have indicated that well-designed mHealth apps can empower patients, improve medication adherence, and reduce the cost of health care [3-6]. However, a previous study pointed out that roughly half of mHealth app users stop using some mHealth apps for various reasons such as hidden costs, high data-entry burden, and loss of interest [7]. Among these factors, high data-entry burden is clearly a usability issue, while loss of interest may also be triggered by poor usability of an app. These facts indicate the importance of good usability of mobile apps.

Among several methods for evaluation of mobile app usability, the usability questionnaire is the most frequently used because of its simplicity in terms of execution and data analysis. Ironically, although millions of mobile apps have been created and released to the public in the past decade [8], *no highly reliable usability questionnaire has been specifically designed for mobile apps.*

There have been studies on creating new mobile app *usability models* because of the special components in mobile apps such as small form factors, connectivity issues, battery issues, limited computation power, and the unique security and privacy challenges associated with highly portal mobile devices [9,10].

Studies have also tried to adjust existing usability frameworks and questionnaires, so that the models can be used for evaluation of mobile app usability [11,12]. However, in practice, many authors of the published questionnaire-based mobile app usability studies chose one of the following two options: use of well-validated usability questionnaires designed for general software systems or creation of their own usability questionnaire according to the general guidelines of usability assessment.

Both choices have their advantages and disadvantages. The advantage of the former case is that the questionnaires have been used in many other studies, and therefore, certain usability aspects of the mobile app can be reliably measured. Frequently used usability questionnaires are the System Usability Scale (SUS) and Post-Study System Usability Questionnaire (PSSUQ) [13,14]. However, for the aspects that are unique to mobile apps, these questionnaires cannot provide the desired unique information in mobile apps.

In the second case, when authors create their own usability questionnaire, they have the flexibility to cover the unique features offered by their mobile apps; however, because the major focus of those usability studies was not on validation of their self-written questionnaires, the authors typically only recruited a small number of study participants; therefore, the obtained data were not sufficient for a reliable psychometric analysis. In other words, the reliability and validity of results from the latter usability studies are questionable. Therefore, there is a need to create a reliable usability questionnaire specifically for mobile apps. This study will specifically focus on creating an mHealth app usability questionnaire.

Requirements for a New Usability Questionnaire

When designing a user-based mHealth app usability questionnaire, one has to consider the type of users and type of mobile apps.

All mHealth apps can be arranged into one of two categories according to the type of target users: patients or health care providers. The type of user is not determined by the occupation of the user, but rather by the *purpose* for using the app. Here, patients are people who use an mHealth app to maintain, improve, or manage their own health, while health care providers are people who use an mHealth app to *deliver* health care services such as medication prescription, laboratory ordering, consultation, and patient education. In the usability questionnaire, certain words need to be customizable to reflect different target users of mHealth apps.

The mHealth apps can be further grouped according to the nature of the *interaction* between the patients and health care providers in the app: *interactive* mHealth apps and *standalone* mHealth apps. In interactive mHealth apps, the app users can send and receive information from their health care providers or patients via the app. The communication between patients and providers can be synchronous or asynchronous. In standalone mHealth apps, the app users enter/collect/store health information about themselves or other people. The standalone apps may generate reminders or show a summary or details about the collected health information, but these apps do not send the data to the user's health care providers or patients. In other words, the major difference between these two types of apps is the level of interactivity. For interactive mHealth apps, there is direct interaction between patients and their health care providers. Therefore, the questionnaire for interactive apps should have statements about the quality of interaction, while the questionnaire for standalone apps may not need those statements.

Hence, *four different versions* of the mHealth app usability questionnaire are needed to evaluate the usability of mHealth apps designed for different users (patients or health care providers) and different interaction modes (interactive or standalone).

Objectives

In this study, the goal was to create a short, reliable, and customizable (via four different versions) questionnaire for assessing the usability of mHealth apps.

Methods

Process Overview

To create the desired questionnaire, we reviewed a large number of published papers about mHealth apps and then created an mHealth app usability questionnaire (MAUQ) based on the questionnaires used in these published studies, taking into consideration the uniqueness of mobile devices and mHealth apps. We then used this newly developed questionnaire in a usability study on two mHealth apps. A psychometric analysis

was performed to evaluate the reliability of the MAUQ as well as the correlation between the results from the MAUQ and those from the SUS and PSSUQ.

Usability Questionnaire Development

Step 1: Collecting Usability Questionnaire Statements From Existing Usability Questionnaires Used in mHealth App Studies

To design the desired mHealth app usability questionnaire, we first used the keywords "mobile app" and "usability" to search for published usability studies on mobile apps in PubMed, CINAHL (Cumulative Index to Nursing & Allied Health Literature), IEEE Xplore, ACM Digital Library, and InSpec [15]. From the 1271 articles obtained, we identified 125 questionnaire-based mHealth app usability studies and collected 38 individual questionnaires including well-validated questionnaires such as the SUS, PSSUQ, After Scenario Questionnaire [16], Perceived Usefulness and Ease of Use [17], Usefulness, Satisfaction, and Ease of use [18], Software Usability Measurement Inventory [19], Questionnaire for User Interaction Satisfaction [20], Computer Usability Satisfaction Questionnaire [16], Health IT Usability Evaluation Scale (Health ITUES) [21,22], and NASA Task Load Index [23] as well as a number of self-written usability questionnaires.

Step 2: Creating a Draft of the New mHealth App Usability Questionnaire and Assessing Each Statement's Relevance and Clarity

We collected 312 *unique* questionnaire statements (including similar statements with different wording) from these 38 questionnaires and arranged them into categories according to the general guidelines for usability assessment and the unique features of mobile apps. The categories included items such as usefulness, ease of learning, ease of use, effectiveness, satisfaction, interface quality, interaction quality, reliability, error messages and online support, internet connectivity, and social settings of mobile app use [9,24-26]. These 312 statements were preprocessed by removing Yes/No questions, negatively phrased questions, and redundant questions. The remaining 140 statements were then distributed to seven researchers with extensive experience in mobile app usability studies (referred to as "usability experts" in the following descriptions) to collect their feedback on the *relevance* and *clarity* of these statements on a scale of 1-4, where 1 indicates no relevance or clarity and 4 indicates high relevance or clarity. If four or more usability experts rated the relevance of a statement 1 or 2, it was removed from the questionnaire. If any one of the usability experts rated the clarity of a statement 1 or 2, the wording of the statement was adjusted. After this step, 53 statements remained in the questionnaire, 19 of which needed the wording to be adjusted.

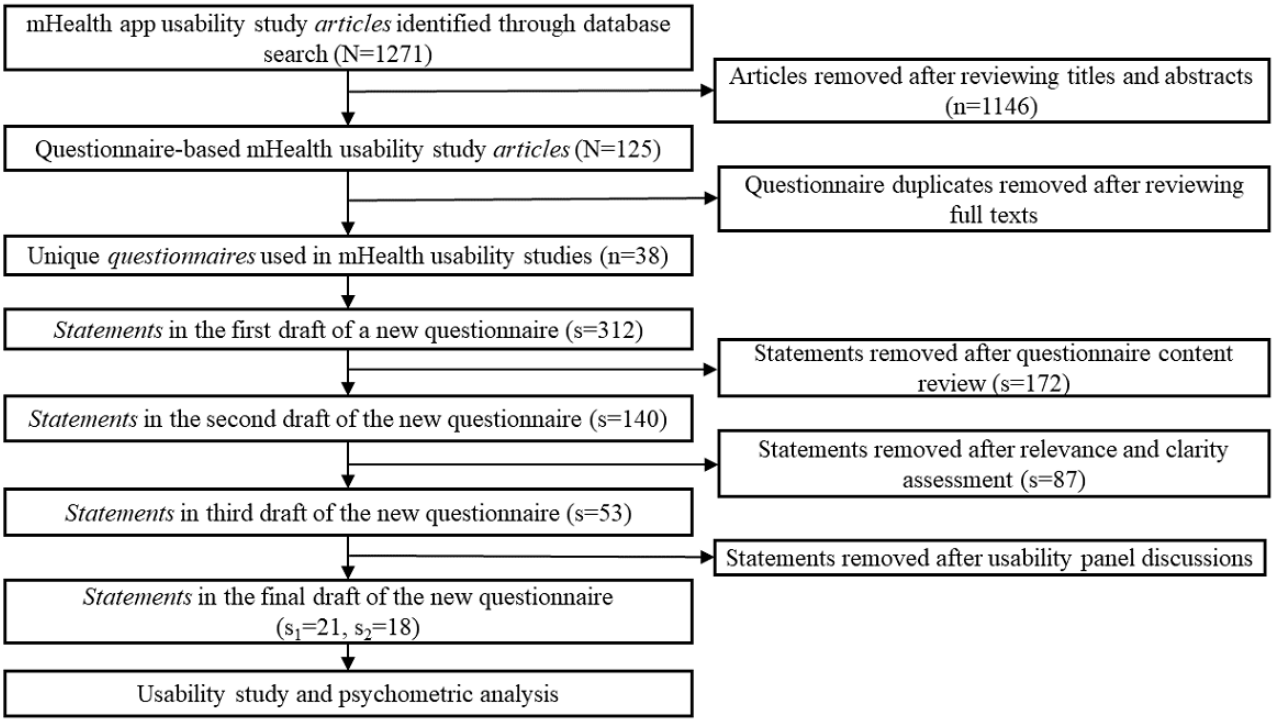
Step 3: Conducting Further Refinement on the Draft Questionnaire

After the necessary adjustments to the 19 statements, the research team held several face-to-face meetings with the seven usability experts to extensively discuss each of the 53 statements remaining. At the end of the discussion, the entire research team decided to reduce the number of statements for interactive mHealth apps to 21 and the number of statements for standalone mHealth apps to 18.

Step 4: Performing Usability Studies and Psychometric Analysis

In this step, we recruited a group of study participants to take part in usability studies using this new questionnaire (MAUQ) and two commonly used usability questionnaires (PSSUQ and SUS). The data obtained from this study were used to evaluate the reliability and validity of this new questionnaire. Figure 1 visually depicts each specific step in the development and validation of the new usability questionnaire. Details of the usability study and the psychometric analysis are presented in the following sections.

Figure 1. A flow chart of the new usability questionnaire development and validation. mHealth: mobile health; s: number of statements; s1: number of statements in the questionnaire for interactive mobile apps; s2: number of statements in the questionnaire for standalone mobile apps.



Study Design and Setting

After the new usability questionnaire was ready, a usability study on mobile apps was designed using the newly developed MAUQ and two widely used usability questionnaires, PSSUQ and SUS. Each of the study participants was asked to first use two mHealth apps: one was an interactive mobile app—iMHere 2.0 [3,27,28]—and the other was a standalone mobile app—Fitbit app version 2.36 on a 10-inch 32 GB iPad Air 2 (iOS version 10.3.2).

During the study, a general introduction to the purpose of the study and the two mHealth apps (iMHere 2.0 and Fitbit) was provided, along with a brief demo of these two mHealth apps and an explanation of the new usability questionnaire MAUQ. After the introduction, the study participants were asked to finish several tasks using these two mobile apps and provide responses to the three usability questionnaires (MAUQ, PSSUQ, and SUS) through the Web-based Qualtrics survey system (Qualtrics, Provo, UT). The demographic information of the study participants was also collected during this usability study via the Qualtrics system. Statistical analysis was performed on the collected data. This study protocol was approved by the institutional review board office at the University of Pittsburgh. A brief description about iMHere 2.0 and Fitbit apps and the tasks completed by the study participants in the usability study are provided below.

Apps Used in the Usability Study

The iMHere 2.0 system has multiple major components, including a mobile app for patients and a Web portal for clinicians. Patients can use the mobile app to perform a number of self-care tasks such as managing their medication schedules, reporting minor skin issues, performing mental health self-assessment, tracking physical activities and nutrition, sending messages to clinicians, receiving messages from clinicians, and receiving brief education in health-related topics. Clinicians can use the Web portal to view all the entered patient information and communicate with patients. Clinicians can also issue or change interventions for patients via the Web portal, which are reflected on the mobile app for patients in real time. Therefore, this mobile app in iMHere 2.0 is considered an interactive mHealth app.

The Fitbit app accompanies the Fitbit wearable device. It can display the data collected by the wearable device, such as the number of steps, heart rate, and sleep duration. The user can also enter some data that cannot be collected by the device, such as food and drink consumed in a day. Although the Fitbit app has some social media features, the user cannot directly communicate with his/her health care provider via the app. Therefore, this Fitbit app is considered a standalone mHealth app.

Tasks Performed by Subjects in the Usability Study

When using the iMHere app, study participants were asked to finish four tasks: (1) schedule a reminder for taking medication in the MyMeds module of the app; (2) report a skin problem in the Skincare module; (3) add medical history records such as medication history, allergies, immunizations, social history, and family medical history in the Personal Health Record module; and (4) adjust settings of the app, such as making the font bigger or changing the personal profile picture. The app would generate a reminder according to the indicated time point in the medication reminder schedule, and the study participant was required to respond to the reminder. After the study participants reported one skin problem, the data would show up on a Web-based clinician portal, a clinician provided a response to

the subject, and the information was shown in the message section of the iMHere 2.0 app. Study participants were required to read the message and could choose to reply to the clinician, if necessary.

When using the Fitbit app, study participants were asked to finish the following tasks: determine the numbers of steps and distances walked in the past week and past month; check their sleep history and add one new sleep log; add one new record about food eaten that day; add one new record about beverages they had drunk that day; check messages received in the app; and change the setting of goals such as the desired number of steps, number of sleep hours, and current and desired weight.

Study Participants

Participants were recruited through flyers distributed in the Greater Pittsburgh area and at the Pitt + Me website at the University of Pittsburgh [29]. Participants were screened using the following selection criteria: ability to speak fluent English; high school or higher education; age between 18 and 65 years; ability to communicate with others orally and in writing; and at least a few years of experience using smart devices such as a smartphone, tablet, or smartwatch. People who meet these selection criteria are the majority of smart device owners and mobile app users [30], and therefore, they are good candidates for providing valuable information in this usability study. All data of the eligible candidates were stored in one Excel file, and the actual study participants were randomly selected from this list.

Statistical Methods

Descriptive Analysis

Data Preprocessing

Responses to the statements on the MAUQ and PSSUQ ranged from 1 (strongly agree) to 7 (strongly disagree). The responses to the statements on the SUS ranged from 1 (strongly disagree) to 5 (strongly agree). Nine (of 128, 7.0%) study participants missed one to a few statements, and the other 119 (of 128, 93.0%) study participants provided responses to all the statements on the MAUQ. In our analysis, we used a value of 4 (indicating neither agreement nor disagreement) to fill the position of the missing data for the MAUQ.

Score Conversion and Descriptive Statistics

A descriptive analysis was first performed on the collected data to gain an understanding of the demographic characteristics of the study participants and the overall performance of the three usability questionnaires (MAUQ, PSSUQ, and SUS). The means and SD for individual statements and the entire questionnaire for the MAUQ and PSSUQ were calculated. The PSSUQ version 3 used in this study had 16 items, where items 1-6 were the first subscale (PSSUQ1), items 7-12 were the second subscale (PSSUQ2), and items 13-15 were the third subscale (PSSUQ3) [14,31]. The scores for these three subscales and the total score for the entire scale for PSSUQ were calculated. For the SUS, we used the standard score conversion procedure to convert each study participant's answers to one score between 0 and 100 [31].

Correlation Coefficient Calculation

The correlation coefficients among the scores obtained using the MAUQ, PSSUQ, and SUS were calculated, including the correlation coefficients of their subscales, if applicable, and the intersubscale correlation coefficient within the MAUQ. The former correlation coefficients were to be used to determine the criterion validity of the MAUQ, while the latter were to be used to determine the construct validity of the MAUQ [31]. In the criterion validity evaluation, a correlation coefficient as low as 0.30 or 0.40 is sufficient. In the construct validity evaluation, a low correlation is good for divergent validity and a high correlation is good for convergent validity.

Psychometric Analysis

An exploratory factor analysis (EFA) was performed on the data collected from all study participants using the MAUQ. We expected multiple factors for the MAUQ and that these factors would not be totally independent. Therefore, we used maximum likelihood as the method for factor extraction and quartimin with Kaiser Normalization for oblique rotation in EFA [32]. The factor loadings obtained in the EFA were used to determine whether each item should be included in the usability questionnaire and one specific construct. Here, 0.32 was used as a guiding value for the evaluation [33]. However, in certain cases, we overruled this value and chose to keep a statement in the questionnaire, even if the factor loadings were smaller than 0.32 or multiple factor loadings were greater than 0.32, using judgement skills gained from our extensive experience in mHealth app usability studies.

To evaluate the internal consistency of the MAUQ, we calculated the values of Cronbach alpha for the entire questionnaire and its subscales. Cronbach alpha is a commonly used measurement of internal consistency for questionnaires. For research and exploratory studies, Cronbach alpha values of 0.7-0.8 are acceptable, while a value of around 0.9 is excellent [34].

All these statistical analyses were performed using R 3.3 (R Foundation for Statistical Computing, Vienna, Austria) and IBM SPSS, version 24 (IBM Corp, Armonk, NY).

Results

Participants

In total, 128 participants were recruited from the Greater Pittsburgh area for the study. The demographic information of these participants is summarized in Table 1.

Table 1. Demographic characteristics of the study participants (N=128).

Characteristic	Value
Age (years), mean (SD)	**32.6 (13.12)**
18-28	67 (52.3)
29-50	42 (32.8)
51-65	19 (14.8)
Gender, n (%)	
Male	49 (38.3)
Female	79 (61.7)
Race, n (%)	
African American	22 (17.2)
White	81 (63.3)
Asian	22 (17.2)
Other	3 (2.3)
Education, n (%)	
High school diploma	5 (3.9)
Some college credits, no degree	28 (21.9)
Associate degree	9 (7.0)
Bachelor’s degree	45 (35.2)
Master’s degree	31 (24.2)
Professional degree	5 (3.9)
Doctoral degree	5 (3.9)
Marital status, n (%)	
Single	87 (68.0)
Married or long-term committed relationship	34 (26.6)
Divorced or separated	7 (5.5)
Living place, n (%)	
Urban	83 (64.8)
Suburban	39 (30.5)
Rural	6 (4.7)
Employment, n (%)	
Employed	92 (71.9)
Not employed	29 (21.9)
Retired or Disabled	8 (6.2)
Household income (US $ per annum), n (%)	
<10,000	24 (18.8)
10,000-50,000	52 (40.6)
50,001-100,000	27 (21.1)
>100,000	17 (13.3)
Decline to answer	8 (6.3)
Occupation, n (%)	
Student	40 (31.3)
Researcher	19 (14.8)
Administrative Personnel	18 (14.1)

Characteristic	Value
Customer Service Personnel	10 (7.8)
Other	39 (30.5)
Years of Using Mobile Devices, mean (SD)	6.86 (2.34)
Year of Using Mobile Apps, mean (SD)	6.64 (2.28)

Descriptive Data

The distribution of data from the MAUQ-based usability study was not normal according to both the Shapiro-Wilk test and the Kolmogorov-Smirnov test, where $P<.01$. Therefore, when evaluating the impact of demographic characteristics on the answers in the MAUQ, we used a nonparametric Kruskal-Wallis test. The test results indicated that none of the demographic factors (age, gender, race, education, income, marital status, living place, occupation, employment status, and experience using smart device and mobile apps) had a statistically significant impact on the answers to the individual statements or the overall score on the MAUQ ($P>.05$ in all cases).

Psychometric Analysis Results

mHealth App Usability Questionnaire for Interactive mHealth Apps (Patient Version)

Table 2 shows the factor loadings for the 21-item MAUQ designed for interactive mHealth apps. Evidently, there are three factors for the MAUQ when 0.32 is used as the cut-off value for factor loadings, with one exception—all three factor loadings for the ease of learning statement (“ *It was easy for me to learn to use the app”*) were smaller than 0.32; however, according to our experience working with usability studies and evidence from numerous other previous studies, we believe this statement is very important for a mobile app usability study and is closely related to the ease of use statement and the design of the interface. Therefore, we decided to retain this statement in the first construct of the questionnaire.

Table 2. Exploratory factor analysis results for the 21 items on the mHealth App Usability Questionnaire designed for interactive mHealth apps (overall Cronbach alpha=0.932, 21 items). Values >0.32 for each factor are italicized.

Item	Factor 1	Factor 2	Factor 3
Ease of use and satisfaction (alpha=0.895), 8 items (MAUQ_E)			
I1. The app was easy to use.	*0.633*	–0.001	0.271
I2. It was easy for me to learn to use the app.	0.234	–0.248	0.268
I3. I like the interface of the app.	*0.729*	0.002	0.010
I4. The information in the app was well organized, so I could easily find the information I needed.	*0.523*	–0.097	0.196
I5. I feel comfortable using this app in social settings.	*0.538*	0.022	0.123
I6. The amount of time involved in using this app has been fitting for me.	*0.588*	–0.041	0.222
I7. I would use this app again.	*0.800*	–0.112	–0.136
I8. Overall, I am satisfied with this app.	*0.855*	0.016	0.136
System information arrangement (alpha=0.829), 6 items (MAUQ_S)			
I9. Whenever I made a mistake using the app, I could recover easily and quickly.	0.148	0.057	*0.672*
I10. This mHealth app provided an acceptable way to receive health care services.	0.114	–0.189	*0.512*
I11. The app adequately acknowledged and provided information to let me know the progress of my action.	0.281	0.007	*0.414*
I12. The navigation was consistent when moving between screens.	0.178	–0.143	*0.466*
I13. The interface of the app allowed me to use all the functions (such as entering information, responding to reminders, viewing information) offered by the app.	0.111	–0.129	*0.473*
I14. This app has all the functions and capabilities I expect it to have.	0.180	–0.250	*0.485*
Usefulness (alpha=0.900), 7 items (MAUQ_U)			
I15. The app would be useful for my health and well-being.	0.235	*– 0.667*	–0.102
I16. The app improved my access to health care services.	0.202	*– 0.750*	–0.165
I17. The app helped me manage my health effectively.	0.308	*– 0.806*	–0.225
I18. The app made it convenient for me to communicate with my health care provider.	–0.196	*– 0.951*	0.109
I19. Using the app, I had many more opportunities to interact with my health care provider.	–0.103	*– 0.797*	0.146
I20. I felt confident that any information I sent to my provider using the app would be received.	–0.046	*– 0.541*	0.189
I21. I felt comfortable communicating with my health care provider using the app.	–0.033	*– 0.487*	0.257

The three factors correspond to three constructs, or subscales, on the MAUQ: ease of use and satisfaction (8 items, MAUQ_E), system information arrangement (6 items, MAUQ_S), and usefulness (7 items, MAUQ_U). Their Cronbach alpha values were 0.895, 0.829, and 0.900, respectively, which indicate strong internal consistency.

The correlation coefficients among the scores of the MAUQ, the scores of the MAUQ's three subscales, the scores of the PSSUQ and its subscales, and the score of the SUS are shown in Table 3. The table shows that the three subscales in MAUQ are correlated. In addition, MAUQ_S is strongly correlated with PSSUQ1, which is related to system quality; MAUQ_E is strongly correlated with PSSUQ3, which is related to interface quality; MAUQ_U is correlated with the three subscales of the PSSUQ, but not as strongly as the other two MAUQ subscales. The reason is that MAUQ_U is mainly about the usefulness of the app for *health care*, which is not covered by the PSSUQ. For overall scores, the MAUQ is strongly correlated with the PSSUQ (r=0.8448) and also correlated with the SUS (r=0.6425). The correlation between the overall scores of the PSSUQ and the SUS is 0.6703. These correlation coefficient values show the criterion validity and construct validity of the MAUQ.

Table 3. Correlation coefficients among scores from the mHealth App Usability Questionnaire, Post-Study System Usability Questionnaire, System Usability Scale, and their subscales. Italics indicate strong correlations among the scales and their subscales.

Scales	MAUQ_U[a]	MAUQ_E[b]	MAUQ_S[c]	MAUQ[d]	PSSUQ1[e]	PSSUQ2[f]	PSSUQ3[g]	PSSUQ[h]
MAUQ_E	0.5775							
MAUQ_S	0.5745	0.7139						
MAUQ	0.8573	0.8794	0.8527					
PSSUQ1	0.5608	0.7657	*0.7124*	0.7781				
PSSUQ2	0.5194	0.6078	0.7044	0.6940	0.6466			
PSSUQ3	0.5129	*0.8077*	0.6474	0.7531	0.6857	0.5943		
PSSUQ	0.6078	0.8154	0.7957	*0.8448*	0.8815	0.8956	0.8267	
SUS[i]	0.4734	0.6367	0.5677	*0.6425*	0.6402	0.5238	0.6099	*0.6703*

[a]MAUQ_U: mHealth App Usability Questionnaire – usefulness.

[b]MAUQ_E: mHealth App Usability Questionnaire - ease of use and satisfaction.

[c]MAUQ_S: mHealth App Usability Questionnaire - system information arrangement.

[d]MAUQ: mHealth App Usability Questionnaire.

[e]PSSUQ1: Post-Study System Usability Questionnaire - subscale 1.

[f]PSSUQ2: Post-Study System Usability Questionnaire - subscale 2.

[g]PSSUQ3: Post-Study System Usability Questionnaire - subscale 3.

[h]PSSUQ: Post-Study System Usability Questionnaire.

[i]SUS: System Usability Scale.

mHealth App Usability Questionnaire for Standalone mHealth Apps (Patient Version)

For the MAUQ designed for standalone apps, a similar analysis was performed and again, three factors were found. We noticed that there were a few cross-loading items (overall satisfaction, information organization, usefulness, and time) when 0.32 was used as the cut-off value of factor loadings. We chose not to remove these items, since our experience and numerous other usability studies indicate the importance of measuring overall satisfaction, information organization, time spent on the app, and usefulness of the app. In the future, we will conduct studies with larger samples to further evaluate these statements and determine whether they should be kept in the standalone mHealth app usability evaluation.

The results for the EFA and correlation coefficient calculation are shown in Tables 4 and 5. Again, the values of Cronbach alpha in the overall questionnaire and the three subscales show strong internal consistency in the questionnaire. The correlation coefficients among MAUQ, PSSUQ, SUS, and their subscales also indicate the criterion validity and construct validity of the MAUQ.

Table 4. Exploratory factor analysis results for the 18 items on the mHealth App Usability Questionnaire designed for standalone mHealth apps (overall Cronbach alpha=0.914). Values >0.32 for each factor are italicized.

Item	Factor 1	Factor 2	Factor 3
Ease of use (alpha=0.847), 5 items (MAUQ_E)			
S1. The app was easy to use.	*0.788*	0.121	–0.064
S2. It was easy for me to learn to use the app.	*0.811*	0.090	–0.058
S3. The navigation was consistent when moving between screens.	*0.708*	0.020	0.090
S4. The interface of the app allowed me to use all the functions (such as entering information, responding to reminders, viewing information) offered by the app.	*0.640*	–0.055	0.271
S5. Whenever I made a mistake using the app, I could recover easily and quickly.	*0.417*	–0.019	0.278
Interface and satisfaction (alpha=0.908), 7 items (MAUQ_I)			
S6. I like the interface of the app.	0.223	*0.841*	–0.247
S7. The information in the app was well organized, so I could easily find the information I needed.	0.525	*0.464*	–0.017
S8. The app adequately acknowledged and provided information to let me know the progress of my action.	0.147	*0.450*	0.198
S9. I feel comfortable using this app in social settings.	–0.020	*0.508*	0.186
S10. The amount of time involved in using this app has been fitting for me.	0.321	*0.515*	0.092
S11. I would use this app again.	-0.046	*0.680*	0.251
S12. Overall, I am satisfied with this app.	0.323	*0.442*	0.301
Usefulness (alpha=0.717), 6 items (MAUQ_U)			
S13. The app would be useful for my health and well-being.	–0.121	0.354	*0.584*
S14. The app improved my access to health care services.	0.050	0.021	*0.390*
S15. The app helped me manage my health effectively.	0.099	–0.002	*0.679*
S16. This app has all the functions and capabilities I expected it to have.	0.210	0.178	*0.379*
S17. I could use the app even when the Internet connection was poor or not available.	–0.014	0.025	*0.526*
S18. This mHealth[a] app provided an acceptable way to receive health care services, such as accessing educational materials, tracking my own activities, and performing self-assessment.	0.068	–0.016	*0.326*

[a]mHealth: mobile health.

Table 5. Correlation coefficients among scores from the mHealth App Usability Questionnaire, Post-Study System Usability Questionnaire, System Usability Scale, and their subscales. Italics indicate strong correlations among the scales and their subscales.

Scale	MAUQ_U[a]	MAUQ_E[b]	MAUQ_I[c]	MAUQ[d]	PSSUQ1[e]	PSSUQ2[f]	PSSUQ3[g]	PSSUQ[h]
MAUQ_E	0.5357							
MAUQ_I	0.5582	*0.7526*						
MAUQ	0.8053	0.8658	0.9122					
PSSUQ1	0.4684	0.8053	*0.8203*	0.8110				
PSSUQ2	0.5887	*0.7285*	0.7423	0.7961	0.8087			
PSSUQ3	0.5143	0.6411	0.7954	0.7654	0.7697	0.7604		
PSSUQ	0.5660	0.7979	0.8479	*0.8585*	0.9451	0.9373	0.8782	
iSUS	0.3832	0.7322	0.7362	*0.7168*	0.8513	0.7491	0.7476	*0.8523*

[a]MAUQ_U: mHealth App Usability Questionnaire - usefulness.

[b]MAUQ_E: mHealth App Usability Questionnaire - ease of use and satisfaction.

[c]MAUQ_I: mHealth App Usability Questionnaire - interface and satisfaction.

[d]MAUQ: mHealth App Usability Questionnaire.

[e]PSSUQ1: Post-Study System Usability Questionnaire - subscale 1.

[f]PSSUQ2: Post-Study System Usability Questionnaire - subscale 2.

[g]PSSUQ3: Post-Study System Usability Questionnaire - subscale 3.

[h]PSSUQ: Post-Study System Usability Questionnaire.

[i]SUS: System Usability Scale.

Textbox 1. Adjusted statement on the mHealth App Usability Questionnaire for use of mHealth apps by health care providers.

Statements on the mHealth App Usability Questionnaire adjusted for interactive mobile apps designed for health care providers:

I10. This app provided an acceptable way to *deliver* health care services.

I15. The app would be useful for my health *care practice*.

I16. The app improved my access to *delivering* health care services.

I17. The app helped me manage my *patients'* health effectively.

I18. The app made it convenient for me to communicate with my *patients*.

I19. Using the app, I had many more opportunities to interact with my *patients*.

I20. I felt confident that any information I sent to my *patients* using the app will be received.

I21. I felt comfortable communicating with my *patients* using the app.

Statements on the mHealth App Usability Questionnaire adjusted for standalone apps designed for health care providers:

S13. The app would be useful for my *health care practice*.

S14. The app improved my access to *delivering* health care services.

S15. The app helped me manage my *patients'* health effectively.

S18. This mHealth app provided an acceptable way to *deliver* health care services, such as accessing educational materials, tracking my own activities, and performing self-assessment.

mHealth App Usability Questionnaire for Interactive and Standalone mHealth Apps (Provider Version)

In the versions of the MAUQ questionnaires described in the previous sections, some statements were generic, while some statements were specific to mHealth apps designed for patients. When evaluating the usability of mHealth apps designed for health care providers, some statements need to be slightly modified, for instance, changing the statement from *receiving* health care services to *delivering* health care services, and from interacting with *health care providers* to interacting with *patients*. Further evaluation may be needed for the health care provider version, since the two mHealth apps in this study were both for patients. Textbox 1 shows the statements on the MAUQ adjusted to evaluate mHealth apps designed for providers. The adjusted words are in italicized font.

A Website for the mHealth App Usability Study Using the mHealth App Usability Questionnaire

To make it convenient for others to utilize the MAUQ in their mHealth app usability studies, we created a website [35] that includes the four versions of the MAUQ (interactive or standalone, patient or provider), and some optional demographic questions and open-ended questions typically used in usability

studies. Anyone who wants to use this usability questionnaire can create an account on the website, create a customized usability questionnaire by simply selecting versions of the MAUQ and optional demographic questions, and add more demographic and open-ended questions if needed. The data collected in a usability study are stored on a secure Web server. The user can view a brief summary of the data collected in his/her usability study on the website and download the collected dataset to a local computer for further analysis.

Discussion

Principal Results

The aim of this study was to develop and validate a new mHealth app usability questionnaire—the MAUQ. This new questionnaire is highly reliable (reflected in the Cronbach alpha value and correlation with PSSUQ and SUS), as we used information gathered and summarized from many previous studies and combined the experience of several usability study experts to create it. Below is a summary of the unique contribution of our work.

First, the development of the MAUQ was based on a number of usability questionnaires used in mHealth app usability assessment described in published journal articles. Some of these questionnaires are well validated, and therefore, many statements in these questionnaires have been approved in numerous studies, making the use of a great resource a better choice than creation of a new questionnaire completely from scratch.

Second, the draft of the MAUQ was created considering components specific to mobile devices and mHealth apps. Some proposed mHealth app usability models were also consulted while questionnaire statements were being selected. This makes the MAUQ an mHealth app–specific questionnaire.

Third, a group of usability experts discussed the draft of the questionnaire and built the draft of the MAUQ. These usability experts had extensive experience in the study of mHealth app usability, and therefore, their knowledge and experience were integrated into the MAUQ during the discussion and questionnaire statement selection.

Fourth, four different versions of the MAUQ were created for four different scenarios (interactive app for patients, interactive app for health care providers, standalone app for patients, and standalone app for health care providers). This makes it feasible for others to easily choose one version for their usability study according to the desired context.

Fifth, this newly built questionnaire was tested in a usability study with 128 study participants and two mHealth apps—one, standalone (Fitbit) and the other, interactive (iMHere 2.0). This sample size is much bigger than typical usability studies; hence, it was feasible for an EFA. The data analysis results indicated that the MAUQ has strong construct validity and criterion validity, and the internal consistency of the three subscales and the entire questionnaire is high.

Sixth, the performance of the MAUQ was compared with two frequently used usability questionnaires—PSSUQ and SUS. The correlation coefficients among the MAUQ, PSSUQ, and SUS were high because they all were used to measure the overall usability of the mHealth apps. Two subscales of the MAUQ were also highly correlated with two subscales of the PSSUQ measuring similar aspects of usability. For the MAUQ subscale, with features unique to mHealth apps, the correlation was not that strong with any of the PSSUQ subscales, which is expected. In other words, the MAUQ can reliably measure usability of mHealth apps.

Comparison With Prior Work

According to our knowledge, there is no single highly reliable usability questionnaire specifically designed for mHealth apps. Researchers from other teams have investigated usability models for mobile apps and also tried to modify existing usability questionnaires for use in mobile app usability studies. However, to date, none of these have been widely adopted by other researchers.

The 21 statements in the final version of the MAUQ for interactive mobile apps were compared with the ones in several other frequently used usability questionnaires. The result indicated that 5 statements in the MAUQ were highly similar to the ones in multiple other questionnaires; 10 statements in the MAUQ were similar to the ones in a few other questionnaires, but more specific to mobile app–based health care activities; and the remaining 6 statements were unique to the MAUQ.

A few years ago, a mobile app rating scale (MARS) was created for assessing the *quality* of mHealth apps [36], which is broader than usability. The target users of MARS were experts in one field (researchers, clinicians, and other professionals) who wanted to identify high-quality mobile apps in that field. Shortly thereafter, a simpler user version of MARS (uMARS) was created for the general population for the same purpose [37]. Both MARS and uMARS included some usability components such as engagement, functionality, and information quality; however, they were not specifically designed to study usability with end users.

One recently published work adjusted an existing usability questionnaire to apply it for mHealth apps [12]. In that study, the authors evaluated the psychometric properties of the existing usability questionnaire designed for health information technology systems by analyzing usability study data obtained from a group of patients with HIV. The authors indicated that this customizable health information technology usability questionnaire, the Health IT Usability Evaluation Scale (Health ITUES), worked well in an mHealth app usability study. However, it is tricky to customize the statements in a usability questionnaire, since the change may impact the responses of the study participants. In the original Health ITUES and the modified version for mHealth apps, several statements need to be customized by the questionnaire user [12,22]. Although customization makes it convenient to measure unique features offered by an individual health information technology system or mHealth app, it also creates challenges for questionnaire users, especially ones who are not experienced in questionnaire development, and this, in turn, will make the reliability and validity of the customized usability questionnaire questionable.

In the MAUQ, we provide four versions for two types of target users (patients and providers) and two major types of mHealth apps (standalone and interactive), which allows MAUQ users to choose the version that fits their needs. Moreover, the website created for the MAUQ makes the administration of the usability questionnaire easy. Researchers and mobile app developers can easily use the website to perform a quick usability/feasibility study with a small number of participants on the prototype of their apps, to conduct a multistage usability study during the process of the app development, or to have a large-scale intensive usability study with many participants over a long period of time. In addition, all the collected data from these usability studies are stored securely on the website for viewing and downloading.

Limitations

The study was performed in the greater Pittsburgh area, but approximately one-third of the study participants were undergraduate and graduate students from different states, making the conclusions obtained in this study generalizable within certain limitations.

In this study, we had more female participants. All the participants were randomly selected from eligible candidates, but there were significantly more female mobile app users who expressed interest in this study. In other words, this study population was not an exact reflection of the US population; however, it did indicate the distribution of mobile app users who were interested in using mHealth apps to take care of their health.

The sample size (N=128) was considered sufficient for EFA, since the ratio between the number of study participants and the number of statements (n=21) was greater than 5. Since EFA is a large-sample procedure, a larger sample size provides more reliable results, for instance, to determine whether the items with low factor loading should be kept in the MAUQ for interactive apps and whether the cross-loading items should be removed from the MAUQ for standalone apps. For this reason, we will use this questionnaire in our usability studies to collect more data in the future. We also created a website to make it convenient for others to use this new questionnaire in their usability studies. We encourage others who use the MAUQ in their usability studies to share their collected data with us, so that we can perform further psychometric analysis on this questionnaire to improve its use for mHealth app usability evaluation.

In this study, we required that the study participants had certain characteristics, such as a high school or higher education, age between 18 and 65 years, and some experience using mHealth apps. This excluded some potential participants, for instance, people older than 65 years or people with a low education level. However, since the purpose of this study was to evaluate the newly created usability questionnaire, the participants selected were representative of the majority of mHealth app users and ones who could provide the most reliable assessment on the questionnaire. A different usability study method may be used for populations not included in this study.

The usability study in this work was performed on two mobile apps: iMHere 2.0 and Fitbit. It is possible that the results might have been *slightly* different for different mHealth apps. This can be assessed in the future after other research teams adopt the MAUQ in their usability studies with other mHealth apps.

Only the patient versions of the MAUQ were tested in this study. Therefore, although we believe the differences in the versions for patients and the versions for providers are not significant, the versions of MAUQ for health care providers were not explicitly evaluated. In the future, we may identify appropriate interactive and standalone mHealth apps designed for health care providers and perform the usability study again to evaluate the versions for providers. We may also utilize usability study data from other research teams to perform further assessment on the versions for providers.

Acknowledgments

We would like to thank Wayan Pulantara, Zakiy Firdaus, and Gede Pramana for reviewing the drafts of the usability questionnaire and evaluating the content relevance and clarity of the statements in the questionnaire. This research was funded in part by grants from the National Institutes of Health (NIH; 1 R43 DC04246-01 and 1 R43 DC04472-01) and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR; 90RE5018, 90DP0064, and 90DPGE0002). The contents of this article do not represent the views of the NIH, the NIDILRR, or the US government.

Conflicts of Interest

None declared.

References

1. Kao C, Liebovitz DM. Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
2. Roess A. The Promise, Growth, and Reality of Mobile Health - Another Data-free Zone. N Engl J Med 2017 Nov 23;377(21):2010-2011. [doi: 10.1056/NEJMp1713180] [Medline: 29116869]
3. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE, McCue MP. iMHere: A Novel mHealth System for Supporting Self-Care in Management of Complex and Chronic Conditions. JMIR Mhealth Uhealth 2013 Jul 11;1(2):e10 [FREE Full text] [doi: 10.2196/mhealth.2391] [Medline: 25100682]

4. Fairman AD, Dicianno BE, Datt N, Garver A, Parmanto B, McCue M. Outcomes of Clinicians, Caregivers, Family Members and Adults with Spina Bifida Regarding Receptivity to use of the iMHere mHealth Solution to Promote Wellness. Int J Telerehabil 2013 Jun;5(1):3-16 [FREE Full text] [doi: 10.5195/ijt.2013.6116] [Medline: 25945209]
5. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Perceptions and experiences of heart failure patients and clinicians on the use of mobile phone-based telemonitoring. J Med Internet Res 2012;14(1):e25 [FREE Full text] [doi: 10.2196/jmir.1912] [Medline: 22328237]
6. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]
7. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015 Nov;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
8. Research2Guidance. 2017. mHealth App Economics 2017 - Current Status and Future Trends in Mobile Health URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2019-04-08] [WebCite Cache ID 77UJZQcgu]
9. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1. [doi: 10.1186/2194-0827-1-1]
10. Ntawanga F, Calitz A, Barnard L. A Context-Aware Model to Improve Usability of Information Display on Smartphone Apps for Emerging Users. Afr J Inf Syst 2015;7(4):30-55.
11. Moumane K, Idri A, Abran A. Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards. Springerplus 2016;5:548 [FREE Full text] [doi: 10.1186/s40064-016-2171-z] [Medline: 27190747]
12. Schnall R, Cho H, Liu J. Health Information Technology Usability Evaluation Scale (Health-ITUES) for Usability Assessment of Mobile Health Technology: Validation Study. JMIR Mhealth Uhealth 2018 Jan 05;6(1):e4 [FREE Full text] [doi: 10.2196/mhealth.8851] [Medline: 29305343]
13. Brooke J. SUS: A quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. Usability Evaluation in Industry. London: Taylor and Francis; 1996:189-194.
14. Lewis JR. Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. International Journal of Human-Computer Interaction 2002 Sep;14(3-4):463-488. [doi: 10.1080/10447318.2002.9669130]
15. Zhou L, Bao J, Parmanto B. Systematic Review Protocol to Assess the Effectiveness of Usability Questionnaires in mHealth App Studies. JMIR Res Protoc 2017 Aug 01;6(8):e151. [doi: 10.2196/resprot.7826]
16. Lewis J. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. International Journal of Human-Computer Interaction 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]
17. Davis F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
18. Lund A. Measuring Usability with the USE Questionnaire. Usability Interface 2001;8(2):3-6.
19. Coleman N. SUMI (Software Usability Measurement Inventory) as a knowledge elicitation tool for improving usability. Ireland: Department of Applied Psychology, University College Cork; 1993.
20. Chin J, Diehl V, Norman K. Development of an instrument measuring user satisfaction of the human-computer interface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM; 1988 Presented at: CHI'88; May 15 - 19; Washington DC p. 213-218.
21. Brown W, Yen P, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
22. Yen P, Sousa KH, Bakken S. Examining construct and predictive validity of the Health-IT Usability Evaluation Scale: confirmatory factor analysis and structural equation modeling results. J Am Med Inform Assoc 2014 Oct;21(e2):e241-e248 [FREE Full text] [doi: 10.1136/amiajnl-2013-001811] [Medline: 24567081]
23. Hart S, Staveland L. Development of NASA-TLX (Task Load Index): Results of empiricaltheoretical research. In: Hancock PA, Meshkati N, editors. Human Mental Workload. Amsterdam: North Holland Press; 1988.
24. Parmanto B, Lewis, Jr. AN, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehab 2016 Jul 01;8(1):3-10. [doi: 10.5195/ijt.2016.6196]
25. Nielsen J. Usability Engineering. Boston: Academic Press Inc; 1993.
26. Wang C, Huang C. A study of usability principles and interface design for mobile e-books. Ergonomics 2015;58(8):1253-1265. [doi: 10.1080/00140139.2015.1013577] [Medline: 25680001]
27. Bendixen RM, Fairman AD, Karavolis M, Sullivan C, Parmanto B. A User-Centered Approach: Understanding Client and Caregiver Needs and Preferences in the Development of mHealth Apps for Self-Management. JMIR Mhealth Uhealth 2017 Sep 26;5(9):e141. [doi: 10.2196/mhealth.7136]
28. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE. Development of mHealth system for supporting self-management and remote consultation of skincare. BMC Med Inform Decis Mak 2015 Dec 30;15:114. [doi: 10.1186/s12911-015-0237-4]
29. Pitt + Me. URL: https://pittplusme.org/ [accessed 2019-04-08]

30. Pew Research Center. Pew Research Center. 2018. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2018-02-12] [WebCite Cache ID 6xDIpUN2z]
31. Sauro J, Lewis J. Quantifying the User Experience: Practical Statistics for User Research. Burlington, MA: Morgan Kaufmann; 2019.
32. Costello AB, Osborn JW. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment Research & Evaluation 2005;10(7):1-9.
33. Tabachnick B, Fidell L. Using Multivariate Statistics. Boston: Allyn and Bacon; 2001.
34. Landauer T. Behavioral research methods in human-computer interaction. In: Helander M, Landauer TK, Prabhu P, editors. Handbook of Human-Computer Interaction. 2nd Edition. Elsevier, Inc: Amsterdam, Netherlands; 1997:203-227.
35. PITT Usability Questionnaire For Telehealth Systems (TUQ) and Mobile Health Apps (MAUQ). URL: https://ux.hari.pitt.edu/app/about [accessed 2019-04-08] [WebCite Cache ID 77UHshH1o]
36. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
37. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]

Abbreviations

CINAHL: Cumulative Index to Nursing & Allied Health Literature
EFA: exploratory factor analysis
Health ITUES: Health IT Usability Evaluation Scale
MARS: mobile app rating scale
MAUQ: mHealth App Usability Questionnaire
mHealth: mobile health
PSSUQ: Post-Study System Usability Questionnaire
SUS: System Usability Scale

Edited by G Eysenbach; submitted 05.07.18; peer-reviewed by A Fairman, T Fazzino, S Kagen MD, J Apolinário-Hagen; comments to author 06.10.18; revised version received 19.01.19; accepted 30.03.19; published 11.04.19.

Please cite as:
Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B
The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study
JMIR Mhealth Uhealth 2019;7(4):e11500
URL: http://mhealth.jmir.org/2019/4/e11500/
doi:10.2196/11500
PMID:30973342

Viewpoint

The Iterative Convergent Design for Mobile Health Usability Testing: Mixed Methods Approach

Meshari F Alwashmi[1], BN, MscMed (Clin Epi); John Hawboldt[1], BSP, PharmD; Erin Davis[1], BSc, Pharm, PharmD; Michael D Fetters[2], MA, MD, MPH

[1]Memorial University, St John's, NL, Canada

[2]Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Meshari F Alwashmi, BN, MscMed (Clin Epi)
Memorial University
230 Elizabeth Ave
St John's, NL, A1B 3X9
Canada
Phone: 1 7096910728
Email: mfa720@mun.ca

Abstract

Although patients express an interest in using mobile health (mHealth) interventions to manage their health and chronic conditions, many current mHealth interventions are difficult to use. Usability testing is critical for the success of novel mHealth interventions. Researchers recognize the utility of using qualitative and quantitative approaches for usability testing, but many mHealth researchers lack the awareness of integration approaches from advances in mixed methods research that can add value to mHealth technology. As efficient usability testing proceeds iteratively, we introduce a novel mixed methods design developed specifically for mHealth researchers. The *iterative convergent mixed methods design* involves simultaneous qualitative and quantitative data collection and analysis that continues cyclically through multiple rounds of mixed methods data collection and analysis until the mHealth technology under evaluation is found to work to the satisfaction of the researcher. In cyclical iterations, early development is more qualitatively driven but progressively becomes more quantitatively driven. Using this design, mHealth researchers can leverage mixed methods integration procedures in the research question, data collection, data analysis, interpretation, and dissemination dimensions. This study demonstrates how the iterative convergent mixed methods design provides a novel framework for generating unique insights into multifaceted phenomena impacting mHealth usability. Understanding these practices can help developers and researchers leverage the strengths of an integrated mixed methods design.

(JMIR Mhealth Uhealth 2019;7(4):e11656) doi:10.2196/11656

KEYWORDS

mHealth; mixed methods; usability; eHealth; methods

Introduction

Published studies indicate that mobile health (mHealth) interventions are beneficial for patients across various diseases and age groups [1-4]. Academics and clinicians have an increasing interest in harnessing these mHealth interventions to improve health outcomes. Although patients express an interest in using mHealth to manage their health and chronic conditions, many current mHealth interventions are difficult to use [5]. Hence, developers of mHealth need efficient and effective approaches for development, but usability research methodology remains in a relatively nascent stage of development [6]. Usability testing is critical for the success of novel mHealth interventions. Although researchers recognize the utility of using qualitative and quantitative approaches for usability testing, many mHealth researchers lack the awareness of integration approaches from advances in mixed methods research (MMR) [7] that can add value to mHealth technology.

This paper advances the existing literature about the combined use of qualitative and quantitative research for mHealth by advancing a specific, integrated approach to mixed methods design appropriate to mHealth. When using qualitative and quantitative procedures without integration, researchers miss the opportunity for added value. Mixed methods methodologists express this as 1+1=2, as the quantitative and qualitative procedures are conducted as 2 independent studies with no particular synergy [8]. By using integrated procedures identified

in the field of MMR, researchers can aspire for and achieve added value, as expressed by 1+1=3 [8,9].

The purpose of this paper is to articulate and illustrate the features of an iterative convergent mixed methods design. As efficient usability testing proceeds iteratively, we introduce a novel mixed methods design developed specifically for mHealth researchers. It offers a novel framework to generate unique insights into multifaceted phenomena related to mHealth usability. Understanding these practices can help developers and researchers leverage the strengths of an integrated mixed methods design.

Background

Mobile Health

Effective health care strategies are required to ensure the right patient receives the right treatment at the right time. Advancements in mobile phones and tablets have led to the emergence of mHealth. The Global Observatory for eHealth of the World Health Organization defines mHealth as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices" [10]. Recent advances allow seamless integration between smartphones and medical devices. This integration enables smartphones to store and analyze objective measurements such as heart rate, lung volume, and medication adherence. Advancements in machine learning and artificial intelligence have the potential to use these measurements, in combination with data collected via smartphones, to improve our understanding of disease etiology [11,12].

The significance of mHealth is highlighted by its ability to deliver timely care over distance to manage diseases. It is particularly important for rural areas with limited access to health care [13,14]. Moreover, mHealth strategies can enhance treatment outcomes while mitigating health care costs [15,16]. Hayes et al [16] illustrated why mHealth could reduce physician visits, resource consumption, and emergency room visits. The literature continues to evolve on applications of mHealth. For example, several published studies indicate that mHealth interventions are beneficial for patients across various diseases and age groups [1-4]. However, research on the development and usability methodology of such interventions remains in a relatively early stage.

Human-Centered Design

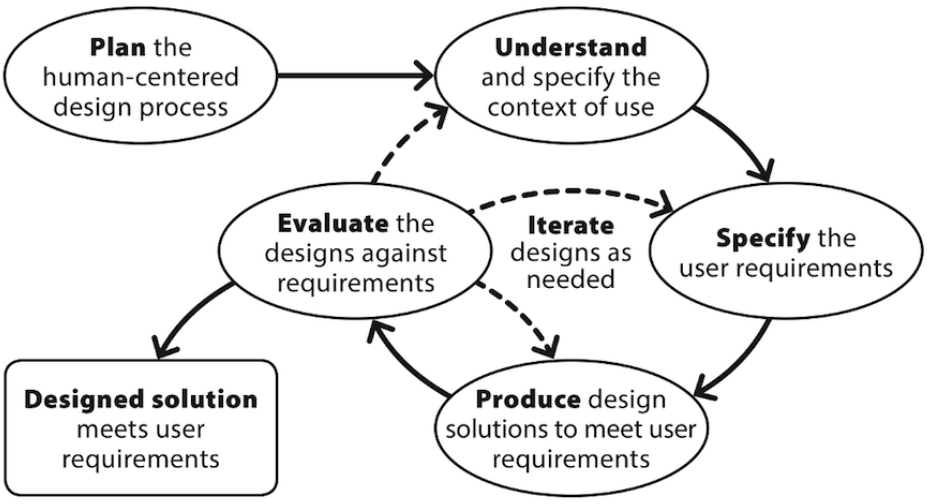
The International Organization for Standardization (ISO) 9241-210 standard defines human-centered design (HCD) as "an approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques" [17]. The ISO uses the term HCD instead of user-centered design as it" addresses impacts on a number of stakeholders, not just those typically considered as users" [17]. However, in practice, these terms are often used synonymously.

HCD has 4 defined activity phases: (1) identify the user and specify the context of use, (2) specify the user requirements, (3) produce design solutions, and (4) evaluate design solutions against requirements. The process model of HCD as defined in ISO 9241-210 is illustrated in Figure 1.

Researchers advocate for involving patients during development who are going to use the mHealth intervention to meet the patient's needs and facilitate successful uptake. Testing mHealth interventions with patients reveals preferences and concerns unique to the tested population [5,18,19]. Developing an mHealth intervention with insights from stakeholders will potentially improve the process and outcome of mHealth interventions. The main goal of HCD is to increase the usability of mHealth technology.

This study offers an in-depth account of the HCD's fourth activity phase, evaluate design solutions against requirements. This clarifies that this framework intends to focus on usability testing as one component of the more extensive design process.

Figure 1. Human-centered design activity phases (ISO, 2010).



Usability

The ISO defines usability as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [20]. Although this definition was published in 1998, it has been updated in 2018 without any changes to the core concepts. The definition is widespread and generalizable [21,22].

mHealth involves the interaction between multiple user groups through a system. As a result, the usability aspect is vital for the effective, efficient, and satisfactory use of mHealth interventions. Although patients express an interest in using mHealth to manage their health and chronic conditions, many mHealth interventions are not easy to use [5,23]. Difficulty in using an mHealth intervention may limit the user retention rate. A high dropout rate is one of the most significant barriers to mHealth adoption [24,25]. The majority of mHealth app publishers (83%) have less than 10,000 users who have used the app at least once a month [26]. These numbers are discouraging as according to a 2018 estimate, the average mHealth app costs $425,000 to develop [26]. By putting a more significant emphasis on usability, iterative improvements can reduce costs and enhance the long-term use and adoption of mHealth interventions [27-29].

Researchers recommend frequent and iterative usability testing to respond to users' preferences, technical issues, and shortcomings [18,30,31]. It is also important to ensure that errors in understanding or using the intervention are addressed before testing the intervention in an efficacy trial [32]. A systematic review investigated the usability evaluation processes described in 22 studies related to mHealth applications [33]. The results suggest that the adoption of automated mechanisms could improve usability and stress the importance of adapting health applications to users' need [33]. Including insights from key users of mHealth has the potential to improve the process and the outcome of the intervention [34].

Contemporary iterative development methods, such as prototyping, reduce the challenges that evolve during the development lifecycle [35]. Prototyping is defined as creating a simulation of the final mHealth technology that is used for testing before launch. Furthermore, researchers suggest including patients who are going to use the mHealth interventions to assist in the development of the intervention [5,18,19,33].

The ISO 9241-11 established usability standards [20]. These standards provide a measure of patients' experienced usability. It focuses on effectiveness, efficiency, and satisfaction [20]. It is easier to quantify effectiveness and efficiency compared with satisfaction. Brooke developed the System Usability Scale (SUS) [36] and noted that "if there is an area in which it is possible to make more generalized assessments of usability, which could bear cross-system comparison, it is the area of subjective assessments of usability" [36]. Thus, the SUS was developed to quantify satisfaction (users' subjective reactions to using the system) [36]. The SUS is an affordable and effective tool for assessing the usability of products [36]. It contains 10 statements that are answered on a 5-point Likert scale (Multimedia Appendix 1). Although this scale was developed in 1996, it is relevant and applicable to current research because it is short and easy to use. Many contemporary mHealth studies have been successful in combining both ISO and SUS to assess usability [37-39]. Table 1 describes the ISO 9241-11 usability constructs.

Table 1. Usability constructs and descriptions.

Constructs[a]	Metrics	Description
Effectiveness	Time to learn and use	Time to read the scenarios and to begin performing tasks
	Data entry time	Time to enter the data necessary for the execution of a task
	Tasks time	Time to accomplish given tasks
	Response time	Time of having the response to the requested information
	Time to install	Installation time of applications or its update
Efficiency	Number of errors	Number of errors made while reading scenarios and during the task execution
	Completion rate	The percentage of participants who correctly complete and achieve the goal of each task
Satisfaction	Usability score	The System Usability Questionnaire

[a]Adapted from Moumane et al [40].

Mixed Methods Research

MMR is gaining popularity and acceptance across disciplines and the world [41]. It draws from multiple scientific traditions and disciplinary backgrounds. MMR is defined as "the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches for the broad purposes of breadth and depth of understanding and corroboration" [42]. MMR combines both closed-ended response data (quantitative) and open-ended personal data (qualitative) [41].

Although quantitative research historically has predominated in health sciences research, many contemporary phenomena in health care are difficult, if not impossible, to measure using quantitative methods alone [43]. The goal of qualitative research is to produce a deep understanding of a phenomenon. It can also be used to generate a hypothesis regarding a phenomenon, its precursors, and its consequences [44]. When the study phenomenon of interest is multifaceted and complex, a mixed methods approach is appropriate [43]. The National Institutes of Health best practices guideline and many mixed methods researchers advise distinguishing the quantitative purpose, the

qualitative research questions, and the mixed methods questions [45]. Consequently, MMR can capitalize on the strengths of both methods, the depth of qualitative research and the breadth of quantitative research. The resulting mixed data can be integrated to balance the strengths and limitations of either method to provide a more comprehensive understanding under potentially complementary sources of evidence [43].

Understanding the principles and practices of integration is essential for leveraging the strengths of MMR. Fetters and Molina-Azorin [7] defined *integration* as the linking of qualitative and quantitative approaches and dimensions together to create a new whole or a more holistic understanding than achieved by either alone. Fetters et al examined vital integration principles and practices in MMR [46]. They provide approaches to integrating both research procedures and data in the design, methods, interpretation, and reporting dimensions of research [46]. Table 2 provides the relevant dimensions of MMR integration and illustrates how researchers can integrate those dimensions. These dimensions are relevant to mHealth, and additional information about MMR dimensions is explained elsewhere [7]. Through increasingly sophisticated approaches, MMR is viewed as an opportunity to address the highly complex, compelling, and even *wicked* research problems facing researchers in the health and social sciences [47]. Investigation of novel mHealth technologies is an important example of a highly complex research challenge that can benefit from a systematic mixed methods approach.

Table 2. Relevant dimensions of the mixed methods research integration.

Integration dimensions[a]	Mixed methods researchers integrate by
Rationale dimension	Citing a rationale for conducting an integrated mixed methods research study (eg, offsetting strengths and weaknesses, comparing, complementing or expanding, developing or building, and promoting social justice)
Study purpose, aims, and research questions dimension	Composing an overarching mixed methods research purpose and stating qualitative, quantitative, and mixed methods aims or multiple mixed methods aims with quantitative aims and qualitative questions
Research design dimension	Scaffolding the work in core (eg, convergent, exploratory sequential, and explanatory sequential), advanced (eg, intervention, case study, evaluation, and participatory), or emergent designs.
Sampling dimension	Sampling through the type, through the relationship of the sources of the qualitative and quantitative data (eg, identical sample, nested sample, separate samples, and multilevel samples), and through the timing (eg, same or different periods for collection of the qualitative and quantitative data)
Data collection dimension	Collecting both types of data with an intent relative to the mixed methods research procedures (eg, comparing, matching, diffracting, expanding, constructing a case, connecting, building, generating and validating a model, or embedding).
Data analysis dimension	Analyzing both types of data using intramethod analytics (eg, analyzing each type of data within the respective qualitative and quantitative methods and core integration analytics), using 1 or more core mixed methods analysis approach (eg, by following a thread, spiraling, and back-and-forth exchanges), or employing advanced mixed methods analysis (eg, qualitative to quantitative data transformation, quantitative to qualitative data transformation, creating joint displays, social network analysis, qualitative comparative analysis, repertory grid/other scale development techniques, geographic information systems mapping techniques, and iterative and longitudinal queries of the data).
Interpretation dimension	Interpreting the meaning of mixed findings (eg, where there are related data and drawing metainferences or conclusions based on interpreting the qualitative and quantitative findings) and examining for the fit of the 2 types of data (eg, confirmation, complementarity, expansion, or discordance). When the results conflict with each other, using procedures for handling the latter including reconciliation, initiation, bracketing, and exclusion.

[a]Adapted from Fetters and Molina-Azorin [7].

Importance of Mixed Methods in Usability Testing

Usability is a complex phenomenon. It is challenging to investigate usability comprehensively using only quantitative methods or qualitative methods in isolation, so-called *monomethod* approaches [6]. The alternative to using a monomethod approach is using diverse methods to generate a complete picture and reveal hidden patterns and novel relationships between variables and concepts [43]. To identify and resolve usability issues, various researchers emphasize the importance of using multiple methods and sources of data [18,48]. Although many studies collect both quantitative and qualitative data to test usability [49-51], mHealth researchers could benefit from advances being made for integration in mixed methods studies [46,52].

Despite the recognized and intuitive value of using mixed methods for mHealth usability testing, mixed methodologists have yet to articulate specific designs that guide the development and testing of mHealth interventions. A core MMR study design that is attractive for usability testing is the convergent design [52]. Also called by some authors as a concurrent parallel study [53] or, historically, a concurrent triangulation design [54], the convergent mixed methods design features the collection and analysis of both types of data and then merging of the data for the final interpretation [41].

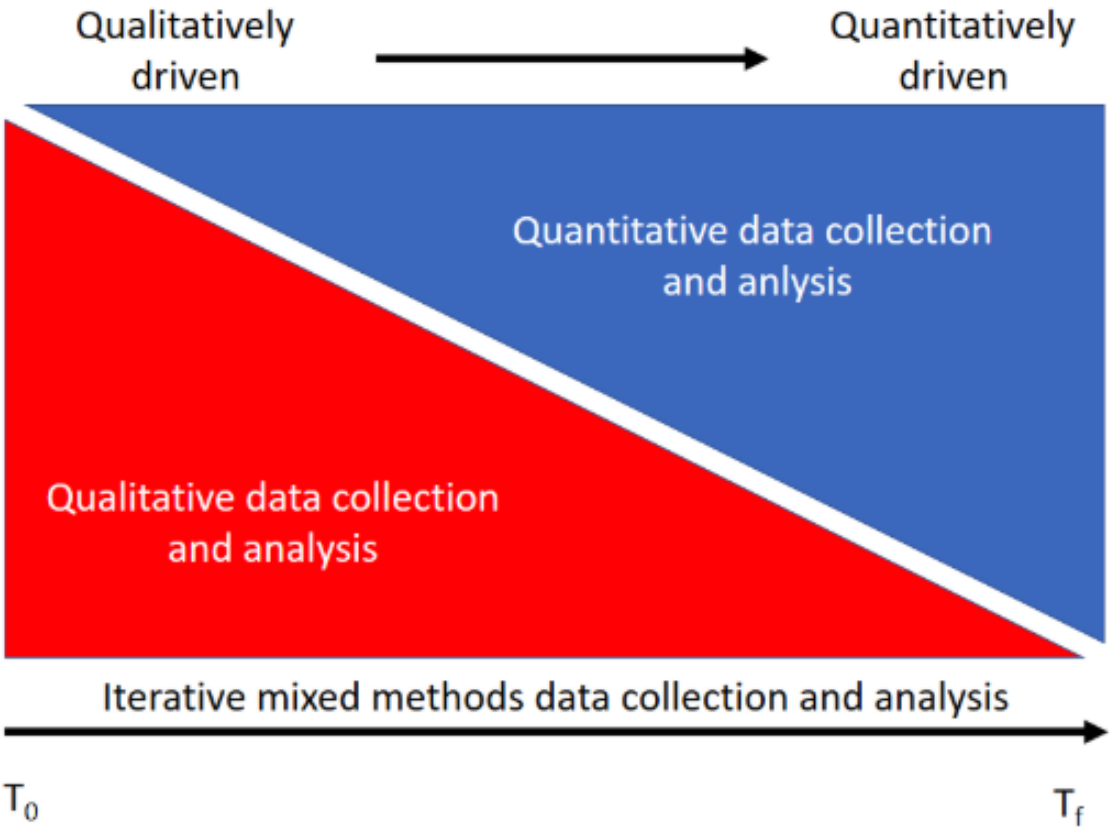
The Iterative Convergent Mixed Methods Design

Owing to the iterative nature of usability testing, we propose a new variation of the convergent design specifically for mHealth, namely, the *iterative convergent mixed methods design*. We

define an iterative convergent mixed methods design as an approach involving simultaneous qualitative and quantitative data collection and analysis that continues cyclically through multiple rounds of mixed methods data collection and analysis until the mHealth technology under evaluation is found to work to the satisfaction of the researcher. In cyclical iterations, early development is more qualitatively driven but progressively becomes more quantitatively driven; see Figure 2 [55]. Thus, the iterative convergent mixed methods design involves simultaneously collecting and analyzing qualitative and quantitative data and, as critically important, taking into consideration the iterative nature of mHealth technology development.

In the following, we articulate the features of an iterative convergent mixed methods design appropriate for mHealth intervention development and usability testing that incorporates an iterative process and is conducted according to the user's health care and usability needs. Leveraging a specific mixed methods design can help fully integrate the 2 forms of data to enhance the understanding of the usability of mHealth interventions.

Figure 2. Evolution in an iterative convergent mixed methods design from qualitatively driven to quantitatively driven.



Methodology

Fetters et al recommend considering the design, data collection procedures, interpretation, and analysis for achieving integration in a mixed methods study [46]. The iterative convergent design includes integration in various dimensions: the research aim/question, data collection, data analysis, and data interpretation. As illustrated in Figure 3, the results of each iteration inform further development of mHealth technology. These integration dimensions, as applied to usability testing, are discussed in more detail below.

Step 1: Integration in the Research Aim/Question Dimension

An iterative convergent mixed methods design should have an MMR aim as well as specific quantitative research aims and qualitative research questions**.**

Mixed Methods Aim

The mixed methods aim is to illustrate, explore, and measure how to improve the usability of an mHealth intervention. A mixed methods aim should imply both qualitative and quantitative data collection methods. For example, illustrate and explore imply qualitative data collection, whereas measure implies qualitative data collection [7].

Quantitative Research Aims

Appropriate quantitative research aims include measuring effectiveness, efficiency, and satisfaction, as illustrated in Table 3. These constructs provide a measure of patients' experienced usability. Effectiveness, efficiency, and satisfaction can be compared across iterations to identify the most usable mHealth technology.

Qualitative Research Questions

As illustrated in Table 3, appropriate qualitative research questions include clarifying and characterizing our

understanding of mHealth intervention usability. Qualitative inquiry is particularly valuable for understanding how and why a phenomenon occurs, a theory explaining a phenomenon, or the nature of someone's experience [56]. In usability testing, specific applications can include how and why participants make certain choices when using a prototype or their overall assessment of the utility. Usability testing may require or suggest a theory for its utility. The quality of the user's experience is critical for an mHealth developer who is creating a desirable user-friendly system.

A recent study by Beatty et al [50] illustrates the mixed methods process as they collected both quantitative and qualitative data to determine the usability of a mobile app for technology-facilitated home cardiac rehabilitation. Quantitative data included the SUS and task completion rate, whereas the qualitative data included questions about the functionality of the mobile app [50].

Figure 3. The iterative convergent mixed methods research design.

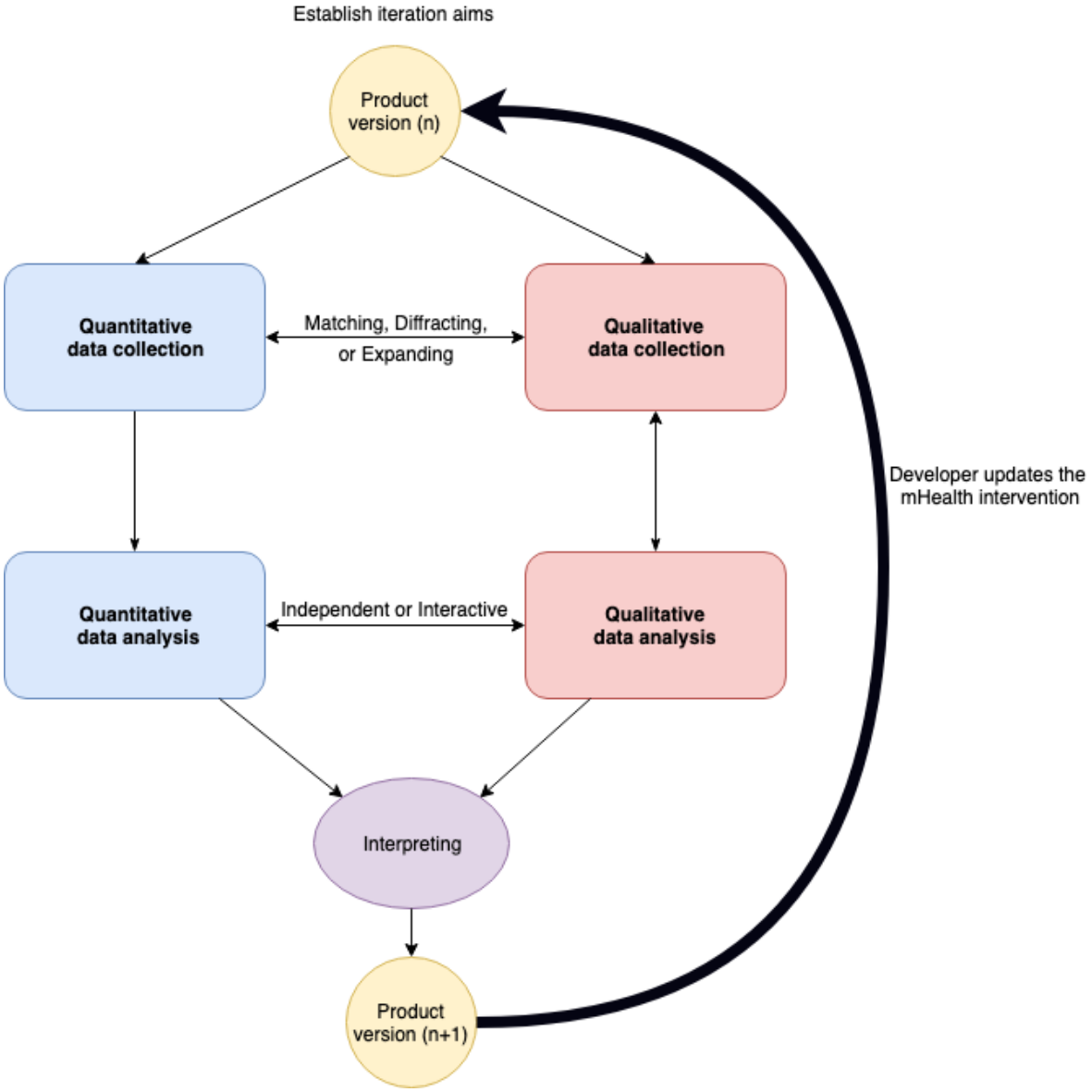


Table 3. Matching of the construct's quantitative variables and qualitative questions in a joint display depicting mixed methods of data collection.

Construct	Quantitative variables	Qualitative questions
Effectiveness	Time to learn and use	How did you learn to use the app? How can we reduce the time it takes to learn the app? What was your experience using the app? How can we reduce the time it takes to use the app?
	Data entry time	How can we reduce the time it takes to enter the data?
	Tasks time	How can we reduce the time it takes to complete the task?
	Response time	How do you feel about the app response time?
	Time to install	What are your thoughts about the time it took to install the app? The time it took to pair the medical device, if applicable?
Efficiency	Number of errors	What can we do to help users avoid the same error?
	Completion rate	What can we do to enhance the completion rate?
Satisfaction	Usability score	How often would you use the app? Why? Why not?; How do you feel about the complexity of the app?; How can we simplify it?; Do you have any recommendations to make the wording and interface easier to use? ; Would you need the support of a technical person to be able to use this system? How would you contact them: phone, email, or messaging? ; How did you find the integration of various functions in this app? How can we make it better?; How did you feel about the consistency of the app?; How can we simplify it?; Did you have any troubles when using the app? Where? How can we fix it? ; Did you feel confident when using the app? How can we make you more confident?; Did the app capture issues of importance to you?; Are there other ways to gather similar information?

Step 2: Integration in the Data Collection Dimension

During usability testing, users will be asked to provide feedback optionally on paper and, later, on working prototypes. Testing usability with 5 participants will generally be sufficient for identifying significant issues for each version [57]. During each session, participants will be given specific tasks. Both quantitative and qualitative data will be collected during and after the completion of the tasks. Researchers have 3 key strategies for integration during data collection: matching, diffracting, and expanding.

Matching

The *matching* integration strategy involves intentionally asking qualitative questions that address the scales or constructs of quantitative instruments such that both instruments will elucidate data about the *same* concepts or domains [7]. For example, the constructs of both the ISO standards and SUS used during quantitative data collection can be matched with similar or related qualitative questions to generate related quantitative and qualitative data, as illustrated in Table 3. A mixed methods data collection joint display includes the major constructs of inquiry in the first column. The latter 2 columns include the quantitative data, for example, scales or items, and the qualitative data, for example, qualitative questions or observation types. For example, Beatty et al [50,58] used matching by integrating the task completion time (quantitative) with asking "I noticed that the _____ feature took you longer than some of the others. Tell me more about that?" (qualitative). They also expanded on the SUS by asking open-ended questions regarding the user's experience with the mHealth intervention.

The qualitative questions in this table include both general and specific questions. Depending on the development needs, more general questions may be used initially, whereas later, more specific questions may be asked. When data become available, the same table structure can be populated with the findings; see *Step 4: Integration in the Data Interpretation Dimension*. The ordering of the columns is flexible according to specific project procedures.

Diffracting

The *diffracting* integration strategy involves intentionally asking qualitative questions that will address *different* aspects of the quantitative data, in the case of mHealth, the usability measure [58]. The intent is to obtain different *cuts of data* that will reveal information about different aspects of the usability that will not be addressed with the quantitative scales or items that are being collected [58]. Hence, for the ISO measures of effectiveness, efficiency, and satisfaction, qualitative questions might explore other facets, for example, animations, color patterns, sounds, and font size.

Diffracting can be used to address external factors to the user; such as the ease of connecting to the internet or connecting medical devices via Bluetooth. It is also important to develop an mHealth intervention that is energy efficient. mHealth interventions that require frequent charging of the smartphone or medical device are not recommended. Finally, developers should ensure that adequate resources are available to address medical and technical difficulties related to the mHealth intervention.

Expanding

The *expanding* integration strategy occurs when the findings from the 2 sources of data diverge and expand upon the phenomenon of interest by addressing both different aspects of a single phenomenon as well as a central phenomenon of interest [46]. Expansion involves intentionally asking qualitative questions that will be the same as the quantitative scales, while also measuring and asking qualitative questions that will address different aspects of usability. In essence, it reflects a hybrid strategy of using both matching, an area of overlap, and diffracting by looking at different aspects or facets of mHealth

during the collection of data. Each of these integration strategies could be used effectively in usability testing.

Quantitative Data Collection

Current prototyping platforms, such as InVision and Adobe XD, integrate with Lookback to enable recording of the user's interaction with a smartphone. These allow recording of the participant's voice, nonverbal reactions, and mobile phone screen display. The researcher asks participants to complete a set of tasks and assess effectiveness, efficiency, and satisfaction. The researcher records the time to learn and use mHealth technology, data entry time, task completion time, response time, and time to install the mHealth technology. The researcher also records the number of errors and task completion rate. After completing the tasks, the researcher administers the SUS to assess the user's satisfaction with the mHealth technology.

Qualitative Data Collection

The methods appropriate for assessment generally involve semistructured interviews during or after the participant's use of the prototype. Researchers can utilize cognitive testing, also called cognitive interviewing [59-61]. The researcher asks participants to use the system while continuously thinking out loud as they move through the user interface [62]. Thinking aloud questions include "Tell me what you are thinking," "What are you looking at?", or "What's on your mind?". The goal is for the users to make their thoughts *transparent* to the researcher. Verbal probing is another alternative for eliciting additional information about mHealth technology. It is a more active form of data collection in which the cognitive interviewer administers a series of probe questions specifically designed to elicit detailed information beyond that which is typically provided by respondents [59].

Another alternative to these approaches involves a postuse debrief where the interviewer observes the user going through the mHealth intervention, notes decisions made, and, after use, enquires about decisions made along each step of the way. The strength of this approach is that the user can go through the version naturally without disruption as a real user would. However, the downside is the risk that the user may forget what specific thoughts or motivations influenced their decisions during real-time use. Postuse debrief questions may include (1) whether the tool captured issues of importance to the user, (2) whether the tool was easy to use and understand regarding question wording and interface, and (3) whether there were other ways the system could be improved.

A different option involves the collection of observations to record information about behavior. This can be done in real time through the collection of notes while observing or through recordings of the user's interactions and using video elicitation interviews [63]. Video elicitation interviews question the user about their experiences and choices at certain points while interacting, for example, specific choices made and reasons for leaving a screen or returning to it. Tobii Pro can allow the researcher to track eye movements that convey behavioral patterns of use while interfacing with the mHealth technology.

Semistructured interviews and cognitive interviewing are suitable in the early stages of development. The goal is to identify *bugs* in the system, that is, anything that is dysfunctional or suboptimal. In later stages, verbal probing and video elicitation interviews are recommended to obtain specific data about the engineered changes.

Step 3: Integration in the Data Analysis Dimension

There are 2 approaches for an integrated analysis: an interactive analysis strategy or an independent intramethod analysis [64].

Interactive Analysis Strategy

The *interactive analysis strategy* [64], also called a crossover-tracks analysis [65], means that the researcher considers the qualitative and quantitative findings, in real time, as the data are collected and analyzed. That is, the data are openly, actively, and interactively considered in the context of each other. Metaphorically, the data are *talking to each other*.

Independent Intramethod Strategy

The *independent intramethod strategy*, also called a parallel-tracks analysis [65], means that the researcher uses an intramethod (ie, within method) qualitative data analysis strategy separately or independently to the quantitative data analysis strategy. First, each type of data is examined using a strategy appropriate for the type of data, for example, statistical analysis of quantitative data and thematic analysis of the qualitative data. After the separate/independent analysis, the findings are then integrated to draw an overarching interpretation, so-called *metainferences* in MMR methodology [65].

For an iterative convergent design, the research can and likely will use both strategies depending on the stage of testing. The interactive analysis strategy is preferred, especially during early prototype testing when the number of users will invariably be smaller and there is an urgency for identifying major issues. As statistical analyses will not be feasible or necessary, this approach allows rapidly assessing user rankings of certain features, for example, using the SUS as well as their qualitative experiences with the system.

In later cycles of testing, the analysis may shift to a more independent intramethod analysis strategy. As a higher number of users engage and real-time automated digital user data emerge, the interactive analysis approach may become more challenging to conduct. Moreover, the independent intramethod analysis may be preferred when the scale of testing expands such that blinded quantitative data collection becomes more important. Doing so can enable the researcher to avoid validity threats to the data quality that could occur by changing the data collection approach or by sharing patterns with users in real time. For example, Kron et al [66] linked the qualitative findings from learners' reflections on their experiences after completing the qualitative and quantitative analyses.

Combined Independent and Interactive Data Analysis

The third option can involve an iteration of both interactive and independent data analyses, that is, user survey and interview data conducted in real time can be looked at interactively, whereas automated data collection that accumulates as the number of users expands may be brought into the results of the interactive analysis after being examined independently. The

exact approach may vary and evolve according to development needs.

The Fit of the 2 Types of Data When Considered Together

Comparing both the qualitative and quantitative findings allows researchers to examine the similarities, differences, or contradictions. This comparison also allows researchers to obtain an expanded understanding of when the qualitative and quantitative findings from the analyses are merged for an interpretation. Similarities occur when there is *convergence* or confirmation between the qualitative and quantitative findings. Differences occur when the 2 types of data illustrate different, nonconflicting interpretations, so-called *complementarity* [67]. There is an expanded understanding, namely, *expansion*, when qualitative and quantitative finding provides a broader understanding of a central commonality [7,46]. Contradictions occur when there is *discordance* or divergence between the findings of the qualitative and quantitative data. To handle discordance, Fetters et al recommend gathering additional data, reanalyzing existing databases to resolve differences, seeking explanations from theory, or challenging the validity of the constructs [7,46].

Step 4: Integration in the Data Interpretation Dimension

A key challenge in mixed methods studies is how to merge the qualitative and quantitative data. A very promising approach of growing popularity among mixed methods researchers is the creation of a joint display [68]. Table 4 provides an example of presenting matched quantitative and qualitative data through a joint display. This joint display is derived from a randomized multisite mixed methods trial designed to compare a medical student's attitudes and experiences regarding the intervention, a virtual human-computer simulation program teaching communication skills, or a control, a computer-based learning module focused on teaching communication skills [66]. The data collection for this project included usability-focused questioning [8].

Joint displays allow researchers to integrate data through visual means to draw out new insights beyond the information gained from the separate quantitative and qualitative results [46,68]. Joint displays are commonly built by organizing quantitative and qualitative findings of a related construct or topic in a table. When matching has been used during data collection, this process follows naturally. In the case of mHealth, the joint display might include the usability constructs, user's perceptions, and an image or even a video representing the task [69]. For example, SUS and ISO metrics can be used to populate the numerical score and SD in the quantitative column. In addition, themes and representative quotations can be used to populate the qualitative column. In the final column, metainferences, an interpretation in consideration of the qualitative and quantitative findings, are made about the findings when analyzed together.

Step 5: Developer Updates the Mobile Health Intervention

After merging the data and drawing interpretations about their cumulative meaning (making metainferences), an iterative convergent mixed methods design then involves the results being communicated to the developers who will include the recommendations in the new iteration of the intervention. Moreover, as the developers make changes, they may also have specific questions to be answered in the subsequent cycle of iterative convergent data collection. Thus, newly emerging questions are added into subsequent rounds of data collection. In general, both qualitative and quantitative data (task completion rate, task completion time, number of errors, completing rate, and the SUS questionnaire) should be compared with each iteration for new mHealth versions.

Step 6: Iterative Evaluation

As illustrated in Figure 3, after 1 cycle of iterative evaluation, the next step is to develop a new version that has incorporated the findings from the previous evaluation. For example, Beatty et al [50] compared the task completion success rate and SUS across iterations of the mHealth intervention. The procedure will be similar by going back to *Step 2: Integration in the Data Collection Dimension*. With the new iteration, there will be new questions to ask, sometimes more general and sometimes more specific, depending on what changes were made. Furthermore, a new iteration is also required when the researcher introduces a new feature or functionality to the mHealth intervention.

On the basis of the results of the usability test, many changes may be required. The researcher should prioritize these changes while focusing on the user's needs. Generally, the magnitude of data collection and intensity will change. In the early rounds of development, the qualitative component of the mixed methods evaluation will weigh more heavily for identifying the macrolevel changes (Figure 1). This is more of a qualitatively driven mixed methods approach [64]. In subsequent iterations, as the prototype moves from paper to digital prototype to product, changes may depend much more heavily on the quantitative automated analyses that can accumulate with increased numbers of users, a quantitatively driven approach. Hence, in later cycles, the quantitative data may help identify problems, whereas the qualitative data can be used to identify solutions.

Reaching Closure in an Iterative Convergent Mixed Methods Design

Many researchers use the concept of saturation when conducting usability testing [70,71]. Saturation represents the point at which the researcher stops collecting data based on the criterion of not finding new information relevant to the development of the mHealth application. Researchers also have the option for longitudinal evaluation by comparing user satisfaction with the new mHealth technology with the result of the SUS across iterations. Similarly, qualitative data about specific features can be compared as well. After usability testing, the final prototype of the mHealth intervention can then be included in a pilot study for final refinement before launching it in a larger trial. During these subsequent phases, the iterative convergent mixed methods design will naturally continue, even into the trial.

Table 4. A joint display adapted from Kron et al's MPathic-VR mixed methods trial comparing a virtual human simulation and a computer-based communications module that illustrates medical students' attitudes and experiences in both trial arms.

Domains	MPathic-VR		Computer Based Learning		Interpretation of mixed methods findings
	Attitudinal item, mean (SD)	Qualitative reflection; illustrative quotes	Attitudinal item, mean (SD)	Qualitative reflection; illustrative quotes	
Verbal communication	5.02 (1.62)	"How to introduce myself without making assumptions about the cultural background of the patient and the family"	3.89 (1.67)	"This educational module was useful for clarifying the use of SBAR and addressing ways that all members of a health care team can improve patient care through better communication skills"	Intervention arm comments suggest deeper understanding of the content than teaching using memorization and mnemonics as in the control, a difference confirmed by higher attitudinal scores
Nonverbal communication	4.11 (1.85)	"Effective communication involves non-verbal facial expression like smiling and head nodding"	2.77 (1.45)	None	Intervention arm comments address the value of learning nonverbal communication, the difference confirmed by attitudinal scores
Training was engaging	5.43 (1.55)	"Reviewing the video review was a great way to see my facial expressions and it allowed me to improve on these skills the second time around"	3.69 (1.62)	"This experience can be improved by incorporating more active participation. For example, there could have been a scenario in which we would have to select the appropriate hand-off information per SBAR guideline"	Intervention arm comments reflect engagement through the after-action review, whereas the control comments suggested the need for interaction, the difference confirmed by higher attitudinal scores
Effectiveness in learning to handle emotionally charged situations	5.13 (1.48)	"I tend to try to smile more often than not in emotionally charged situations and that may result in conveying the wrong message"	2.34 (1.35)	"I anticipate that high-stress situations where time is exceedingly crucial requires modification to the methods presented."	Intervention arm comments indicate awareness of communication in emotionally charged situations, yet control comments indicate the need for additional training, a difference confirmed in attitudinal scores

Discussion

Here, we emphasize the need and process for mHealth researchers to use state-of-the-art mixed methods procedures. Previous single method usability studies were limited in their findings. Some studies have assessed usability using only qualitative data [72-74]. These studies can only elucidate an understanding of how and why participants make certain choices when using a prototype or their overall assessment of the utility. On the contrary, some studies have used quantitative data exclusively to assess usability [75,76]. These studies are limited to specific questions about usability and could miss valuable experiential information.

Features of the Iterative Convergent Mixed Methods Design

Despite the recognized value of using mixed methods for usability testing [49-51,77], researchers have lacked a specific design and clear procedures featuring an integrated approach that is focused on mHealth development. We believe, and the identified literature supports, that many researchers in the field are only using qualitative and quantitative procedures separately without a focus on the features of integration [37,50,51]. This illustrates explicitly why mHealth intervention researchers need a better understanding of how to incorporate the latest advances in MMR methodology, which explicitly emphasizes integration, and has procedures for achieving it.

We suggest the following criteria for evaluating the quality of studies that have used the iterative convergent mixed methods design:

1. The authors report on an empirical mHealth-related usability study.
2. The authors use an integrated mixed methods approach, defined as the collection, analysis, and integration of quantitative and qualitative data [52].
3. The authors compare the results of various iterations of the mHealth intervention.

The iterative convergent mixed methods design provides a clear framework for integrating quantitative and qualitative data to assess usability. As illustrated in Figure 3, there are multiple dimensions during testing, questions, data collection, analysis, and interpretation as well as subsequent data collection that characterize an iterative convergent mixed methods design.

With this design, researchers will start with a mockup, prototype, or the actual mHealth intervention that is represented in the diagram by the circle named the mHealth technology version.

In each round, the researcher will evaluate aspects of the version using both qualitative and quantitative research aims and, importantly, making overarching interpretations or metainferences based on the findings of both types of data that inform the next steps [65]. During data collection, the researcher can use matching, diffracting, or expanding as data are collected. Employing specific data collection approaches, the constructs explored quantitatively with scales can be explored with a similar line of qualitative questioning or inquiry. Once the data are brought together, they are compared so as to examine their confirmation, expansion, differing interpretations, or discordance [46]. As successive versions of the mHealth technology are produced, each version will involve both qualitative and quantitative data collection brought together for an integrated analysis. Iterative qualitative and quantitative data collection can be compared with each iteration to create the most usable version of the mHealth technology.

Limitations

There are potential limitations to the current usability approach. Although the small sample size may resolve the majority of usability issues [57], usability testing with a small number of individuals will generally reveal major flaws or *bugs* in the system. As the mHealth intervention becomes refined and moves from the protocol stage to the actual use stage, access to quantitative data rapidly increases and becomes more of a focus.

Usability testing can be conducted on the Web or in a laboratory setting. The value of Web-based testing is that users can participate from their natural context and use their own devices. It is also more cost-effective, and users can be in any location with an internet connection. In a laboratory setting, the researcher can probe users while they walk through their tasks, gather visual cues, assist stumped users, and ask new questions during the testing session.

We acknowledge that there are other methods, including other mixed methods designs [46], potentially applicable for usability and design research, for example, mixed methods interventions or trials. There are also other scales that can be used to quantify the satisfaction construct, such as the Post Study System Usability Questionnaire [78]. Addressing all of these methods and scales extends beyond the scope of our current focus.

Conclusions

A usable mHealth intervention with high user satisfaction can have a significantly positive effect on mHealth adoption, resulting in improved health outcomes and quality of life and reduced overall health care costs. Effective mHealth interventions are critically important for empowering patients to manage their health and also potentially enable them to participate more actively in shared decision making with their health care providers. This study offers a novel framework to guide mHealth research that has the potential to generate unique insights into multifaceted phenomena related to usability. Understanding these practices can help developers and researchers leverage the strengths of an integrated mixed methods design.

Conflicts of Interest

None declared.

Multimedia Appendix 1

System Usability Scale.

[DOCX File, 13KB - mhealth_v7i4e11656_app1.docx]

References

1. Joe J, Demiris G. Older adults and mobile phones for health: a review. J Biomed Inform 2013 Oct;46(5):947-954 [FREE Full text] [doi: 10.1016/j.jbi.2013.06.008] [Medline: 23810858]
2. Juen J, Cheng Q, Schatz B. A natural walking monitor for pulmonary patients using mobile phones. IEEE J Biomed Health Inform 2015 Jul;19(4):1399-1405. [doi: 10.1109/JBHI.2015.2427511] [Medline: 25935052]
3. Zhang J, Song Y, Bai C. MIOTIC study: a prospective, multicenter, randomized study to evaluate the long-term efficacy of mobile phone-based Internet of Things in the management of patients with stable COPD. Int J Chron Obstruct Pulmon Dis 2013;8:433-438 [FREE Full text] [doi: 10.2147/COPD.S50205] [Medline: 24082784]
4. Alwashmi M, Hawboldt J, Davis E, Marra C, Gamble J, Abu Ashour W. The effect of smartphone interventions on patients with chronic obstructive pulmonary disease exacerbations: a systematic review and meta-analysis. JMIR Mhealth Uhealth 2016 Sep 1;4(3):e105 [FREE Full text] [doi: 10.2196/mhealth.5921] [Medline: 27589898]
5. Sarkar U, Gourley G, Lyles C, Tieu L, Clarity C, Newmark L, et al. Usability of commercially available mobile applications for diverse patients. J Gen Intern Med 2016 Dec;31(12):1417-1426 [FREE Full text] [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
6. Carayon P, Kianfar S, Li Y, Xie A, Alyousef B, Wooldridge A. A systematic review of mixed methods research on human factors and ergonomics in health care. Appl Ergon 2015 Nov;51:291-321 [FREE Full text] [doi: 10.1016/j.apergo.2015.06.001] [Medline: 26154228]
7. Fetters M, Molina-Azorin J. The journal of mixed methods research starts a new decade: the mixed methods research integration trilogy and its dimensions. J Mix Methods Res 2017 Jun 16;11(3):291-307. [doi: 10.1177/1558689817714066]

8. Fetters M. Six equations to help conceptualize the field of mixed methods. J Mix Methods Res 2018 Jun 20;12(3):262-267. [doi: 10.1177/1558689818779433]
9. Fetters M, Freshwater D. The 1 + 1 = 3 integration challenge. J Mix Methods Res 2015 Apr 9;9(2):115-117. [doi: 10.1177/1558689815581222]
10. World Health Organization. 2011. mHealth: New horizons for health through mobile technologies: Based on the findings of the second global survey on eHealth (global observatory for eHealth series, volume 3) URL: http://www.who.int/goe/publications/goe_mHealth_web.pdf [accessed 2019-04-05] [WebCite Cache ID 77PK3pRX0]
11. Fogel A, Kvedar J. Artificial intelligence powers digital medicine. NPJ Digit Med 2018 Mar 14;1(1):5. [doi: 10.1038/s41746-017-0012-2]
12. Watson H. IBM. Watson Health: Get the Facts URL: https://www.ibm.com/blogs/watson-health/watson-health-get-facts/ [accessed 2019-01-28] [WebCite Cache ID 75lm9stxT]
13. Källander K, Tibenderana J, Akpogheneta O, Strachan DL, Hill Z, ten Asbroek AH, et al. Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low- and middle-income countries: a review. J Med Internet Res 2013 Jan 25;15(1):e17 [FREE Full text] [doi: 10.2196/jmir.2130] [Medline: 23353680]
14. Aranda-Jan C, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014 Feb 21;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
15. Vashist S, Luppa P, Yeo L, Ozcan A, Luong J. Emerging technologies for next-generation point-of-care testing. Trends Biotechnol 2015 Nov;33(11):692-705. [doi: 10.1016/j.tibtech.2015.09.001] [Medline: 26463722]
16. Hayes D, Markus H, Leslie R, Topol E. Personalized medicine: risk prediction, targeted therapies and mobile health technology. BMC Med 2014 Feb 28;12(1):37 [FREE Full text] [doi: 10.1186/1741-7015-12-37] [Medline: 24580858]
17. International Organization for Standardization. 2010. ISO 9241-210 ergonomics of human-system interaction -- part 210: Human-centred design for interactive systems URL: https://www.iso.org/standard/52075.html [accessed 2019-04-05] [WebCite Cache ID 77PKzAfQr]
18. Nelson L, Mayberry L, Wallston K, Kripalani S, Bergner E, Osborn C. Development and usability of REACH: a tailored theory-based text messaging intervention for disadvantaged adults with type 2 diabetes. JMIR Hum Factors 2016 Sep 8;3(2):e23 [FREE Full text] [doi: 10.2196/humanfactors.6029] [Medline: 27609738]
19. Steele Gray C, Khan A, Kuluski K, McKillop I, Sharpe S, Bierman AS, et al. Improving patient experience and primary care quality for patients with complex chronic disease using the electronic patient-reported outcomes tool: adopting qualitative methods into a user-centered design approach. JMIR Res Protoc 2016 Feb 18;5(1):e28 [FREE Full text] [doi: 10.2196/resprot.5204] [Medline: 26892952]
20. International Organization for Standardization. Ergonomics of human-system interaction - part 11: Usability: Definitions and concepts URL: https://www.iso.org/standard/63500.html [accessed 2019-04-05] [WebCite Cache ID 77PK8kRmL]
21. Goldberg L, Lide B, Lowry S, Massett HA, O'Connell T, Preece J, et al. Usability and accessibility in consumer health informatics current trends and future challenges. Am J Prev Med 2011 May;40(5 Suppl 2):S187-S197. [doi: 10.1016/j.amepre.2011.01.009] [Medline: 21521594]
22. Bevan N, Carter J, Harker S. ISO 9241-11 revised: what have we learnt about usability since 1998? Int Conf Hum Comput Interact 2015 Jul 21:2015-2151. [doi: 10.1007/978-3-319-20901-2_13]
23. Horton K. The use of telecare for people with chronic obstructive pulmonary disease: implications for management. J Nurs Manag 2008 Mar;16(2):173-180. [doi: 10.1111/j.1365-2834.2008.00845.x] [Medline: 18269548]
24. Mayberry L, Mulvaney S, Johnson K, Osborn C. The MEssaging for diabetes intervention reduced barriers to medication adherence among low-income, diverse adults with type 2. J Diabetes Sci Technol 2017 Dec;11(1):92-99 [FREE Full text] [doi: 10.1177/1932296816668374] [Medline: 27595710]
25. Dorsey E, Chan YF, McConnell MV, Shaw SY, Trister AD, Friend SH. The use of smartphones for health research. Acad Med 2017 Dec;92(2):157-160. [doi: 10.1097/ACM.0000000000001205] [Medline: 27119325]
26. research2guidance. 2018. mHealth economics - how mHealth app publishers are monetizing their apps URL: https://research2guidance.com/product/mHealth-economics-how-mHealth-app-publishers-are-monetizing-their-apps/ [accessed 2019-04-05] [WebCite Cache ID 77PKSpVXL]
27. Usability.gov. Benefits of user-centered design URL: https://www.usability.gov/what-and-why/benefits-of-ucd.html [accessed 2019-04-05] [WebCite Cache ID 77PKmvj2H]
28. Ribeiro N, Moreira L, Barros A, Almeida A, Santos-Silva F. Guidelines for a cancer prevention smartphone application: a mixed methods study. Int J Med Inform 2016 Dec;94:134-142. [doi: 10.1016/j.ijmedinf.2016.07.007] [Medline: 27573321]
29. Smith S, Claridy M, Whitehead M, Sheats JQ, Yoo W, Alema-Mensah EA, et al. Lifestyle modification experiences of African American breast cancer survivors: a needs assessment. JMIR Cancer 2015;1(2) [FREE Full text] [doi: 10.2196/cancer.4892] [Medline: 26380378]
30. Hattink B, Droes R, Sikkes S, Oostra E, Lemstra A. Evaluation of the Digital Alzheimer Center: testing usability and usefulness of an online portal for patients with dementia and their carers. JMIR Res Protoc 2016 Jul 21;5(3):e144 [FREE Full text] [doi: 10.2196/resprot.5040] [Medline: 27444209]

31. Triantafyllidis A, Velardo C, Chantler T, Shah SA, Paton C, Khorshidi R, SUPPORT-HF Investigators. A personalised mobile-based home monitoring system for heart failure: the SUPPORT-HF Study. Int J Med Inform 2015 Oct;84(10):743-753. [doi: 10.1016/j.ijmedinf.2015.05.003] [Medline: 26037921]
32. Lyles C, Sarkar U, Osborn C. Getting a technology-based diabetes intervention ready for prime time: a review of usability testing studies. Curr Diab Rep 2014 Oct;14(10):534 [FREE Full text] [doi: 10.1007/s11892-014-0534-9] [Medline: 25173689]
33. Zapata B, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
34. Hopkins I, Dunn K, Bourgeois F, Rogers J, Chiang V. From development to implementation-a smartphone and email-based discharge follow-up program for pediatric patients after hospital discharge. Healthc (Amst) 2016 Jun;4(2):109-115. [doi: 10.1016/j.hjdsi.2015.11.003] [Medline: 27343160]
35. Krug S. Don't Make Me Think, Revisited: A Common Sense Approach To Web Usability (3rd Edition) (voices That Matter). San Francisco: New Riders; 2014.
36. Brooke J. Universite de Geneve. 1996. SUS-a quick and dirty usability scale URL: https://cui.unige.ch/isi/icle-wiki/_media/ipm:test-suschapt.pdf [accessed 2019-04-06] [WebCite Cache ID 77QdKu9Ik]
37. Farinango C, Benavides J, Cerón JD, López DM, Álvarez RE. Human-centered design of a personal health record system for metabolic syndrome management based on the ISO 9241-210:2010 standard. J Multidiscip Healthc 2018;11:21-37 [FREE Full text] [doi: 10.2147/JMDH.S150976] [Medline: 29386903]
38. Gunter R, Fernandes-Taylor S, Mahnke A, Awoyinka L, Schroeder C, Wiseman J, et al. Evaluating patient usability of an image-based mobile health platform for postoperative wound monitoring). JMIR Mhealth Uhealth 2016 Sep 28;4(3):e113 [FREE Full text] [doi: 10.2196/mhealth.6023] [Medline: 27683059]
39. Georgsson M, Staggers N. Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. J Am Med Inform Assoc 2016 Jan;23(1):5-11 [FREE Full text] [doi: 10.1093/jamia/ocv099] [Medline: 26377990]
40. Moumane K, Idri A, Abran A. Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards. Springerplus 2016;5:548 [FREE Full text] [doi: 10.1186/s40064-016-2171-z] [Medline: 27190747]
41. Creswell J. A Concise Introduction To Mixed Methods Research. Thousand Oaks, California: Sage Publications, Inc; 2015.
42. Johnson R, Onwuegbuzie A, Turner L. Toward a definition of mixed methods research. J Mix Methods Res 2016 Jun 29;1(2):112-133. [doi: 10.1177/1558689806298224]
43. Curry L, Nunez-Smith M. Mixed Methods in Health Sciences Research. Thousand Oaks, California: Sage Publications, Inc; 2015.
44. Curry L, Nembhard I, Bradley E. Qualitative and mixed methods provide unique contributions to outcomes research. Circulation 2009 Mar 17;119(10):1442-1452. [doi: 10.1161/CIRCULATIONAHA.107.742775] [Medline: 19289649]
45. Creswell J, Klassen A, Plano CV, Smith K. Best Practices for Mixed Methods Research in the Health Sciences. Bethesda (Maryland): National Institutes of Health; 2013:541-545.
46. Fetters M, Curry L, Creswell J. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
47. Mertens DM. Mixed methods and wicked problems. J Mix Methods Res 2014 Dec 5;9(1):3-6 [FREE Full text] [doi: 10.1177/1558689814562944]
48. Nitsch M, Dimopoulos C, Flaschberger E, Saffran K, Kruger JF, Garlock L, et al. A guided online and mobile self-help program for individuals with eating disorders: an iterative engagement and usability study. J Med Internet Res 2016 Jan 11;18(1):e7 [FREE Full text] [doi: 10.2196/jmir.4972] [Medline: 26753539]
49. Sage A, Roberts C, Geryk L, Sleath B, Tate D, Carpenter D. A self-regulation theory-based asthma management mobile app for adolescents: a usability assessment. JMIR Hum Factors 2017 Feb 1;4(1):e5 [FREE Full text] [doi: 10.2196/humanfactors.7133] [Medline: 28148471]
50. Beatty AL, Magnusson SL, Fortney JC, Sayre GG, Whooley MA. VA FitHeart, a mobile app for cardiac rehabilitation: usability study. JMIR Hum Factors 2018 Jan 15;5(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.8017] [Medline: 29335235]
51. Alnasser A, Kyle J, Alkhalifah A, Marais D. Relationship between evidence requirements, user expectations, and actual experiences: usability evaluation of the Twazon Arabic weight loss app. JMIR Hum Factors 2018 Apr 17;5(2):e16 [FREE Full text] [doi: 10.2196/humanfactors.9765] [Medline: 29666042]
52. Creswell J, Clark V. Designing and Conducting Mixed Methods Research. Thousand Oaks, California: Sage Publications, Inc; 2018.
53. Tashakkori A, Teddlie C, Teddlie C. Mixed Methodology: Combining Qualitative and Quantitative Approaches. Thousand Oaks, California: Sage Publications, Inc; 1998.
54. Creswell J, Plano CV, Gutmann M, Hanson W. Advanced mixed methods research designs. In: Tashakkori A, Teddlie C, editors. SAGE Handbook of Mixed Methods in Social & Behavioral Research. Thousand Oaks, California: Sage Publications, Inc; 2003:240.
55. Hesse-Biber S, Johnson R. The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry. New York, NY: Oxford University Press; 2015.

56. Teherani A, Martimianakis T, Stenfors-Hayes T, Wadhwa A, Varpio L. Choosing a qualitative research approach. J Grad Med Educ 2015 Dec;7(4):669-670 [FREE Full text] [doi: 10.4300/JGME-D-15-00414.1] [Medline: 26692985]
57. Nielsen J. Nielsen Norman Group. 2018. How many test users in a usability study? URL: https://www.nngroup.com/articles/how-many-test-users/ [accessed 2018-07-24] [WebCite Cache ID 71A29VZwV]
58. Uprichard E, Dawney L. Data diffraction: challenging data integration in mixed methods research. J Mix Methods Res 2019 Jan;13(1):19-32 [FREE Full text] [doi: 10.1177/1558689816674650] [Medline: 30595679]
59. Willis G, Artino JA. What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. J Grad Med Educ 2013 Sep;5(3):353-356 [FREE Full text] [doi: 10.4300/JGME-D-13-00154.1] [Medline: 24404294]
60. Wharton C, Rieman J, Lewis C, Polson P. The cognitive walkthrough method: a practitioner's guide. In: Usability Inspection Methods. New Jersey, United States: Wiley; 1994.
61. Lapka C, Jupka K, Wray R, Jacobsen H. Applying cognitive response testing in message development and pre-testing. Health Educ Res 2008 Jun;23(3):467-476. [doi: 10.1093/her/cym089] [Medline: 18218615]
62. Nielsen J. Nielsen Norman Group. 2012. Thinking aloud: The #1 usability tool URL: https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/ [accessed 2018-07-24] [WebCite Cache ID 71A2GL4VG]
63. Henry S, Fetters M. Video elicitation interviews: a qualitative research method for investigating physician-patient interactions. Ann Fam Med 2012;10(2):118-125 [FREE Full text] [doi: 10.1370/afm.1339] [Medline: 22412003]
64. Moseholm E, Fetters M. Conceptual models to guide integration during analysis in convergent mixed methods studies. Method Innov 2017 Dec 14;10(2):205979911770311. [doi: 10.1177/2059799117703118]
65. Teddlie C, Tashakkori A. Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences. Thousand Oaks, California: Sage Publications, Inc; 2009.
66. Kron F, Fetters M, Scerbo M, White CB, Lypson ML, Padilla MA, et al. Using a computer simulation for teaching communication skills: a blinded multisite mixed methods randomized controlled trial. Patient Educ Couns 2017 Dec;100(4):748-759 [FREE Full text] [doi: 10.1016/j.pec.2016.10.024] [Medline: 27939846]
67. Greene J, Caracelli V, Graham W. Toward a conceptual framework for mixed-method evaluation designs. Educ Eval Policy Anal 2016 Nov 23;11(3):255-274. [doi: 10.3102/01623737011003255]
68. Guetterman T, Fetters M, Creswell J. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015 Nov;13(6):554-561 [FREE Full text] [doi: 10.1370/afm.1865] [Medline: 26553895]
69. Usability.gov. Reporting usability test results URL: https://www.usability.gov/how-to-and-tools/methods/reporting-usability-test-results.html [accessed 2018-07-29] [WebCite Cache ID 71HbqLiLc]
70. Press A, McCullagh L, Khan S, Schachter A, Pardo S, McGinn T. Usability testing of a complex clinical decision support tool in the emergency department: lessons learned. JMIR Hum Factors 2015 Sep 10;2(2):e14 [FREE Full text] [doi: 10.2196/humanfactors.4537] [Medline: 27025540]
71. Voncken-Brewster V, Moser A, van der Weijden T, Nagykaldi Z, de Vries H, Tange H. Usability evaluation of an online, tailored self-management intervention for chronic obstructive pulmonary disease patients incorporating behavior change techniques. JMIR Res Protoc 2013 Jan 16;2(1):e3 [FREE Full text] [doi: 10.2196/resprot.2246] [Medline: 23612363]
72. Lodhia V, Karanja S, Lees S, Bastawrous A. Acceptability, usability, and views on deployment of peek, a mobile phone mHealth intervention for eye care in Kenya: qualitative study. JMIR Mhealth Uhealth 2016 May 9;4(2):e30 [FREE Full text] [doi: 10.2196/mhealth.4746] [Medline: 27160779]
73. Wang J, Chu C, Li C, Hayes L, Siminerio L. Diabetes educators' insights regarding connecting mobile phone- and wearable tracker-collected self-monitoring information to a nationally-used electronic health record system for diabetes education: descriptive qualitative study. JMIR Mhealth Uhealth 2018 Jul 26;6(7):e10206 [FREE Full text] [doi: 10.2196/10206] [Medline: 30049667]
74. Khurana L, Durand E, Gary S, Otero AV, Gerzon MC, Beck J, et al. Subjects with osteoarthritis can easily use a handheld touch screen electronic device to report medication use: qualitative results from a usability study. Patient Prefer Adherence 2016;10:2171-2179 [FREE Full text] [doi: 10.2147/PPA.S94247] [Medline: 27822018]
75. Schobel J, Pryss R, Probst T, Schlee W, Schickler M, Reichert M. Learnability of a configurator empowering end users to create mobile data collection instruments: usability study. JMIR Mhealth Uhealth 2018 Jun 29;6(6):e148 [FREE Full text] [doi: 10.2196/mhealth.9826] [Medline: 29959107]
76. Metelmann B, Metelmann C, Schuffert L, Hahnenkamp K, Brinkrolf P. Medical correctness and user friendliness of available apps for cardiopulmonary resuscitation: systematic search combined with guideline adherence and usability evaluation. JMIR Mhealth Uhealth 2018 Nov 6;6(11):e190 [FREE Full text] [doi: 10.2196/mhealth.9651] [Medline: 30401673]
77. Andargoli AE, Scheepers H, Rajendran D, Sohal A. Health information systems evaluation frameworks: a systematic review. Int J Med Inform 2017 Dec;97:195-209. [doi: 10.1016/j.ijmedinf.2016.10.008] [Medline: 27919378]
78. Lewis JR. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. In: Proceedings of the Human Factors Society 36th Annual Meeting. 1992 Presented at: HFES 1992; October 12-16, 1992; Atlanta, Georgia, USA p. 1259-1260. [doi: 10.1177/154193129203601617]

Abbreviations

HCD: human-centered design
ISO: International Organization for Standardization
mHealth: mobile health
MMR: mixed methods research
SUS: System Usability Scale

Edited by G Eysenbach; submitted 02.08.18; peer-reviewed by C SteeleGray, M Tistad, J Zhang, C Aranda-Jan; comments to author 27.09.18; revised version received 05.11.18; accepted 24.03.19; published 26.04.19.

Please cite as:
Alwashmi MF, Hawboldt J, Davis E, Fetters MD
The Iterative Convergent Design for Mobile Health Usability Testing: Mixed Methods Approach
JMIR Mhealth Uhealth 2019;7(4):e11656
URL: http://mhealth.jmir.org/2019/4/e11656/
doi:10.2196/11656
PMID:31025951

Original Paper

Wearable-Based Mobile Health App in Gastric Cancer Patients for Postoperative Physical Activity Monitoring: Focus Group Study

Jin-Ming Wu[1,2], MD, PhD; Te-Wei Ho[1*], PhD; Yao-Ting Chang[3], MS; ChungChieh Hsu[3], MS; Chia Jui Tsai[3], MS; Feipei Lai[3], PhD; Ming-Tsan Lin[1,2*], MD, PhD

[1]Department of Surgery, National Taiwan University Hospital, National Taiwan University, Taipei, Taiwan

[2]Department of Surgery, College of Medicine, National Taiwan University, Taipei, Taiwan

[3]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan

[*]these authors contributed equally

Corresponding Author:
Ming-Tsan Lin, MD, PhD
Department of Surgery
National Taiwan University Hospital
National Taiwan University
No.7 Chung Shan South Road
Taipei,
Taiwan
Phone: 886 223123456
Email: linmt@ntu.edu.tw

Abstract

Background: Surgical cancer patients often have deteriorated physical activity (PA), which in turn, contributes to poor outcomes and early recurrence of cancer. Mobile health (mHealth) platforms are progressively used for monitoring clinical conditions in medical subjects. Despite prevalent enthusiasm for the use of mHealth, limited studies have applied these platforms to surgical patients who are in much need of care because of acutely significant loss of physical function during the postoperative period.

Objective: The aim of our study was to determine the feasibility and clinical value of using 1 wearable device connected with the mHealth platform to record PA among patients with gastric cancer (GC) who had undergone gastrectomy.

Methods: We enrolled surgical GC patients during their inpatient stay and trained them to use the app and wearable device, enabling them to automatically monitor their walking steps. The patients continued to transmit data until postoperative day 28. The primary aim of this study was to validate the feasibility of this system, which was defined as the proportion of participants using each element of the system (wearing the device and uploading step counts) for at least 70% of the 28-day study. “Definitely feasible,” “possibly feasible,” and “not feasible” were defined as ≥70%, 50%-69%, and <50% of participants meeting the criteria, respectively. Moreover, the secondary aim was to evaluate the clinical value of measuring walking steps by examining whether they were associated with early discharge (length of hospital stay <9 days).

Results: We enrolled 43 GC inpatients for the analysis. The weekly submission rate at the first, second, third, and fourth week was 100%, 93%, 91%, and 86%, respectively. The overall daily submission rate was 95.5% (1150 days, with 43 subjects submitting data for 28 days). These data showed that this system met the definition of “definitely feasible.” Of the 54 missed transmission days, 6 occurred in week 2, 12 occurred in week 3, and 36 occurred in week 4. The primary reason for not sending data was that patients or caregivers forgot to charge the wearable devices (>90%). Furthermore, we used a multivariable-adjusted model to predict early discharge, which demonstrated that every 1000-step increment of walking on postoperative day 5 was associated with early discharge (odds ratio 2.72, 95% CI 1.17-6.32; *P*=.02).

Conclusions: Incorporating the use of mobile phone apps with wearable devices to record PA in patients of postoperative GC was feasible in patients undergoing gastrectomy in this study. With the support of the *mHealth* platform, this app offers seamless tracing of patients’ recovery with a little extra burden and turns subjective PA into an objective, measurable parameter.

(JMIR Mhealth Uhealth 2019;7(4):e11989) doi:10.2196/11989

KEYWORDS

telemedicine; exercise; perioperative care; gastrectomy; stomach neoplasms

Introduction

Background

Physical activity (PA) is a well-established parameter to not only assess the general condition but also monitor the recovery of patients [1,2]. Regular PA can be protective against the risk of gastric and breast cancers [3,4]. Although the double-labeled water method is considered to be the gold standard for assessing total energy expenditure (to reflect PA), it is rarely used because it is expensive, equipment dependent, and time consuming [5,6]. Therefore, other modalities, including self-report questionnaires, self-report activity diaries, direct observation, and devices (accelerometers, pedometers, or armbands), have been implemented to measure PA [6].

Gastric cancer (GC) is the fourth leading cancer worldwide and the second leading cause of cancer-related mortality [7]. In Taiwan, nearly 3800 new cases of GC are reported each year in patients with a median age of 70 years. Complete surgical resection and endoscopic submucosal resection are the only curative therapies that provide better long-term survival. Nevertheless, gastrectomy-related stress and discomfort adversely affect PA and quality of life immediately postoperatively and last for up to 6 months [8], contributing to poor outcomes or early recurrence, particularly in cases with advanced-stage GC with obstruction of the gastrointestinal tract and malnutrition [9,10]. Furthermore, PA declines more markedly in patients who have undergone gastrectomy and are receiving postoperative chemotherapy or chemoradiation [11]. Moreover, patients with GC encounter significant functional impairments and decreased quality of life because of decreased PA and increased gastrointestinal symptoms [12]. Despite the deterioration in physical function, regular PA (to strengthen muscle power, which leads to improved physical function), proper nutritional intervention (to improve food intake, which results in weight gain), and mental support (to preserve self-esteem and maintain social activity) may help in restoring the patients' health status and improving quality of life [13].

Traditional perioperative care depends on medical professionals asking patients about the progression of their PA; however, these self-report measures are not only unreliable in aged adults with cognitive impairment [14] but also time consuming in processing the data [15]. For the purpose of clinical research, self-report questionnaires are the most common method for PA assessment, and they have the advantage of cost-effectiveness and ease of administration [6]. However, compared with using devices for recording PA, potential disadvantages of self-report questionnaires are that they may be less reliable in measuring light or moderate PA and may also be affected by external factors, such as social desirability, age, complexity of the questionnaire, and the participants' recall ability [16,17]. Furthermore, 4 key categories of aging barriers are associated with the use of mobile health (mHealth) in aged adults, including barriers in cognition, motivation, physical ability, and perception [18]. As GC typically comprises an aged population (median age >65 years) with the potential for developing cognitive or memory impairments, it is important to select an easy-to-follow PA device to use in clinical research for this population.

mHealth is a scalable and flexible platform that can assist the practice of medicine and public health with the support of mobile devices [19-21]. Several studies have demonstrated that mHealth technology has improved clinical outcomes in medical patients by improving the control of cardiac function and glycemic hemostasis, enhancing medication compliance, and shortening hospital stay [22-25]. Although the experience of using the mHealth app in surgical care is limited [26,27], it is suggested that surgical patients can benefit from this new technology support and restore the critical decline in physical and medical functions. The mHealth system and its associated mobile apps support many theory-based techniques that have shown to increase PA in behavioral interventions [28-31], with self-monitoring being the most important element associated with success of the intervention [29]. With the support of technology-based trackers, patients are encouraged to self-monitor, and wearing automatic recorders of PA reduces the burden. Furthermore, 1 study investigated the accuracy in step counting among commercially available wearable devices, showing that most devices did not overcount or undercount steps [32]. These findings are particularly important for clinical interventions using such wearable devices for clinical research.

Objectives

Five years ago, our team worked with bioinformatics developers at our university to create a new first-generation mobile phone/tablet app (SurgeryDiary) to accelerate recovery in patients who have undergone gastrointestinal surgery [33]. This next-generation app implemented a wearable device to track daily PA of patients with GC. This pilot study focused on the feasibility and clinical value of the second-generation app in patients with GC. In addition, this study determined the correlation between PA variables collected from the wearable device and outcomes and illustrated how the device can be used to estimate patient recovery.

Methods

Study Population

In this study, eligible participants were adult inpatients (aged ≥20 years) on the general surgery service of an academic teaching hospital. We enrolled patients who were undergoing laparoscopic or open gastrectomy for GC at our institution from January 2016 to December 2017. All patients fulfilling the inclusion criteria were approached to participate in this study. Notably, patients were excluded if they had preoperative walking disorders (paralysis or hemiplegia) or prolonged stay in the intensive care unit (ICU). During the study period, 50 patients were screened; of these patients, we enrolled 43 in this study, excluding 7 with a prolonged ICU stay. This study protocol was approved by the institutional review board of the National Taiwan University Hospital (201412040RIND). A research assistant helped the enrolled patients to install the app on their smartphones and instructed them on how to use the app preoperatively. Before the enrollment of this study, the research assistant would evaluate the patients' familiarity with wearable devices and smartphones. If the patients were not confident about using these devices, we would provide further instruction to their caregivers.

Data of patients' demographics and oncological factors were obtained by 2 medical professionals after reviewing charts (discharge summaries, imaging reports, and pathological reports). Regarding comorbidities, we collected the following data on the comorbidity of patients before gastrectomy using the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes: anemia (ICD-9-CM: 285.x), myocardial infarction (410.x and 412.x), mild liver disease (571.2 and 571.4-571.6), hyperlipidemia (272.0-272.2), diabetes mellitus (250.0-250.3 and 250.7), chronic obstructive pulmonary disease (490.x-496.x), renal failure (584.x-586.x), and hypertension (401.x-405.x). Next, we used the Charlson comorbidity index to calculate baseline comorbidity scores for each patient [34]. This was used to compare the baseline comorbidity between the 2 groups. Moreover, we used the Eastern Cooperative Oncology Group performance status to represent the baseline daily living ability [35]. Higher scores in both scales implied that the patients had poor medical and physical functions. Furthermore, the American Society of Anesthesiologists (ASA) score, which ranges from 1 to 5, assesses the preoperative physical status of patients [36]. The definition of ASA 1, 2, 3, 4, and 5 is "a normal healthy patient," "a patient with mild systemic disease," "a patient with severe systemic disease," "a patient with severe systemic disease that is a constant threat to life," and "a moribund patient who is not expected to survive without the operation," respectively. For the oncological variables, including cancer histology and cancer stage, the American Joint Committee on Cancer staging was implemented to determine cancer stages of this study's population [37]. Satisfaction survey of the wearable device and app was measured using items adapted from a previous study [38], including 5 questions (patient comfort using the wearable device, if the patient would continue to wear the device, if it was convenient for the patient to use the app, if the app was user friendly, and if the patient would recommend this app). All responses were made on a Likert-type scale from 1 (strongly disagree) to 5 (strongly agree) as the study period was completed.

The App

SurgeryDiary is an iOS/Android app that facilitates patients who have undergone gastrectomy to access educational information of surgical procedures, record perioperative clinical variables (PA, associated discomfort, body weight, and drain amount), and transmit digital images of the surgical wound to the medical staff. PA (Figure 1, left) and heart rate (Figure 1, right) could be graphed for diverse study periods (ranging from days to years), and patients/caregivers could view the change in PA or heart rate by time increment. This app was developed by surgical professionals and software programmers to fulfill the need of the patients. The wearable device for recording PA (number of steps) is Apple Watch for iOS and Samsung Gear S2 for Android. Reportedly, both devices could reliably measure the number of steps as effective health evaluation indicators [39]. Figure 2 presents the system architecture. The wearable devices initially collected daily step count data, which were then transferred to the original customized wearable apps in the smartphones using the sync functionality. Furthermore, our designed app captured the data stored in the original customized wearable app. Next, data in our designed app were sent to the server when internet access was available for these smartphones. This function of interdevice data transmission worked well with no abnormal events reported by the patients. Furthermore, 1 case manager monitored the synchronous data during daytime on weekdays and would call the patients if their step counts had decreased or were completely missed. Patients continued to transmit data until postoperative day 28, and 1 medical staff who did not participate in designing this system independently reviewed and analyzed data.

Figure 1. Screenshots of the app showing the number of walking steps (left) and heart rates (right) at different periods.

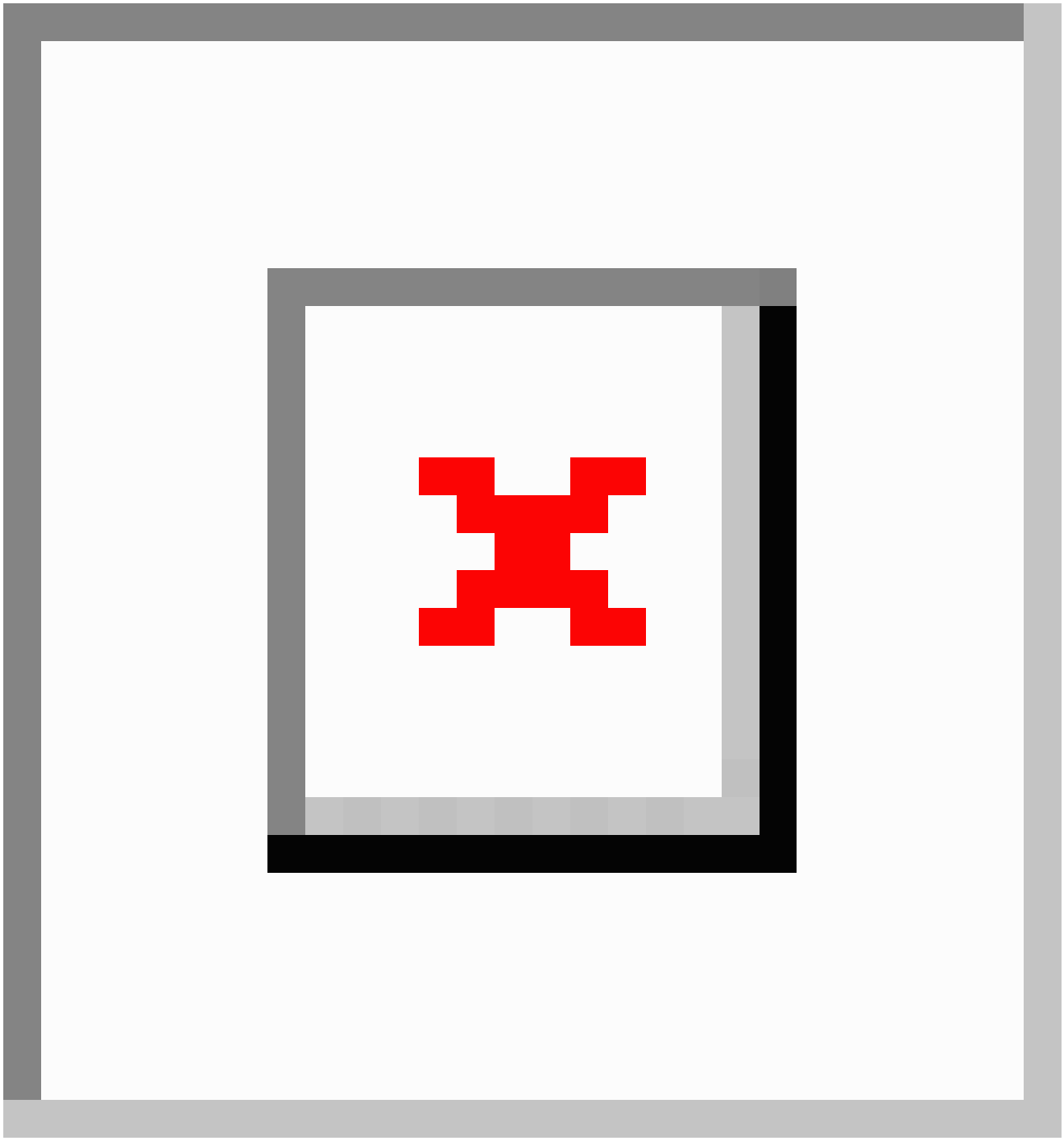
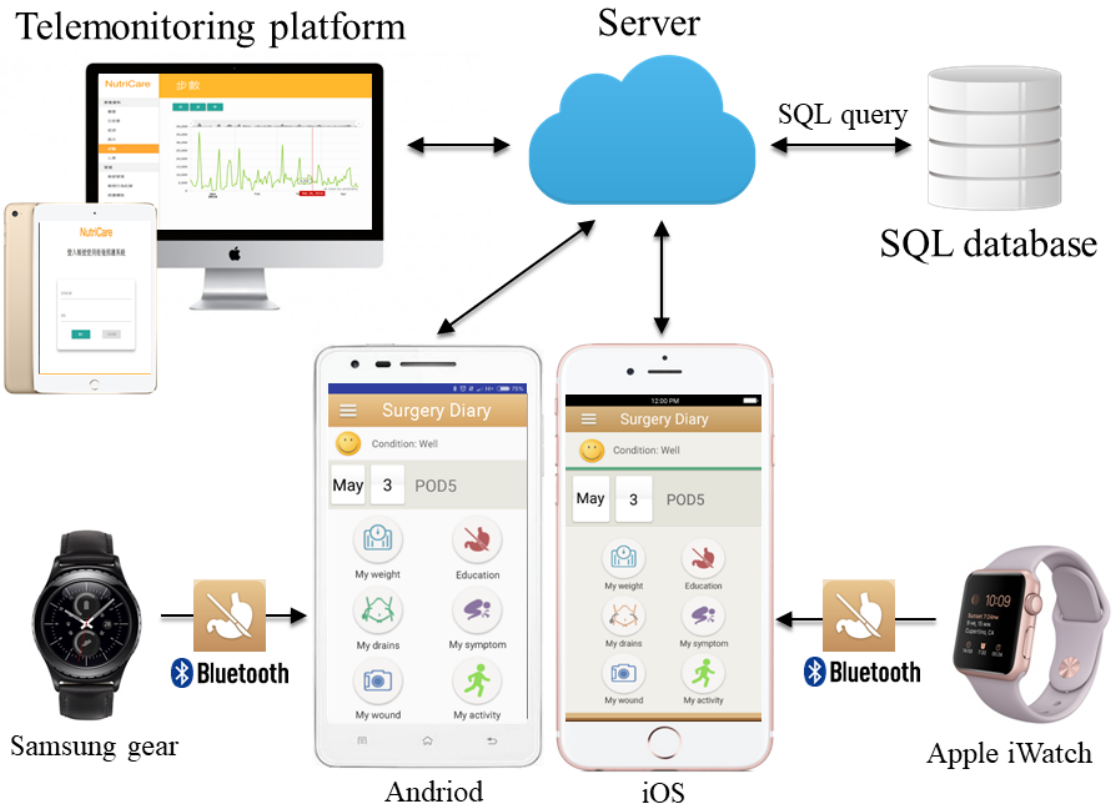


Figure 2. The system architecture: the wearable device for recording physical activity (number of steps) is Apple Watch for iOS and Samsung Gear S2 for Android. The wearable devices connected and sent data to the original customized apps in mobile phones, which were further connected to our designed app. SQL: standard database language.



Outcomes

Feasibility: Protocol Completion

In this study, feasibility was defined as the proportion of participants using each element of the system (wearable device and uploading step counts) for at least 70% of the 28-day study period [40]. “Definitely feasible,” “possibly feasible,” and “not feasible” were defined as ≥70%, 50%-69%, and <50% participants meeting criteria, respectively.

Clinical Value: Association of Walking Steps and Early Discharge

The outcome of clinical value was to validate whether the improvement of PA based on wearable devices was a reliable parameter to estimate early discharge (length of hospital stay <9 days) in patients who had undergone gastrectomy. Our hypothesis was based on 1 report addressing that physical functional recovery, including adequate pain control, ability to mobilize, tolerance of oral intake, and no abnormal physical findings or laboratory tests, could be achieved on postoperative day 5 in patients with GC who had undergone gastrectomy; this was a parameter to predict if patients would qualify for early discharge [41]. Furthermore, we collected demographic (age, gender, body mass index, and comorbidities), type of gastrectomy, minimal invasive surgery, cancer histology, and cancer stage information for statistical adjustment.

Results

Participants’ Characteristics

Of the 43 analyzed patients (Table 1), 51% (22/43) were males, and the median age was 68 years, which is consistent with that observed in the population with GC in Taiwan. The median age of male and female patients was 72 years and 59 years, respectively. Among the study population, the median body weight was 65 kg and the median body mass index was 22.3. Regarding surgical procedures, there were 7 (16%) wedge resections, 30 (70%) subtotal gastrectomies, and 6 (14%) total gastrectomies. We observed a low value of serum albumin (<4.0 mg/dl; parameter of malnutrition) in 15 (35%) patients, and 22 (51%) patients had an ASA physical status classification system score of >2. Notably, 30 patients (70%) depended on the assistance of active caregivers to perform this task because of old age or limited prior experience with technology.

All questions received a mean rating over 4 out of 5, but 2 (5%) patients reported a rating of 3 for 1 question (“patient comfort using the wearable device”).

Feasibility: Protocol Completion

All participants submitted step data every day in the first week. The weekly submission rate in the second, third, and fourth week was 93%, 91%, and 86%, respectively. The overall daily submission rate was 95.5% (1150 of days, given 43 subjects submitting data for 28 days). Of 54 missed transmission days, 6 occurred in the second week, 12 occurred in the third week, and 36 occurred in the fourth week. The leading reason for not

sending data was that patients/caregivers forgot to charge the wearable devices (>90%).

Clinical Value: Association of Walking Steps and Early Discharge

Table 2 outlines the comparison of study subjects with early discharge and without early discharge. There were 19 (44%) patients with early discharge. The patients with early discharge exhibited a significantly higher proportion of patients with minimal invasive surgery (68% vs 29%; *P*=.01), cancer histology of gastrointestinal stromal cancer (32% vs 4%; *P*=.01), and more walking steps on postoperative day 5 (5823 vs 4311; *P*=.01) compared with the patients without early discharge. Furthermore, we validated whether walking steps could predict early discharge using the logistic regression model (Table 3). We included significant variables on univariable analysis with a *P*<.05 into the multivariable analysis, which demonstrated that every 1000-step increment of walking on postoperative day 5 was associated with early discharge (odds ratio 2.72, 95% CI 1.17-6.32; *P*=.019) after adjustment of both baseline physical status and aforementioned variables.

Table 1. Demographic characteristics of study subjects, participant adherence to the protocol, and clinical outcomes.

Characteristic	Data (N=43)
Age, median (IQR[a])	68.0 (58.0-75.0)
Male, n (%)	22 (51)
Body mass index, median (IQR)	22.3 (19.9-25.8)
Value of serum albumin (mg/dl) <4, n (%)	15 (35)
Smoking history, n (%)	39 (91)
American Society of Anesthesiologists score >2, n (%)	22 (51)
Minimal invasive surgery, n (%)	20 (47)
Charlson Comorbidity Index >2, n (%)	10 (23)
Education level, n (%)	
Illiterate	1 (2)
Up to high school	30 (70)
>High school	12 (28)
Employment level, n (%)	
Full-time employment	25 (59)
Part-time employment	7 (16)
Retired	10 (23)
Unemployed	1 (2)
Financial situation, n (%)	
Self-pay	41 (95)
Social support	2 (5)
Place of residence, n (%)	
Urban area	30 (70)
Rural area	13 (30)
Eastern Cooperative Oncology Group performance status scales, n (%)	
0	38 (88)
1	5 (12)
Method of gastrectomy, n (%)	
Wedge resection	7 (16)
Subtotal gastrectomy	30 (70)
Total gastrectomy	6 (14)
Histological classification, n (%)	
Adenocarcinoma	36 (84)
Malignant gastrointestinal stromal tumor	7 (16)
System of American Joint Committee on Cancer staging, n (%)	
Stage I	15 (35)
Stage II	20 (47)
Stage III	8 (18)
Major complication	6 (14)
Method of participation, n (%)	
Caregiver	13 (30)
Independent	30 (70)
App system, n (%)	

Characteristic	Data (N=43)
Android	38 (88)
iOS	5 (12)
Participant compliance, n (%)	
At first week	43 (100)
At second week	40 (93)
At third week	39 (91)
At fourth week	37 (86)
Overall daily compliance	
Total submissions (person-day)	1204
Days submitted, n (%)	1150 (95)
Days missed, n (%)	54 (5)

[a]IQR: interquartile range.

Table 2. Comparison of the study subjects with early discharge and without early discharge.

Characteristic	Without early discharge (n=24)	With early discharge (n=19)	*P* value
Age, median (IQR[a])	68.0 (58.5-74.5)	66.0 (56.0-77.0)	.84
Gender (male:female)	9:15	12:7	.09
Body mass index, median (IQR)	22.3 (20.9-23.6)	21.7 (19.8-28.1)	.90
Smoking history, n (%)	21 (88)	18 (95)	.42
Value of serum albumin (mg/dl) <4, n (%)	10 (42)	5 (26)	.29
American Society of Anesthesiologists score >2, n (%)	14 (58)	8 (42)	.29
Charlson Comorbidity Index >2, n (%)	4 (17)	6 (32)	.25
Eastern Cooperative Oncology Group performance status scales, n (%)			.67
0	21 (88)	17 (90)	—[b]
1	3 (12)	2 (10)	—
Minimal invasive surgery, n (%)	7 (29)	13 (68)	.01
Method of gastrectomy, n (%)			.05
Wedge resection	1 (4)	6 (32)	—
Subtotal gastrectomy	19 (79)	11 (58)	—
Total gastrectomy	4 (17)	2 (11)	—
Cancer histology, n (%)			.01
Adenocarcinoma	23 (96)	13 (68)	—
Gastrointestinal stromal cancer	1 (4)	6 (32)	—
Method of participation, n (%)			.40
Caregiver	6 (25)	7 (37)	—
Independent	18 (75)	12 (63)	—
Walking steps (X1000) on postoperative day 5, median (IQR)	4.3 (4.1-4.7)	5.8 (4.5-6.1)	.01

[a]IQR: interquartile range.

[b]Not applicable.

Table 3. Adjusted multivariate analysis to predict early discharge.

Characteristic	Odds ratio (95% CI)	*P* value
Every 1000-step increment of walking	2.72 (1.17-6.32)	.02
Minimal invasive surgery	4.14 (0.71-24.03)	.11
Method of gastrectomy (reference: wedge resection)		
Subtotal gastrectomy	0.40 (0.03-5.36)	.49
Total gastrectomy	0.42 (0.01-13.26)	.63
Cancer pathology (reference: adenocarcinoma)		
Gastrointestinal stromal cancer	1.03 (0.71-3.17)	.99
Eastern Cooperative Oncology Group performance status scales (reference: 0)	0.58 (0.12-2.67)	.49

Discussion

Principal Findings

This study demonstrated the feasibility of using a mobile app connected to a wearable device to record perioperative numbers of steps in patients after major gastrectomy for GC. The PA variable based on the number of steps is a reliable parameter for predicting early discharge from the hospital. To the best of our knowledge, this is one of the innovative studies focusing on the development of comprehensive app functions to track the PA of cancer subjects [42], and we focus on the cancer patients undergoing gastrectomy. This system continues to monitor PA during the crucial perioperative period when several complications and functional/physical disorders occur and result in delayed recovery.

From our results, improved PA was associated with early discharge of the GC patients undergoing gastrectomy. It was because the patients with improved PA had resumed physical function, which was the main factor to evaluate if the patients were qualified to be discharged. One study investigating factors associated with early discharge of patients with GC undergoing gastrectomy found that several factors, such as improved PA, laboratory variables, minimally invasive surgery, and body temperature, could predict early discharge [43]. For the patients undergoing surgical procedures other than gastrectomy, the reports also supported the relationship between improved PA and early discharge [44,45]. Our study suggested that the surgical staff should regard PA as an important clinical parameter, which has previously been long ignored as research is often hindered by the challenge of adapting an easy, valid, and reliable measure to record PA [6]. However, because of the relatively small sample, the results of the statistical tests should be interpreted with care; more patients should be enrolled to validate the relationship between improved PA and early discharge.

Although medical professionals acknowledge the significance of PA in the prevention of GC and maintenance of chronic diseases [3,46], few incorporate PA counseling into routine clinic visits/care [47]. The gap between knowledge and implementation of PA in daily practice can be attributed to several reasons, including transitions of care from an inpatient stay to the community, labor of self-recording, and difficult access for medical professionals to check data. The ability to wirelessly interface with wearable and mobile devices and the application of platforms/apps provide surgical patients with a method to share health information with their surgeons. Our app is unique in the breadth and convenience of PA data that it can capture.

In contrast to the rapid growth of the field of medical mHealth, research on surgical subjects who are in much need of continuous care because of a marked decline in general conditions is limited. Among surgical mHealth, the application of wound care is the main topic for validating its clinical effect [48-50] because wound care is a unique issue for surgical patients in comparison with medically ill patients, and telephone conversations or questionnaires cannot be used to access the visual component. In addition to our developed system, a recent study used another wearable device to track step counts of patients who had undergone diverse abdominal surgeries within 1 month after discharge and showed that the mHealth app could effectively track recovery [27]. Both studies established the feasibility of PA generated by wearable devices, which should be routinely implemented in clinical services and daily life to monitor the degree of recovery of surgical patients. However, our system monitored the step counts specifically in patients who had undergone gastrectomy immediately after gastrectomy, which provided more detailed perioperative data of PA for further analysis. In this study, we developed an app that is compatible with both Android and iOS devices to enable its use with any type of mobile devices. Another strength of our study is the largest cases series addressing the use of mHealth app in the GC patients to date.

Several studies have used electronic assessment of patient-report outcomes, which has proved to be as accurate as paper and pencil administration in 1 multicenter observational cohort study [51,52]. By collecting information that was unavailable and data that were unavailable in the traditional process of perioperative care, both health care organizations and medical staff can undertake quality improvement initiatives and conduct a comprehensive analysis. Conversely, health care workers in Taiwan are exposed to high levels of occupational stress and heavy workloads [53]. Although PA is a reliable predictor of outcomes for surgical patients, collection of data and records can be another burden for health care workers. At our hospital, we have considered implementing our electronic PA system with the health information system in the future, to attenuate

workloads of the medical staff and expedite data extraction and analysis.

Limitations

This study should be evaluated in the context of some limitations. First, this study was conducted in a single medical center. The findings of this study might be specific to our subjects, but they warrant additional research in other populations. Thus, we will design a prospective randomized clinical trial to validate the efficacy of this app in patients with GC who are undergoing gastrectomy (a study group treated by wearable devices connected to the research program and a control group treated by wearable devices who are not affiliated with the research program). Second, 4.5% of patient days exhibited a technological problem that hindered data collection mostly because of the issue of forgetting to charge the battery of the wearable devices. Usually, patients needed to charge the battery every 2 to 3 days, depending on which wearable device they used; in future, devices should be improved, particularly for their battery capacity. Third, a connection between this app and other types of wearable devices was the main barrier that limited the number of participants. Currently, we provided a wearable device to patients, but continuous modification of the system is essential with upgrades of the software and associated devices, including smartphones and wearable devices. Finally, in our study, the metric for PA is daily step count. Moderate to vigorous PA (MVPA) is also an important parameter for assessing PA in surgical patients in conjunction with daily step counts. However, MVPA is relatively more stressful for early postoperative patients compared with simple measures of daily step counts. In future studies, we will validate the role of MVPA on late postoperative surgical patients as they recover to competent physical and mental functions.

Conclusions

The use of mobile apps implementing wearable devices for recording PA in patients with postoperative GC was feasible in patients undergoing gastrectomy in this study. With the support of the mHealth platform, this app offers seamless tracing of patients' recovery with a little extra burden and turns subjective PA into an objective, measurable parameter.

Acknowledgments

The authors would like to thank all members of the Department of Surgery at the National Taiwan University Hospital for their efforts and contribution. This study was supported by grants (Numbers 104-2218-E-002-028 and 106-2628-E-002-004-MY3) from the Ministry of Science and Technology, Taiwan, which had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

References

1. Macfarlane DJ, Lee CC, Ho EY, Chan KL, Chan D. Convergent validity of six methods to assess physical activity in daily life. J Appl Physiol (1985) 2006 Nov;101(5):1328-1334 [FREE Full text] [doi: 10.1152/japplphysiol.00336.2006] [Medline: 16825525]
2. Lee IM, Hsieh CC, Paffenbarger Jr RS. Exercise intensity and longevity in men. The Harvard Alumni Health Study. J Am Med Assoc 1995 Apr 19;273(15):1179-1184. [doi: 10.1001/jama.1995.03520390039030] [Medline: 7707624]
3. Abioye AI, Odesanya MO, Abioye AI, Ibrahim NA. Physical activity and risk of gastric cancer: a meta-analysis of observational studies. Br J Sports Med 2015 Feb;49(4):224-229. [doi: 10.1136/bjsports-2013-092778] [Medline: 24434186]
4. Lahart IM, Metsios GS, Nevill AM, Carmichael AR. Physical activity levels in women attending breast screening, receiving chemotherapy and post-breast cancer treatment; a cross-sectional study. Int J Environ Res Public Health 2014 May 20;11(5):5487-5496 [FREE Full text] [doi: 10.3390/ijerph110505487] [Medline: 24852599]
5. Westerterp KR. Assessment of physical activity: a critical appraisal. Eur J Appl Physiol 2009 Apr;105(6):823-828. [doi: 10.1007/s00421-009-1000-2] [Medline: 19205725]
6. Sylvia LG, Bernstein EE, Hubbard JL, Keating L, Anderson EJ. Practical guide to measuring physical activity. J Acad Nutr Diet 2014 Feb;114(2):199-208 [FREE Full text] [doi: 10.1016/j.jand.2013.09.018] [Medline: 24290836]
7. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010 Dec 15;127(12):2893-2917 [FREE Full text] [doi: 10.1002/ijc.25516] [Medline: 21351269]
8. Munene G, Francis W, Garland SN, Pelletier G, Mack LA, Bathe OF. The quality of life trajectory of resected gastric cancer. J Surg Oncol 2012 Mar 15;105(4):337-341. [doi: 10.1002/jso.22139] [Medline: 22095440]
9. Crew KD, Neugut AI. Epidemiology of gastric cancer. World J Gastroenterol 2006 Jan 21;12(3):354-362 [FREE Full text] [Medline: 16489633]
10. Price TJ, Shapiro JD, Segelov E, Karapetis CS, Pavlakis N, Van Cutsem E, et al. Management of advanced gastric cancer. Expert Rev Gastroenterol Hepatol 2012 Apr;6(2):199-208. [doi: 10.1586/egh.11.103] [Medline: 22375525]
11. O'Neill L, Moran J, Guinan EM, Reynolds JV, Hussey J. Physical decline and its implications in the management of oesophageal and gastric cancer: a systematic review. J Cancer Surviv 2018 Aug;12(4):601-618. [doi: 10.1007/s11764-018-0696-6] [Medline: 29796931]

12. Brenkman HJ, Tegels JJ, Ruurda JP, Luyer MD, Kouwenhoven EA, Draaisma WA, LOGICA Study Group. Factors influencing health-related quality of life after gastrectomy for cancer. Gastric Cancer 2018 May;21(3):524-532 [FREE Full text] [doi: 10.1007/s10120-017-0771-0] [Medline: 29067597]
13. Yu W, Park KB, Chung HY, Kwon OK, Lee SS. Chronological changes of quality of life in long-term survivors after gastrectomy for gastric cancer. Cancer Res Treat 2016 Jul;48(3):1030-1036 [FREE Full text] [doi: 10.4143/crt.2015.398] [Medline: 27004956]
14. Campbell NL, Zhan J, Tu W, Weber Z, Ambeuhl R, McKay C, et al. Self-reported medication adherence barriers among ambulatory older adults with mild cognitive impairment. Pharmacotherapy 2016 Feb;36(2):196-202. [doi: 10.1002/phar.1702] [Medline: 26890914]
15. Pettee GK, McClain JJ, Schmid KK, Storti KL, Ainsworth BE. Reliability and convergent validity of the past-week modifiable activity questionnaire. Public Health Nutr 2011 Mar;14(3):435-442. [doi: 10.1017/S1368980010002612] [Medline: 20843404]
16. Jacobs Jr DR, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med Sci Sports Exerc 1993 Jan;25(1):81-91. [Medline: 8423759]
17. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med 2003 Jun;37(3):197-206 [FREE Full text] [Medline: 12782543]
18. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
19. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
20. Wayne N, Ritvo P. Smartphone-enabled health coach intervention for people with diabetes from a modest socioeconomic strata community: single-arm longitudinal feasibility study. J Med Internet Res 2014 Jun;16(6):e149 [FREE Full text] [doi: 10.2196/jmir.3180] [Medline: 24907918]
21. Kim BY, Lee J. Smart devices for older adults managing chronic disease: a scoping review. JMIR Mhealth Uhealth 2017 May 23;5(5):e69. [doi: 10.2196/mhealth.7141] [Medline: 28536089]
22. Hung CS, Yu JY, Lin YH, Chen YH, Huang CC, Lee JK, et al. Mortality benefit of a fourth-generation synchronous telehealth program for the management of chronic cardiovascular disease: a longitudinal study. J Med Internet Res 2016 May 13;18(5):e102 [FREE Full text] [doi: 10.2196/jmir.5718] [Medline: 27177497]
23. Doocy S, Paik KE, Lyles E, Hei TH, Fahed Z, Winkler E, et al. Guidelines and mHealth to improve quality of hypertension and type 2 diabetes care for vulnerable populations in Lebanon: longitudinal cohort study. JMIR Mhealth Uhealth 2017 Oct 18;5(10):e158 [FREE Full text] [doi: 10.2196/mhealth.7745] [Medline: 29046266]
24. Ahmed I, Ahmad NS, Ali S, Ali S, George A, Saleem DH, et al. Medication adherence apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 16;6(3):e62 [FREE Full text] [doi: 10.2196/mhealth.6432] [Medline: 29549075]
25. Machado SM, Wilson EH, Elliott JO, Jordan K. Impact of a telemedicine eICU cart on sepsis management in a community hospital emergency department. J Telemed Telecare 2018 Apr;24(3):202-208. [doi: 10.1177/1357633X17691862] [Medline: 29278979]
26. Gustavell T, Langius-Eklöf A, Wengström Y, Segersvärd R, Sundberg K. Development and feasibility of an interactive smartphone app for early assessment and management of symptoms following pancreaticoduodenectomy. Cancer Nurs 2018 Mar 27 (forthcoming). [doi: 10.1097/NCC.0000000000000584] [Medline: 29596113]
27. Symer MM, Abelson JS, Milsom J, McClure B, Yeo HL. A mobile health application to track patients after gastrointestinal surgery: results from a pilot study. J Gastrointest Surg 2017 Sep;21(9):1500-1505. [doi: 10.1007/s11605-017-3482-2] [Medline: 28685388]
28. Hartman SJ, Nelson SH, Weiner LS. Patterns of Fitbit use and activity levels throughout a physical activity intervention: exploratory analysis from a randomized controlled trial. JMIR Mhealth Uhealth 2018 Feb 05;6(2):e29 [FREE Full text] [doi: 10.2196/mhealth.8503] [Medline: 29402761]
29. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
30. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
31. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. JMIR Mhealth Uhealth 2016;4(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4461] [Medline: 27122452]
32. Modave F, Guo Y, Bian J, Gurka MJ, Parish A, Smith MD, et al. Mobile device accuracy for step counting across age groups. JMIR Mhealth Uhealth 2017 Jun 28;5(6):e88 [FREE Full text] [doi: 10.2196/mhealth.7870] [Medline: 28659255]
33. Wu JM, Yu HJ, Ho TW, Su XY, Lin MT, Lai F. Tablet PC-enabled application intervention for patients with gastric cancer undergoing gastrectomy. Comput Methods Programs Biomed 2015 Apr;119(2):101-109. [doi: 10.1016/j.cmpb.2015.03.004] [Medline: 25819034]

34. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. Am J Epidemiol 2011 Mar 15;173(6):676-682. [doi: 10.1093/aje/kwq433] [Medline: 21330339]
35. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol 1982 Dec;5(6):649-655. [Medline: 7165009]
36. Fitz-Henry J. The ASA classification and peri-operative risk. Ann R Coll Surg Engl 2011 Apr;93(3):185-187 [FREE Full text] [doi: 10.1308/rcsann.2011.93.3.185a] [Medline: 21477427]
37. Wang J, Dang P, Raut CP, Pandalai PK, Maduekwe UN, Rattner DW, et al. Comparison of a lymph node ratio-based staging system with the 7th AJCC system for gastric cancer: analysis of 18,043 patients from the SEER database. Ann Surg 2012 Mar;255(3):478-485. [doi: 10.1097/SLA.0b013e31824857e2] [Medline: 22330040]
38. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: a randomized controlled pilot trial. JMIR Mhealth Uhealth 2017 Mar 06;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]
39. Xie J, Wen D, Liang L, Jia Y, Gao L, Lei J. Evaluating the validity of current mainstream wearable devices in fitness tracking under various physical activities: comparative study. JMIR Mhealth Uhealth 2018 Apr 12;6(4):e94 [FREE Full text] [doi: 10.2196/mhealth.9754] [Medline: 29650506]
40. Meyer N, Kerz M, Folarin A, Joyce DW, Jackson R, Karr C, et al. Capturing rest-activity profiles in schizophrenia using wearable and mobile technologies: development, implementation, feasibility, and acceptability of a remote monitoring platform. JMIR Mhealth Uhealth 2018 Oct 30;6(10):e188. [doi: 10.2196/mhealth.8292] [Medline: 30377146]
41. Jeong O, Ryu SY, Park YK. Postoperative functional recovery after gastrectomy in patients undergoing enhanced recovery after surgery: a prospective assessment using standard discharge criteria. Medicine (Baltimore) 2016 Apr;95(14):e3140 [FREE Full text] [doi: 10.1097/MD.0000000000003140] [Medline: 27057836]
42. Kiss N, Baguley BJ, Ball K, Daly RM, Fraser SF, Granger CL, et al. Technology-supported self-guided nutrition and physical activity interventions for adults with cancer: systematic review. JMIR Mhealth Uhealth 2019 Feb 12;7(2):e12281 [FREE Full text] [doi: 10.2196/12281] [Medline: 30747720]
43. Park JH, Son YG, Kim TH, Huh YJ, Yang JY, Suh YJ, et al. Identification of candidates for early discharge after gastrectomy. Ann Surg Oncol 2017 Jan;24(1):159-166. [doi: 10.1245/s10434-016-5447-1] [Medline: 27448119]
44. Jones D, Musselman R, Pearsall E, McKenzie M, Huang H, McLeod RS. Ready to go home? Patients' experiences of the discharge process in an enhanced recovery after surgery (ERAS) program for colorectal surgery. J Gastrointest Surg 2017 Nov;21(11):1865-1878. [doi: 10.1007/s11605-017-3573-0] [Medline: 28932946]
45. Rae A. Reasons for delayed patient discharge following day surgery: a literature review. Nurs Stand 2016 Nov 09;31(11):42-51. [doi: 10.7748/ns.2016.e10292] [Medline: 27848403]
46. Hébert ET, Caughy MO, Shuval K. Primary care providers' perceptions of physical activity counselling in a clinical setting: a systematic review. Br J Sports Med 2012 Jul;46(9):625-631. [doi: 10.1136/bjsports-2011-090734] [Medline: 22711796]
47. Diaz KM, Krupka DJ, Chang MJ, Peacock J, Ma Y, Goldsmith J, et al. Fitbit®: an accurate and reliable device for wireless physical activity tracking. Int J Cardiol 2015 Apr 15;185:138-140. [doi: 10.1016/j.ijcard.2015.03.038] [Medline: 25795203]
48. Gunter R, Fernandes-Taylor S, Mahnke A, Awoyinka L, Schroeder C, Wiseman J, et al. Evaluating patient usability of an image-based mobile health platform for postoperative wound monitoring. JMIR Mhealth Uhealth 2016 Sep 28;4(3):e113 [FREE Full text] [doi: 10.2196/mhealth.6023] [Medline: 27683059]
49. Fernandes-Taylor S, Gunter RL, Bennett KM, Awoyinka L, Rahman S, Greenberg CC, et al. Feasibility of implementing a patient-centered postoperative wound monitoring program using smartphone images: a pilot protocol. JMIR Res Protoc 2017 Feb 22;6(2):e26 [FREE Full text] [doi: 10.2196/resprot.6819] [Medline: 28228369]
50. Gunter RL, Fernandes-Taylor S, Rahman S, Awoyinka L, Bennett KM, Weber SM, et al. Feasibility of an image-based mobile health protocol for postoperative wound monitoring. J Am Coll Surg 2018 Mar;226(3):277-286. [doi: 10.1016/j.jamcollsurg.2017.12.013] [Medline: 29366555]
51. Rosner BI, Gottlieb M, Anderson WN. Accuracy of internet-based patient self-report of postdischarge health care utilization and complications following orthopedic procedures: observational cohort study. J Med Internet Res 2018 Jul 20;20(7):e10405 [FREE Full text] [doi: 10.2196/10405] [Medline: 30030212]
52. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value Health 2009 Jun;12(4):419-429. [doi: 10.1111/j.1524-4733.2008.00470.x] [Medline: 19900250]
53. Tsai YC, Liu CH. Factors and symptoms associated with work stress and health-promoting lifestyles among hospital staff: a pilot study in Taiwan. BMC Health Serv Res 2012 Jul 16;12:199 [FREE Full text] [doi: 10.1186/1472-6963-12-199] [Medline: 22799278]

Abbreviations

ASA: American Society of Anesthesiologists

GC: gastric cancer
ICD-9-CM: International Classification of Diseases, 9th Revision, Clinical Modification
ICU: intensive care unit
mHealth: mobile health
MVPA: moderate to vigorous physical activity
PA: physical activity

Edited by G Eysenbach; submitted 20.08.18; peer-reviewed by A Chwałczyńska, L Grepo, N Gell, A Mohammadabadi; comments to author 13.10.18; revised version received 13.11.18; accepted 24.03.19; published 23.04.19.

Please cite as:
Wu JM, Ho TW, Chang YT, Hsu C, Tsai CJ, Lai F, Lin MT
Wearable-Based Mobile Health App in Gastric Cancer Patients for Postoperative Physical Activity Monitoring: Focus Group Study
JMIR Mhealth Uhealth 2019;7(4):e11989
URL: http://mhealth.jmir.org/2019/4/e11989/
doi:10.2196/11989
PMID:31012858

Original Paper

The Use of Wearable Activity Trackers Among Older Adults: Focus Group Study of Tracker Perceptions, Motivators, and Barriers in the Maintenance Stage of Behavior Change

Anastasia Kononova[1], PhD; Lin Li[2], MA; Kendra Kamp[3], PhD, RN; Marie Bowen[4], MA; RV Rikard[2], PhD; Shelia Cotten[2], PhD; Wei Peng[2], PhD

[1]Department of Advertising and Public Relations, Michigan State University, East Lansing, MI, United States

[2]Department of Media and Information, Michigan State University, East Lansing, MI, United States

[3]Department of Biobehavioral Nursing and Health Informatics, University of Washington, Seattle, WA, United States

[4]Center for Innovation and Research, Michigan State University, East Lansing, MI, United States

Corresponding Author:
Anastasia Kononova, PhD
Department of Advertising and Public Relations
Michigan State University
Room 319
404 Wilson Road
East Lansing, MI, 48824
United States
Phone: 1 5174325129
Email: kononova@msu.edu

Abstract

Background: Wearable activity trackers offer the opportunity to increase physical activity through continuous monitoring. Viewing tracker use as a beneficial health behavior, we explored the factors that facilitate and hinder long-term activity tracker use, applying the transtheoretical model of behavior change with the focus on the maintenance stage and relapse.

Objective: The aim of this study was to investigate older adults' perceptions and uses of activity trackers at different points of use: from nonuse and short-term use to long-term use and abandoned use to determine the factors to maintain tracker use and prevent users from discontinuing tracker usage.

Methods: Data for the research come from 10 focus groups. Of them, 4 focus groups included participants who had never used activity trackers (n=17). These focus groups included an activity tracker trial. The other 6 focus groups (without the activity tracker trial) were conducted with short-term (n=9), long-term (n=11), and former tracker users (n=11; 2 focus groups per user type).

Results: The results revealed that older adults in different tracker use stages liked and wished for different tracker features, with long-term users (users in the maintenance stage) being the most diverse and sophisticated users of the technology. Long-term users had developed a habit of tracker use whereas other participants made an effort to employ various encouragement strategies to ensure behavior maintenance. Social support through collaboration was the primary motivator for long-term users to maintain activity tracker use. Short-term and former users focused on competition, and nonusers engaged in vicarious tracker use experiences. Former users, or those who relapsed by abandoning their trackers, indicated that activity tracker use was fueled by curiosity in quantifying daily physical activity rather than the desire to increase physical activity. Long-term users saw a greater range of pros in activity tracker use whereas others focused on the cons of this behavior.

Conclusions: The results suggest that activity trackers may be an effective technology to encourage physical activity among older adults, especially those who have never tried it. However, initial positive response to tracker use does not guarantee tracker use maintenance. Maintenance depends on recognizing the long-term benefits of tracker use, social support, and internal motivation. Nonadoption and relapse may occur because of technology's limitations and gaining awareness of one's physical activity without changing the physical activity level itself.

(JMIR Mhealth Uhealth 2019;7(4):e9832) doi:10.2196/mhealth.9832

KEYWORDS

aging; wearable electronic devices; biobehavioral sciences; transtheoretical model of behavior change; exercise; physical activity

Introduction

Trackers Are Beneficial to Older Adults, Yet, Underused

The focus of this study was on exploring the nuances of maintaining the use of wearable activity trackers and reasons to discontinue activity tracker use in the population of adults who are aged 65 years or older. Academic and industry research has shown that the use of activity trackers can increase physical activity through continuous monitoring of activity progress, motivational messages, social support, and many other empirically tested behavioral change techniques [1-4]. Activity trackers facilitate physical activity, which is beneficial for older adults because of the protective power of physical activity against diseases associated with older ages (eg, heart rate) [5].

Despite the evident benefits of activity trackers for older generations, digital care today is more available to younger populations, leaving older adults on the periphery of the industry [6]. As little as 7% of older adults owned an activity tracker in 2014 [7]. Although many adults are now aware of this technology with its increased popularity, this population still shows slow rates of adoption that depends on many factors, including activity tracker trial and price [8,9]. Only a few studies have been done to understand how and why older adults maintain the use of activity trackers and why they choose not to use or stop using this wearable technology [10-12]. Even if individuals decide to use activity trackers or are given a tracker at no cost, it does not guarantee that they will continue using them on a long-term basis. Overall, 1 in 3 activity tracker users of all adult ages stop using the device within 6 months after purchase [7,13]. The length of use correlates with age, where adults who are aged 70 years or older quit using activity trackers in only 2 weeks [14].

Little comprehensive evidence exists with regard to the long-term use of activity trackers by older adults [10,11]; motivations for long-term use [2,3]; and differences between nonuse and short- and long-term use. To study how older adults maintain the use of activity trackers and why their motivation to maintain activity tracker use grows stronger or fades over time, we applied the transtheoretical model (TTM) of behavior change, with a focus on the maintenance of desirable health behavior and relapse (discontinued activity tracker use). In the literature review below, we first introduce the theoretical framework of TTM, then discuss physical activity and activity tracker use in older adults, and finish by addressing previous research on the maintenance of activity tracker use behaviors.

Transtheoretical Model of Behavior Change

This study explored the motivations and barriers associated with the sustained use of activity trackers through the lens of the TTM. The TTM employs the stages of change to integrate processes and components of change across major theories of intervention [15-17]. A key element to the TTM is the stage of change as the construct is the temporal dimension to the framework. The TTM proposes that change is a process that unfolds over time and progresses through a series of 6 stages. In the precontemplation stage, people do not plan to change their behavior in the near term because they are either uninformed or underinformed about the consequences of their behavior. For example, older adults might not be aware of the existence of activity trackers or realize what health benefits activity tracker use entails. The costs and benefits of change potentially produce uncertainty, and some people remain in contemplation for extended periods of time. For instance, activity tracker use could be perceived as beneficial, but the price and lack of technology skills could stop older adults from adopting it. People begin to take significant steps toward the behavior change in the preparation stage. There is a plan of action and a critical stage for recruitment into action-oriented programs. In the action stage, an individual makes specific evident lifestyle modifications, such as wearing an activity tracker every day and changing daily routines to increase trackable activities.

The continuation of the specific behavior and lack of relapse is the maintenance stage. Relapse, however, is a common occurrence in the maintenance stage where individuals abandon new behaviors and return to old ones (former tracker users in this study). Older adults may successfully maintain the use of an activity tracker for a period of time; however, some may stop at some point. What makes older adults continue to use the technology and what contributes to its abandonment is the focus of this study. During the maintenance stage, individuals have made specific and apparent changes to their lifestyles, and their continuous efforts to prevent against relapse no longer require frequent applications of the change processes as one would during the action stage. Individuals in the maintenance stage (long-term activity tracker users in this study) are less likely to be tempted to revert to previous behaviors, and they have increased confidence and self-efficacy in keeping up with the changes. When making a behavior-related decision, individuals in the maintenance stage are more likely to consider and be influenced by the pros rather than the cons associated with the behavior [18]. The estimated duration of the maintenance stage is about 6 months to 5 years before the termination stage [17].

Wearable Activity Trackers Increase Physical Activity

Activity trackers refer to sensor-based wearable devices that automatically track and monitor various indicators of physical activity, such as steps taken, stairs climbed, duration and quality of sleep, pulse or heart rate, calories consumed or burned, and even mood [19]. Activity trackers synchronize these data with users' personal accounts, ensuring easy access from any device. Although the specific features available depend on tracker brands and models, older adults typically use activity trackers to monitor the distance covered, steps taken, calories burned, sleep time, and heart rate [11], making activity trackers a convenient tool for this age group that provides feedback about physical activity amount and intensity.

Physical activity in older adults reduces the risk of chronic diseases, such as cardiovascular disease, stroke, obesity, and hypertension; improves cognitive and mental health; lowers the chance of falls; and helps maintain a longer independent life [20-22]. The minimum recommended level for older adults is 150 min of moderate-to-vigorous physical activity per week [23]. Despite this recommendation, older adults constitute the most sedentary age group [24-26]. Almost 84% of older adults aged 65 years and older do not meet the aerobic and muscle-strengthening physical activity requirements [27], which makes activity trackers a particularly relevant technology for this age group. In particular, physical activity recommendations for older adults focus on moderate-intensity aerobic and muscle-strengthening activities such as walking, jogging, bicycle riding, yard work, and gardening [22,28]. Some of these activities are tracked by wearable technology.

Activity trackers have the advantage of boosting physical activity through the integration of empirically tested behavioral change techniques such as goal setting, self-monitoring, social support, social comparison, feedback, and rewards [3,4], in contrast to antecedent technologies, such as pedometers. Self-monitoring and goal setting have been especially effective in promoting self-efficacy and physical activity in interventions [29]. It has been found that even though wearing a new piece of health technology is a novel activity for older adults, they appreciate the activity tracker's contribution to self-awareness and goal setting. Activity trackers provide older adults with relatively unbiased data about basic activities. In addition, older adults view activity trackers as helpful motivators in achieving walking goals and competing with themselves [8]. Another important advantage of activity tracker use in the *65+* population is social connection. For a population that is characterized by social isolation and loneliness [30], technology that addresses social connectedness needs is perceived as helpful in overcoming barriers to increase physical activity [31].

Adults in general and, specifically, older adults who started using wearables have been shown to increase daily activity levels [1,10]. A 7-month study of 18 participants (aged 36 to 73 years) who were given a wearable tracker [32] found that 16 participants continued to use it after 7 months. The benefits of use included weight loss, social connection, and increased activity awareness. Participants aged 60 years and older who were given a tracker reduced waist circumference and increased step count during another 12-week study [2]. African American and Hispanic older female participants, who tested a newly developed tracking device in a 7-week study, increased their physical activity level, lost weight, and lowered blood pressure levels [33]. Activity trackers have been found to be more effective than their *predecessors*, where sedentary female older adults who used digital trackers significantly increased their physical activity compared with those who used pedometers [1]. A tracker that delivered prompts via short message service has also been found effective in increasing moderate-to-vigorous physical activity among overweight and obese adults [34].

Although activity trackers can be helpful in increasing physical activity, this technology is not ideal. For example, in a study with 8 adults who were aged 75 years or older, 3 participants experienced technical problems with the activity trackers, preventing them from gathering any activity feedback [35]. Participants reported that they could only get the activity tracker to work 78% of the time.

Activity Tracker Use Maintenance Among Older Adults

To date, there is a dearth of research employing the TTM to examine the efficacy of activity tracker–based interventions to maintain physical activity level, along with activity tracker use, among older adults [36,37]. Research based on a general population sample has shown that nearly three-fourth of the participants discontinued using activity trackers after 100 days from the initiation date, with most dropouts explained by technical failures and loss of activity trackers. In that study, 4 participants aged older than 65 years were among those who stopped using activity trackers within the first 100 days. Age, along with positive user experience, perceived activity tracker effects, and playing individual sports with family, positively predicted activity tracker use duration among those who continued using this technology after 100 days [38]. Another study with an older sample of patients recovering from a myocardial infarction (average age=56 years) found that activity tracker use was successfully maintained for over 1 year. Generally, participants used activity trackers several days a week but not on all days. Activity tracker use on all days was common only for the initiation period and only for a few patients [39].

Study Objective and Research Questions

Although multiple TTM studies have investigated the adoption of health-related behaviors, fewer studies have examined behavior maintenance and its abandonment after a period of long use [16]. In this study, we focused on *activity tracker use* (not physical activity per se) as a beneficial health behavior and explored the factors that contribute to the successful maintenance of this behavior and, on the contrary, the failure to maintain it among older adults. We conducted focus groups with 4 types of activity tracker users who were aged 65 years or older: long-term, short-term, former users, and nonusers. Insights from long-term users helped us understand the strategies that they employed to maintain activity tracker use and activity tracker features that encouraged such activity tracker use sustainability. Talking with short-term users allowed us to compare activity tracker use strategies and perceptions at the initial and maintenance stages of tracker use. Experiences of nonusers and former tracker users were analyzed to examine whether and how activity trackers encouraged physical activity in older adults and what were the reasons for not using or abandoning this technology.

Research question (RQ) 1a: *What features and functions do adults who are 65 and older consider useful and would like to have on their activity trackers?*

RQ1b: *What are similarities and dissimilarities in perceptions of tracker features and functions betweenadults who are 65 and olderat the maintenance stage of use (current long-term users), those who stopped using activity trackers (former users), those who were only at the initial stages of use (short-term users and non-users after the activity tracker trial), and those who are*

not familiar with the technology (non-users before the activity tracker trial)?

RQ2a: *How do perceived benefits and motivators associated with older age drive activity tracker use in adults who are 65 and older?*

RQ2b: *How do adults who are 65 and older who are tracker long-term, short-term, former users, and non-users compare in terms of perceived benefits and motivators?*

RQ3a: *What barriers do adults who are 65 and older experience during tracker use?*

RQ3b: *What are similarities and differences in perceptions of tracker use barriers amongadults who are 65 and older that are long-term, short-term, former users, and non-users of activity tracker?*

Methods

Recruitment, Participants, and Procedures

A total of 10 focus groups with adults aged 65 years and older were conducted. Furthermore, 4 activity tracker trial focus groups were conducted with tracker nonusers, where each participant attended 2 meetings. After the initial focus group, nonusers were offered an activity tracker to use for several weeks and then attended a follow-up focus group meeting. A total of 6 additional focus groups that did not involve an activity tracker trial were conducted with short- and long-term users and former users (2 focus groups were held per user type). Each focus group lasted approximately 2 hours. Each participant received US $20 for participation.

Up to 10 people were invited to each focus group with the use of convenience sampling strategies [40-42]. Participants were recruited through local senior centers, university's listservs, and an online recruitment system at Michigan State University. The online recruitment system (Sona) provides access to over 7000 community-dwelling individuals who are interested in research participation (4% are aged 61 years or older). Flyers were distributed at local churches and senior centers, posted on Facebook networking groups, and published in a local newspaper. The general participant selection criterion was that participants had to be aged 65 years or older. Focus group locations included a local community center, hospital building, senior center, and university campus.

Older adults who were interested in our study contacted the research team through phone or email. A research team member provided a brief overview of the study and asked each participant screening questions regarding the use of activity trackers (eg, length of use) and date and time preferences for the focus group. Potential participants were invited to focus groups based on tracker usage. Individuals who had never used an activity tracker were invited to the tracker nonuser group. Short-term users were the current users who had been using a tracker for less than 6 months. Long-term users were the current users who had been using the technology for 6 months or longer. Former users were individuals who had used an activity tracker in the past but stopped using it.

The focus groups were conducted by the members of the research team who developed and refined the focus group protocol. The first focus groups were conducted by researchers with substantial experience in conducting focus groups and in-depth interviews. Those who were new to the procedure observed the focus groups first and then led their own under the supervision of more experienced colleagues. A structured guide was used for each focus group.

All materials and study procedures were approved by the institutional review board at the university where the study was conducted. Participants in each focus group were given time to read the consent form, which provided the option of not participating and withdrawing at any time. Participants were given an opportunity to not answer any of the questions asked. Data confidentiality was guaranteed. None of the participants refused to answer questions or stopped participation. Participants consented to audio and video recording by signing the consent form and a video release form.

After participants signed the consent form, focus group procedure began. Participants filled out a short survey to collect information about demographics, medical conditions or disabilities, baseline activity levels, and health status. Then, they were provided instructions about the focus group procedure (eg, there were no right and wrong answers and positive and negative opinions were equally important). The conversation started with an icebreaker question about favorite personal technology. After the tone of the focus group was set, they shared their associations with the term *wearable activity tracker*, described their experiences using it or observing others using it and provided ideas about an ideal activity tracker. Then the conversation moved to an in-depth discussion of reasons to start using the technology, or the lack of thereof, and motivations for continued activity tracker use. Benefits and barriers of activity tracker use and its influence on users' lives were discussed as well. Former users also discussed reasons for abandoning the technology. At the end of each focus group, participants shared their opinions about the role of physical activity and wearable technology in maintaining one's good health and were provided an opportunity to add anything else to the discussion. After the focus group was over, participants received compensation. Nonusers in the initial focus groups also received a tracker and instructions for its use.

Focus Groups With Activity Tracker Nonusers

Overall, 4 focus groups (2 initial and 2 follow-up focus groups) were conducted with older adults who had never used an activity tracker. The first set of focus groups ran in February (N1=10) and the second set ran in May (N2=7) to account for seasonal effects on physical activity [34]. At the end of the first focus group, participants were given an activity tracker, Garmin Vivofit 2, to use for 2 weeks in February and 4 weeks in May and then they returned for a follow-up focus group to discuss benefits and barriers to activity tracker use and acceptability of the technology. May focus group participants kept their activity trackers at the end of the follow-up focus group as part of the participation incentive. The 2 trials (February and May) differed because the suggestions from February focus group participants were incorporated into the May trial procedure. The suggestions

included (1) making the trial period longer, (2) clarifying activity tracker use instructions and creating instructional and motivational videos, and (3) incentivizing participants by giving away an activity tracker as a gift (see Results for more details). Participants in the February group received a demonstration of how to put on the activity tracker, a brief overview of the features, and a written instruction book. Participants in the May group received everything from the February group plus instructional and motivational videos. The instructional videos demonstrated how to use the activity tracker (eg, how to put the tracker on and off, taking the tracker out of the wrist band, and how to use the tracker). Motivational videos were recorded with activity tracker long-term users (participants from the first long-term use focus group) who shared positive activity tracker use experiences and tips to maintain activity tracker use.

Garmin Vivofit 2 tracker is a wrist wearable device that records steps taken, calories burned, and distance traveled. The Vivofit 2 has a yearlong battery life and is water resistant up to 50 meters deep. As the Vivofit 2 does not need to be charged frequently, participants were expected to be less likely to forget the trackers on the chargers or stop wearing them. We also chose the waterproof option because of past evidence that water-related activities (eg, water aerobics) are especially popular among older adults [34]. The Vivofit 2 has a large display screen to ensure ease of reading for older adults.

Focus Groups With Activity Tracker Short-Term Users

A total of 2 focus groups (N3=2 and N4=7) were conducted with older adults who had started using activity trackers within 6 months before the focus groups (on average, they had used their activity trackers for less than 3 months). In addition, 4 of the short-term participants in the second focus group were the same participants who took part in the nonusers trial focus group in May. These participants kept their activity trackers and continued using them, which qualified them to participate in the short-term use focus groups. Garmin was the most used activity tracker in this group (n=5), followed by Fitbit (n=4).

Focus Groups With Activity Tracker Long-Term Users

A total of 2 focus groups (N5=7 and N6=4) were conducted with older adults who had been using a tracker for over 6 months at the time of focus groups. Participants had experience using either Fitbit (n=9) or Garmin (n=2). One participant had used Apple Watch in addition to her Fitbit. On average, they had used activity trackers longer than 12 months.

Focus Groups With Activity Tracker Former Users

Older adults who had previously used an activity tracker for any period of time but stopped using them before the focus groups occurred were invited to participate in 1 of 2 focus groups (N7=9 and N8=2). Participants predominantly had used Fitbit (n=7), but they also mentioned Garmin, Jawbone, Misfit, Nike, Gear Fit, manual pedometers, and special medical technology provided by a local hospital (eg, health management app). On average, their tracker use lasted for nearly 10 months.

Data Coding and Analysis

Focus group conversations were audio recorded. Audio files were deidentified and transcribed using a Web-based service. Data were inductively analyzed; an exploratory thematic analysis was performed to indicate common nodes. A total of 5 coders iteratively analyzed the transcripts using NVivo, a software for qualitative data analysis. First, all coders analyzed 2 randomly selected transcripts. After the general agreement over the codes was established based on an in-depth discussion of each code [43], 2 coders analyzed 4 additional transcripts and another 2 coders worked with the remaining 4 transcripts. The fifth coder checked the quality of coding to ensure consistency. The coders met regularly until they finished coding all transcripts to discuss their interpretations of the existing codes, introduce new codes, reach mutual agreement on them, and create new code labels [43]. Some codes were aggregated into code categories during these meetings. When a disagreement was identified, it was discussed, and then the coders recoded the transcripts following the adjusted coding procedure. Thus, the coding rubric was being refined throughout the coding process because each new set of focus groups introduced new ideas and meanings. Intercoder agreement was ensured by extensive discussions of each new code [43]. At the end of the coding procedure, NVivo results from the coders of the same transcripts were compared to ensure that the coders agreed on the most prominent codes. Nodes were aggregated in more inclusive categories, which were used to derive themes. Multimedia Appendix 1 presents the map of the most prominent codes and code categories by focus group type. These codes were used to conduct thematic analysis and further understand differences among the 4 types of users. Quotations most representative of each theme were selected per agreement of all researchers from a larger pool of quotes that corresponded to each node and node category. Research team members reviewed multiple quotes and made suggestions about which ones to include in the article.

Results

Sample Description: Demographics and Beyond

Table 1 displays information about participants' age, gender, race or ethnicity, and occupation by focus group type. Most of the focus group participants were female, especially in the long-term and former user groups. Participants in these groups were also younger. The majority of participants were white. In addition, 2 out of 3 short- and long-term users had completed graduate degrees, whereas about one-third of nonusers had only graduated high school and had some college education.

Information about chronic conditions, physical activity types and frequency, and activity tracker use length by focus group type is provided in Table 2. More short-term and former activity tracker users than other participants reported having a chronic condition. Former users showed a greater diversity of physical activity, whereas nonusers were mostly focused on walking. Walking and gardening were the most popular activities for long- and short-term users. Nonusers exhibited the lowest level of physical activity frequency.

Table 1. Participants' demographic information.

Characteristic	Nonusers (n=17)	Short-term users (n=9)	Former users (n=11)	Long-term users (n=11)	All users (N=48)
Average age (years), mean (SD)	72.9 (7.5)	72.2 (9.9)	68.9 (2.5)	68.0 (3.1)	70.8 (6.7)
Age, range	66-94	66-94	67-73	65-73	65-94
Race or ethnicity, n (%)					
White	15 (88)	7 (82)	11 (100)	9 (82)	42 (88)
African American	0 (0)	2 (18)	0 (0)	2 (18)	4 (8)
Hispanic	1 (6)	0 (0)	0 (0)	0 (0)	1 (2)
Asian	1 (6)	0 (0)	0 (0)	0 (0)	1 (2)
Gender, n (%)					
Male	7 (41)	2 (18)	2 (18)	2 (18)	13 (27)
Female	10 (59)	7 (82)	9 (82)	9 (82)	35 (73)
Education, n (%)					
High School	2 (12)	0 (0)	0 (0)	0 (0)	2 (4)
Some College	4 (18)	2 (18)	2 (18)	2 (18)	10 (21)
College	7 (41)	4 (46)	2 (18)	5 (46)	18 (38)
Graduate	4 (24)	3 (36)	7 (64)	4 (36)	18 (38)

Table 2. Participants' chronic conditions, physical activity levels, and activity tracker use length.

Health and physical activity descriptive	Nonusers (n=17)	Short-term users (n=9)	Former users (n=11)	Long-term users (n=11)	All users (N=48)
Chronic conditions, n (%)					
Arthritis	8 (47)	3 (33)	8 (73)	6 (55)	25 (52)
High blood pressure	7 (41)	6 (67)	6 (55)	3 (27)	22 (46)
Obese	2 (12)	4 (44)	2 (18)	3 (27)	11 (23)
Thyroid condition	3 (18)	3 (33)	2 (18)	2 (18)	10 (21)
Heart disease	1 (6)	2 (22)	3 (27)	1 (9)	7 (15)
Diabetes	1 (6)	4 (44)	2 (18)	0 (0)	7 (15)
Physical activities, n (%)					
Biking	3 (18)	6 (68)	7 (64)	2 (18)	18 (38)
Callisthenic classes	3 (18)	1 (11)	5 (45)	2 (18)	11 (23)
Weight lifting	4 (24)	3 (33)	8 (73)	3 (27)	18 (38)
Gardening	4 (24)	6 (68)	7 (64)	6 (55)	23 (48)
Walked	13 (77)	5 (56)	10 (91)	9 (82)	37 (77)
Water aerobics	6 (35)	1 (11)	2 (18)	3 (27)	12 (25)
Frequency of physical activity, n (%)					
0 times per week	1 (6)	0 (0)	1 (10)	1 (9)	3 (6)
Once per week	2 (12)	0 (0)	1 (10)	0 (0)	3 (6)
2-3 times per week	6 (35)	4 (44)	3 (30)	4 (36)	17 (36)
4-5 times per week	7 (41)	3 (33)	1 (10)	2 (18)	13 (28)
More than 5 times a week	1 (6)	2 (22)	4 (40)	4 (36)	11 (23)
Average length of activity tracker use	0 months	Less than 3 months	10 months (before abandonment)	Over 12 months	8 months

Table 3. Technology ownership by focus group type.

Access to technology type	Nonusers, n (%)	Short-term users, n (%)	Former users, n (%)	Long-term users, n (%)	All users, n (%)
Access to a landline phone (N=46)	11 (65)	6 (89)	8 (72)	4 (36)	29 (63)
Access to a mobile phone (N=48)	15 (88)	8 (89)	11 (100)	10 (91)	44 (92)
Access to a desktop computer (N=48)	13 (77)	8 (89)	8 (72)	8 (72)	37 (77)
Access to internet-enabled laptop computer (N=47)	11 (65)	6 (78)	9 (81)	9 (81)	35 (75)
Access to tablet computer (N=48)	11 (65)	7 (78)	10 (91)	11 (100)	39 (81)
Access to an activity tracker (N=48)	0 (0)	9 (100)	8 (72)	11 (100)	28 (58)

As part of focus group discussion, participants talked about the meaning of health and the role of physical activity and activity trackers in such perceptions. There were little differences across groups as most participants associated being healthy with freedom and independence, ability to move, being in good mental health, having high-quality life, and holding positive attitudes (Multimedia Appendix 1).

Many participants occasionally referred to themselves as being *laggards* and *luddites*, that is, not being technologically savvy. Short-term users described themselves as *having health issues* and *being active*. Long-term users also mentioned having health issues but were more likely to describe themselves as being technologically savvy and early technology adopters. Talking about their favorite technology during the icebreaker, participants mostly referred to information and communication devices, such as desktop and laptop computers, mobile phones, and tablet computers. Technology access by focus group type is shown in Table 3. Notably, the majority of long-term tracker users did not have access to landline phones. Most participants associated activity trackers with pedometers and other health management technology.

Ideal Tracker: Prettier, Bigger, and More Comfortable

RQs 1a and 2b asked about the perceptions of tracker features and functions. Nonusers before the activity tracker trial did not have any experience with the technology and many did not know what it was, so they could not clearly identify what features they favored. The only experience some nonusers had previously had with wearables was vicarious where they generated their knowledge through observing others. In their descriptions of an ideal tracker, nonusers expressed great interest in a possibility to *monitor diverse activities*, such as biking, water aerobics, golf, and others. Nonuser participants also suggested that an ideal activity tracker should count *calories burned*, monitor *heart rate*, *look good*, and be *user friendly*.

After trying activity trackers, nonusers found 4 features and functions to be useful: it could be used as a *watch* to show date and time, it had a *comfortable band*, it tracked *sleep*, and it was *waterproof*. Interestingly, the step count feature was not a favorite feature among the nonusers as most perceived step quantification to be inaccurate.

After the activity tracker trial, nonusers' preferences for an ideal activity tracker notably changed. Although they continued to suggest that activity trackers should be aesthetically appealing and not *ugly*, they focused on activity tracker band comfort, which, according to many participants, Garmin VivoFit 2 lacked. Nonusers described the band as *plastic* (ie, cheap and of bad quality), *clunky*, *annoying*, *rigid*, and *uncomfortable*:

> *It's very rigid. The design is poor. It collects water underneath. I end up having a really loose bracelet. Which could have some effect on accuracy. I don't know. I found it totally uncomfortable. It's really ugly.* [Female participant, nonuser]

Another characteristic of an ideal tracker that was mentioned by nonusers after the VivoFit 2 trial was related to design features for *better vision*. Those referred to multiple aspects of activity tracker display that made reading easier, from the size of symbols to the direction of symbol placement on the screen (horizontal vs vertical) and light.

Looking *nice*, *cool*, and *fashionable*, as well as looking like something else (eg, *watch*, *bracelet*, or *necklace*) and having a comfortable band were equally important for participants in all 4 types of focus groups. Short-term and former activity tracker users identified simple usage patterns by focusing on the importance of step count. They also wished that activity trackers were more *user friendly* for older adults when they discussed device maintenance:

Short-term users expressed an appreciation of heart rate monitoring feature and wished to have features for better vision and a *stopwatch* function on their ideal activity trackers. Former users liked the sleep-tracking feature and wanted an activity tracker that measures *blood pressure* and provides *accurate information*.

Long-term tracker users, who were in the maintenance stage, showed a much more diverse usage of activity trackers and had a more elaborate *wish list*. In addition to features popular in other focus groups, long-term users identified *motivational messages and reminders*, *pace tracking*, and *calories burned* as being important:

> *I use mine as a, as a watch, um, when I wear it. I use it for distance, I use it for time, and I use it for my pace and calories expended.* [Female participant, long-term user]

Time feature, in addition to physical activity tracking, was crucial to long-term users. They also wanted activity trackers to be waterproof, have longer *battery life*, track *physical*

progress in real time, automatically count *calories consumed*, and measure heart rate.

Maintenance: Trial, Opportunity, Togetherness, and Internal Motivation

RQs 2a and 2b asked about participants' perceptions of activity tracker use benefits and motivational strategies that they used, would use, or had used to sustain tracker usage over time. The overarching idea expressed by nonusers before the activity tracker trial was that trackers *would not motivate* them to increase physical activity. Some nonusers suggested they would be interested in *increasing awareness* of their daily activity as it could encourage them to move more. *Goal setting*, or knowing exactly what outcome to strive for, could also strengthen the activity tracker use motivation. Several nonuser participants named *long-term benefits*, such as staying away from medications, feeling better, losing weight, and improving physical indicators of health (heart rate and blood pressure), as motivations that could drive tracker use.

After trying Garmin VivoFit 2, the *activity tracker does not motivate* theme became less pronounced. Participants agreed that tracker use motivated them to walk more, driven by quantifying activity (counting steps) and continuously making users "more conscious of extra walking" (female participant, nonuser):

> *I did find myself looking all the time, how many steps I had, and did try to. In a couple days, I had like 2,000 steps and felt guilty as all get out. I really enjoyed it. I didn't think I would, I thought maybe for a week or two, the novelty you know. In fact, I'm monitoring it as I came upstairs today, so I'm really enjoying it.* [Male participant, nonuser]

The overwhelming tactic that the nonusers used was to slightly *modify daily activities* to take more steps. Those included taking stairs instead of elevators, parking far away from destination buildings, or walking to an area instead of driving.

Short-term and former users talked about the same tactics that nonusers did, with an addition of 2 new *tricks*: using a *red-line* activity tracker alert of inactivity and shopping.

Short-term and former users were also motivated by the *presence of others*: either participants competed with their friends and family members in numbers of steps taken, were supported by them (eg, being given a tracker as a gift), or simply observed them using activity trackers:

> *I would say one of the things that I did, my kids had gotten some wearables, and so we had a little contest. That encouraged me, instead of having the red line or something like that. It was like, "Oh, so how many did you do today?" It was a non-threatening way of getting everybody to see what they did, so, "You only did so many steps today? What's going on with you?"* [Female participant, former user]

The presence of others was the top motivator for long-term users. Although competition was an important motivator, the importance of "togetherness" and cooperation, that is, doing something enjoyable together to achieve individual fitness goals, was more prevalent in conversations with long-term users:

> *When we lived in Houston, and we both worked and carried pagers and cell phones or whatever, Friday night was date night and we started date night with a run together. And, cooled down, and then went out to dinner. So, it's part of our marriage in a way, and it was part of our dating relationship and that's when I think about, it's just like brushing your teeth every morning.* [Female participant, long-term user]

Achieving goals was very satisfactory to long-term users who derived gratifications from using activity trackers that confirmed completion of daily activity tasks and gave *rewards*:

> *It's just amazing when it tells me I've gotten my 10,000 steps in, I've gotten my ten stairs, I've burned my 2,000+ calories, whatever, and I get this little green flashy thing on my phone, it's stupid I get it, but I look for that at night. "Oh did I get my flashy thing?" Ya know? It's, it's motivating.* [Female participant, long-term user]

Long-term users were more likely than participants in other focus groups to see long-term benefits of regular physical activity and, consequently, activity tracker use, such as physical and emotional health (eg, stress reducer), reduced pain, independence and mobility, and long life:

> *When you start seeing your classmates in the obituaries, I think that...* [Female participant, long-term user]
>
> *It's an eye-opener, isn't it?* [Female Participant, long-term user]
>
> *Or you see them and you think, "Boy, do they look old!" [laughs] you know? "I don't want to be that person...* [Female participant, long-term user]

Long-term users, compared with other participants, perceived external factors, such as change of weather, as *opportunities*. Seasons, especially summer, facilitated physical movement. One participant found an opportunity to be physically active in winter by shoveling snow (male participant, long-term user). Long-term users, however, expressed that external factors did not fully determine motivation. *Forcing* oneself, being consistently active day to day, allowing bad days but *reimbursing* for them later was indicative of *internal motivation* playing the most important role in maintaining physical activity. Technology came in second, only as a facilitator:

> *It's something that clicks in your mind. You have to make that commitment, and once you do, the technology is very motivating. But, until you take the step of getting it and tuning into it, it isn't going to work.* [Female participant, long-term user]

What's Stopping Them: Data Inaccuracy and Lack of Adequate Instructions

RQs 3a and 3b asked about perceived barriers to tracker use. Participants in the former and shot-term user focus groups appeared to like their routines and established exercise schedule, to which activity trackers did not add. To them, physical activity

came before the use of the device. It was not that participants were not motivated to be active (they actually were), but activity tracker use did not seem to be the driver of that motivation:

> *[...] I've done 9,000 steps and it's 8:30 at night and it's my time to sit and quilt. I'm not going to do anymore. [laughs] I figured out I've done plenty.* [Female participant, short-term user]

Tracker nonusers mostly generated questions and speculations about how trackers worked: "Are they waterproof?"; "Does it monitor your sleep?"; "Do we have to keep it on our wrist?"; "Is it physically moving or is it your wrist that tracks us?"; "Does it tell time?"; and "Is it counting calories you're using?" With the lack of knowledge came skepticism as nonusers questioned the ways in which activity trackers worked. Inaccuracy in counting steps raised most of the concerns. Participants explained that physical activities were very diverse and had different levels of intensity that no technology could grasp.

After the activity tracker trial was over, nonusers continued to ask questions about how the device worked. Their skepticism toward and disappointment with the technology increased because they found that activity trackers had given inaccurate information. It also had limitations in terms of features offered. Participants compared actual step recordings with their perceived activity and found obvious inconsistencies:

> *I thought they were highly inaccurate. I clocked 1,350 steps just driving out to eat in Rapids one day.* [Female participant, nonuser]

After the tracker trial, nonusers also suggested to improve activity tracker use *instructions* tailored to the population of older adults. They expressed the need to be better educated about various device features and functions, as well as synchronization with other devices, and suggested using step-by-step detailed instructions and in-group learning environment. In the first nonuser set of focus groups held in February, participants suggested that video instructions (similar to YouTube videos) would work well to guide and motivate older adults in tracker use. Nonuser participants in May groups, thus, were provided with such videos (see Methods). However, most May nonuser participants did not watch the videos, and those who watched them were not motivated more than they normally would be. Some participants experienced a classic case of the third-person effect [44]; they were sure that others "but not me" would be motivated by the videos.

Other reasons for nonadoption included high price, physical limitations that older adults face as well as inactive lifestyle, little interest and curiosity in trying activity trackers, band discomfort, and technology use difficulties. Technology-related barriers were discussed more extensively after the nonusers tried activity trackers, especially with regard to mobile phone ownership and synchronizing activity trackers with it.

Short-term and former users added that defects in activity trackers could or did make them stop using the activity tracker. One short-term user briefly discussed the issue of keeping personal data private when information from the activity tracker is synchronized with a mobile phone or computer and shared on social networks.

Former users indicated that long-term use of activity trackers was not necessary as one gets an impression of numbers related to certain activities. Many former users were found to maintain the same daily routines, so they quickly learned how many steps, calories, and miles daily activities were associated. After that, using an activity tracker was not a priority for them:

> *I know that I get about 13,000 steps. I don't use it every day. I started wearing it again when I said I would sign up for it just to make sure that I was still doing what I thought I was doing because I walk in the morning before I go to work, and then do my activities.* [Female participant, former user]

A few former users also indicated that having a short battery life and losing activity trackers were the reasons why they stopped using the device.

Conversations with nonusers and short-term and former users centered on barriers to adopt or continue using an activity tracker, but this theme was less pronounced in long-term users who predominantly discussed activity tracker features and motivations to use. No unique additional barriers were discovered in focus groups with long-term users.

Discussion

Principal Findings

From the Basics to Sophistication

The results indicated that focus group participants, who were aged 65 years and older, favored and expected their activity trackers to have a wider variety of features than what had been shown in previous research within this population [8,14]. Participants with different levels of activity tracker use experience liked and wished to use different activity tracker features. Long-term users who were in the maintenance stage of activity tracker use indicated a more diverse and sophisticated usage of activity trackers. This could be explained by the length and richness of experience with activity trackers as well as the higher level of technological savviness mentioned by the long-term users. Nonusers expressed disappointment with the technology as it did not meet their expectations regarding accurate step count and did not track various activities automatically. Health-related monitoring, such as weight watching, hear rate, and blood pressure were not the most popular features despite many participants reported having chronic conditions.

From Effort to Effortlessness

Nonusers and short-term users had to *trick* themselves into using activity trackers to countercondition inactivity by slight modifications in daily routines. Long-term users were more habitual, automatic, and effortless in their activity tracker use. Long-term users emphasized the importance of internal motivation (*Just do it*) where activity trackers were serving as secondary facilitators and expressed the enthusiasm about modifying the environment to *keep going*.

Maintenance Through Internal Motivation

It is recommended to focus on the importance of intrinsic over extrinsic motivation for tracker use maintenance in future studies. The distinction between the 2 types of motivation is made within the framework of self-determination theory (SDT). SDT studies have consistently demonstrated that intrinsic motivations are predictors of long-term physical activity adherence and weight loss [45,46]. SDT posits that motivation is driven by individuals striving to satisfy 3 essential needs: autonomy (independence in the world of external constraints), competence (self-efficacy), and social relatedness [47-49]. Intrinsic motivation refers to engaging in an activity out of one's authentic interest in it, which brings inherent satisfaction and feelings of enjoyment, accomplishment, and excitement [47,50]. For example, the long-term users in this study not only emphasized the *Just Do It* aspect of sustained activity tracker use but also referred to positive emotional rewards they would get after completing their goals. Extrinsic motivation refers to engaging in activities that provide rewards from the outside. Extrinsic motivation with a greater degree of autonomy (eg, increased socialization via the activity tracker network) leads to a higher likelihood of behavior maintenance than external motivation with low autonomy levels (eg, activity tracker use because of a doctor's prescription) that is only effective in the short run [47,51-55]. The results of this study suggest that long-term use of activity trackers depends mostly on intrinsic motivation and extrinsic motivation at greater level of autonomy, suggesting that this behavior can be successfully maintained over time. Satisfaction of the social relatedness need through tracker use is another powerful driver of maintenance.

Observe-Compete-Collaborate

Social support was not a pronounced motivation for nonusers who often observed others using activity trackers, and it was only secondary for short-term and former users who emphasized the importance of competition with others. Long-term users indicated social support to be the main motivational factor, with the focus on building relationships around daily activity routines. Long-term users were better prepared to modify the social environment around them to maintain an active lifestyle, receive positive feedback, and seek accountability from others.

Awareness Before Health Benefits

Short-term, long-term, and former participants exhibited high levels of physical activity, which could have contributed to curiosity about activity tracking but not to the understanding of its health management value. Nonusers were unaware of tracker potential in helping to meet physical activity goals. Rising consciousness occurred among some tracker nonusers as they realized that activity tracker use could benefit them in terms of feeling better, losing weight, and lowering blood pressure. Nonusers reported lower physical activity frequency and said activity tracker use helped them increase physical activity levels. This result suggests that tracker use could be more useful for those older adults who are not physically active and seek to change that. Our short-term and former users who showed high levels of physical activity indicated greater skepticism and disappointment with the technology. Activity trackers provided them with information about physical activity, but it did not motivate to make already active users be more active. It could be that because activity trackers did not drastically change physical activity levels in the group of former users, they failed to maintain its use.

Achieving a Pros-and-Cons Balance

Long-term users were more engaged in the discussion of activity tracker features and motivations, whereas participants in other groups focused on problems with the activity tracker related to the lack of activity tracker knowledge and tracker use skills as well as inaccuracy and defects in activity tracker measures. Long-term users also saw many more benefits of activity tracker use than their counterparts from other focus groups. This result is in line with the proposition of the TTM that the maintenance stage of behavior change is associated with seeing a greater number of benefits than costs [18]. Former users not only reported the same high level of physical activity as short-term and long-term users but also exhibited a greater diversity of physical activity, which indicated that the activity tracker was not the primary motivator.

Comparison With Previous Work

This study contributes to the existing knowledge about wearable technology use in 4 meaningful ways. First, we analyzed the experiences of older adults at different levels of tracker use with a special focus on the maintenance stage of behavior change and relapse (ie, discontinuing tracker use) [16]. Previous studies have explored activity tracker use mostly at the initial stages of use [2,8,10,56] and have not systematically evaluated the spectrum of features used, liked, and wanted by adults aged 65 years and older. In addition, little research has been previously done on tracker use relapse in older adults, and our study filled that gap.

Second, we outlined several new activity tracker features and characteristics that adults in the *65+* age group would want to benefit from on a long-term basis. One of the significant findings of this study was that our participants valued tracker device comfort and aesthetics more than the basic features of step and calorie count, sleep and heart rate tracking, and water resistance that were found important for older adults in previous research [8,11]. The suggestion of making the activity tracker accommodate the older target groups in terms of features for better vision came afterward. A possible explanation of these results could be that older adults felt underappreciated by the current consumer technology developers that often target younger populations. Thus, they required devices to meet basic consumer needs in comfort and pleasant appearance to break the association with *bulky* and *ugly* devices *for older adults*. Another possibility is that our sample is skewed to be more highly educated than the average older adults; thus, their expectations may vary in terms of activity tracker characteristics. Further research with more diverse samples of older adults might yield different findings.

Third, this study indicated that health benefits provided to adults who are aged 65 years and older by activity tracker use were not greatly important to them. Participants in all focus groups put awareness of physical activity and curiosity as the primary reasons to start and continue activity tracker use. In general, the

motivators to continuously use the tracker mirrored those mentioned in previous literature: physical activity awareness, goal setting, positive reinforcement, and social connection [1-4,8,10,29-34,57,58]. Participants in our study, like participants in other studies [59], saw the benefits of physical activity awareness via self-monitoring and goal setting. Our findings regarding social support indicated that social support is not only one of the most important and consistent predictors of physical activity adherence [60-63] but also a crucial factor to maintain tracker use. Furthermore, we found that nonusers expected activity trackers to track diverse activities and were disappointed that the technology did not do it. This is echoed by the findings of another study advocating for a more tailored approach to tracking activities salient to older adults [64].

Finally, this study added to the list of the known barriers to activity tracker technology adoption and continued use and explained why relapse in activity tracker use happens. In addition to technology defects, lack of technology use skills, physical and psychological limitations, and financial restrictions [35,65], our participants added distrust in the activity tracker's capability of accurate data tracking. The questions that participants posed about activity tracker functions, speculations about how this technology works, and skepticism about it indicated a need for future studies of older adults' technology literacy [66]. Former users were motivated by knowing the numbers behind their daily physical routines, but they were not eager to use activity trackers to modify or increase these daily physical activities.

Study Limitations and Suggestions for Future Research

Though our study was extensive in the type of activity tracker users included, the sample was more educated, slightly younger, and less ethnically and socioeconomically diverse than older adults in general. For example, 23% of older adults in the *65+* age group in the United States were members of racial or ethnic minority populations in 2016 [67], which is higher than the percentage of our sample participants who were not white (less than 15%). In terms of educational attainment, approximately 86% of older adults completed high school and about 30% completed a bachelor's degree or higher in 2017 [67]. Our sample was highly educated, with over 40% of participants holding a graduate degree. Furthermore, 44 of the 48 participants (92%) had access to a mobile phone, which is higher than the national average of 78% [42]; 39 out of 48 (81%) had access to a tablet computer, which is also higher than the national rate for older adults (32%) [42]; and 36 out of 47 (75%) reported having access to an internet-enabled laptop computer, which is also higher than the national average of 55% [42]. Higher levels of education, income, technology access, younger age, and greater homogeneity in terms of race or ethnicity could influence the level of physical activity in older adults and, as a result, emphasize certain perceptions of activity trackers detected in this study. For example, only 6% of our participants did not engage in physical activity and another 6% engaged in it only once a week. This is much lower than the national average where 84% of adults aged 65 years and older live sedentary lifestyles [24-26]. Perhaps demographic and physical activity characteristics of our sample led to skeptical perceptions of activity trackers, emphasis on internal motivation to maintain activity tracker use and physical activity, as well as viewing the activity tracker as the tool to increase the awareness of physical activity rather than physical activity itself. That is why, possibly, activity tracker use was found to be most beneficial to nonusers who were less active and had lower technology access. Future studies should focus on sedentary older populations to explore activity tracker use maintenance and relapse. Studies with more diverse samples of older adults might reveal other facilitators and barriers to using activity trackers. In addition, the study solely focused on adults who are aged 65 years and older and did not collect empirical data from younger activity tracker users. This shortcoming should be addressed in future studies.

We had difficulty recruiting short-term tracker users. This leads us to question activity tracker uptake among older adults, at least in the Midwest state where this study took place. To our surprise, it was easier to find participants who had been using activity trackers for over 6 months than those who had just started using this technology. In addition, we used only 1 type of activity tracker with our nonuser participants. The limitations of Garmin Vivofit 2 could have had an effect on nonusers' evaluations of activity trackers in general.

Finally, including participants with specific types of chronic diseases could yield specific activity tracker use facilitators and barriers for these groups that might influence future studies. We were not able to focus on specific chronic disease groups in this study. Long-term randomized controlled trials that incorporate multiple-level interventions are needed to enhance physical activity among older adults.

Practical Implications

The findings of this study have several practical implications for activity tracker design and health interventions involving activity trackers. Many commercial trackers do not come with a detailed manual other than simple book leaflets and a URL for the users to refer to. It is assumed that the use is intuitive, or users can use the internet to find additional information. As one of the primary barriers to the device adoption among older adults was lack of knowledge, activity tracker manufacturers or researchers attempting to use the technology as health promotion devices may consider developing detailed manuals with screen captures and visual illustrations of features, buttons, and navigation in the app, especially if the target users are older adults who are used to having a hard-copy manual. Step-by-step instructions of use may be a good way to educate older adult users who are not technologically savvy [68]. If the promotional materials of the activity tracker manufacturers highlight the ease of use, older adults may be more likely to take an action to adopt the activity tracker.

Skepticism regarding the accuracy of activity trackers and their ability to capture various physical activities was another major barrier to adoption and activity tracker use maintenance. Activity tracker designers and manufacturers could include in their promotional materials explanations of how the activity tracker works and acknowledge that activity trackers cannot capture all activities including bicycling and swimming because of the nature of the accelerometers used in activity trackers for tracking activity. However, users should be able to manually enter the data of these activities not captured by trackers and still get an

estimate of equivalent "step count" based on type, intensity, and duration.

Physical appearance and comfort were cited by many users, both long-term and short-term. Therefore, making the band of the activity tracker to be similar to a bracelet or another piece of jewelry that can match outfits or express identity seems to be a good strategy. This may be especially useful when individuals are already at the maintenance stage of behavior change to keep them engaged.

Many former users complained about the battery life of activity trackers, and it was identified as one of the reasons for abandonment. Battery capacity is an issue when the activity tracker is designed to be small. However, as we know that older adults prefer to have a bigger screen, we suggest that when activity tracker designers are faced with the dilemma of a small-sized activity tracker and larger battery capacity, they should put priority on battery life when creating technology for older populations to ensure sustained use.

Social support was considered to be a major facilitating factor among long-term users to use activity trackers to keep physically active. Although friendly competition was mentioned by some long-term participants, the majority of them relied on not competition but cooperation or collaboration—working together to be physically active. This indicates that activity tracker designers may add features to activity trackers and their associated apps to facilitate social support and the concept of working out or walking together with family and friends.

Conclusions

Our research suggests that there is no *magic bullet* approach to ensuring that older adults will use activity trackers on a long-term basis. Focusing on individual-, interpersonal-, and community-level factors that predict the maintenance of activity tracker use behavior will likely be needed to help older adults effectively use activity trackers to become and stay more physically active. In addition, activity tracker developers and manufacturers should consider the design aspects that may be most relevant for older adults, given the rapidly increasing size of this demographic group around the world.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Number of sources (focus groups) and references (mentions) for codes and code categories.

[DOCX File, 152KB - mhealth_v7i4e9832_app1.docx]

References

1. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a Fitbit-based physical activity intervention for women. Am J Prev Med 2015 Sep;49(3):414-418 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]
2. O'Brien T, Troutman-Jordan M, Hathaway D, Armstrong S, Moore M. Acceptability of wristband activity trackers among community dwelling older adults. Geriatr Nurs 2015 Apr;36(2 Suppl):S21-S25. [doi: 10.1016/j.gerinurse.2015.02.019] [Medline: 25771957]
3. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. JMIR Mhealth Uhealth 2016 Apr 27;4(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4461] [Medline: 27122452]
4. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug 15;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
5. US Department of Health and Human Services. 2008. Physical Activity Guidelines Advisory Committee Report, 2008: To the Secretary of HealthHuman Services URL: https://health.gov/paguidelines/report/pdf/CommitteeReport.pdf [accessed 2018-12-15] [WebCite Cache ID 74gvHtEni]
6. Levine DM, Lipsitz SR, Linder JA. Trends in seniors' use of digital health technology in the United States, 2011-2014. J Am Med Assoc 2016 Aug 02;316(5):538-540. [doi: 10.1001/jama.2016.9124] [Medline: 27483069]
7. Ledger D, McCaffrey D. Endeavour Partners. Cambridge, MA: Endeavour Partners; 2014. Inside Wearables: How the Science of Human Behavior Change Offers the Secret to Long-Term Engagement URL: https://medium.com/@endeavourprtnrs/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3 [accessed 2019-03-20] [WebCite Cache ID 770VXqxR9]
8. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]

9. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]
10. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. J Am Med Assoc 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
11. Li L, Peng W, Kamp K, Bowen M, Cotten S, Rikard R, et al. Poster: Understanding long-term adoption of wearable activity trackers among older adults. : ACM; 2017 Presented at: Proceedings of the 2017 Workshop on Wearable Systems Applications; June 19-23, 2017; New York, NY p. A. [doi: 10.1145/3089351.3089819]
12. Mercer K, Li M, Grindrod KA. Do wearable activity trackers have a place in pharmacies? Can Pharm J (Ott) 2015 May;148(3):134-137 [FREE Full text] [doi: 10.1177/1715163515579221] [Medline: 26150885]
13. Karapanos E, Gouveia R, Hassenzahl M, Forlizzi J. Wellbeing in the making: peoples' experiences with wearable activity trackers. Psychol Well Being 2016;6(1):4 [FREE Full text] [doi: 10.1186/s13612-016-0042-6] [Medline: 27376017]
14. American Association of Retired Persons (AARP) CATALYST. 2015. Building a Better Tracker: Older Consumers Weigh in on Activity and Sleep Monitoring Devices. Project Catalyst URL: https://www.aarp.org/content/dam/aarp/home-and-family/personal-technology/2015-07/innovation-50-project-catalyst-tracker-study-AARP.pdf [accessed 2018-12-15] [WebCite Cache ID 74gvwO5sq]
15. 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]
16. Prochaska JO, Reddings CA, Evers KE. The transtheoretical model and stages of change. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass; 2008.
17. 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]
18. Prochaska J. Decision making in the transtheoretical model of behavior change. Med Decis Making 2008;28(6):845-849. [doi: 10.1177/0272989X08327068] [Medline: 19015286]
19. Asimakopoulos S, Asimakopoulos G, Spillers F. Motivation and user engagement in fitness tracking: heuristics for mobile healthcare wearables. Informatics 2017 Jan 22;4(1):5 [FREE Full text] [doi: 10.3390/informatics4010005]
20. Manini TM, Pahor M. Physical activity and maintaining physical function in older adults. Br J Sports Med 2009 Jan;43(1):28-31 [FREE Full text] [doi: 10.1136/bjsm.2008.053736] [Medline: 18927164]
21. Centers of Disease Control and Prevention. 1999. Physical Activity and Health: A Report of the Surgeon General URL: https://www.cdc.gov/nccdphp/sgr/olderad.htm [accessed 2018-06-07] [WebCite Cache ID 701kx2eA2]
22. Nelson ME, Rejeski WJ, Blair SN, Duncan PW, Judge JO, King AC, American College of Sports Medicine, American Heart Association. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. Circulation 2007 Aug 28;116(9):1094-1105. [doi: 10.1161/CIRCULATIONAHA.107.185650] [Medline: 17671236]
23. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007 Aug;39(8):1423-1434. [doi: 10.1249/mss.0b013e3180616b27] [Medline: 17762377]
24. Craig CL, Russell SJ, Cameron C, Bauman A. Twenty-year trends in physical activity among Canadian adults. Can J Public Health 2004 Feb;95(1):63. [doi: 10.1007/BF03403636] [Medline: 14768744]
25. Caspersen CJ, Pereira MA, Curran KM. Changes in physical activity patterns in the United States, by sex and cross-sectional age. Med Sci Sports Exerc 2000 Sep;32(9):1601-1609 [FREE Full text] [doi: 10.1097/00005768-200009000-00013] [Medline: 10994912]
26. Bennett J, Winters-Stone K, Nail L, Scherer J. Definitions of sedentary in physical-activity-intervention trials: a summary of the literature. J Aging Phys Act 2006 Oct;14(4):456-477. [Medline: 17215562]
27. Center for Disease Control and Prevention. Adult Participation in Aerobic and Muscle-Strengthening Physical Activities - United States, 2011 URL: https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6217a2.htm [accessed 2018-01-29] [WebCite Cache ID 701l1G4a4]
28. Elsawy B, Higgins KE. Physical activity guidelines for older adults. Am Fam Physician 2010 Jan 01;81(1):55-59 [FREE Full text] [Medline: 20052963]
29. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
30. Newall NE, Menec VH. Loneliness and social isolation of older adults: why it is important to examine these social aspects together. J Soc Pers Relationships 2017 Dec 27;36(3):925-939. [doi: 10.1177/0265407517749045]
31. Fan C, Forlizzi J, Dey A. Considerations for technology that support physical activity by older adults. In: Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility. New York, NY: ACM; 2012 Presented at: 14th international ACM SIGACCESS conference on Computers and accessibility; October 22-24, 2012; Boulder, CO p. 33-40. [doi: 10.1145/2384916.2384923]

32. Randriambelonoro M, Chen Y, Pu P. Can fitness trackers help diabetic and obese users make and sustain lifestyle changes? Computer 2017 Mar;50(3):20-29. [doi: 10.1109/MC.2017.92]
33. Sookhai L, Coppola JF, Gaur C. Intergenerational activity tracker program: impact with health related outcomes on older adults. 2015 Presented at: Eleventh Annual 2015 IEEE Long Island Systems, Applications and Technology Conference; May 1, 2015; Farmingdale, NY p. 7 URL: https://www.researchgate.net/publication/315956465_Can_Fitness_Trackers_Help_Diabetic_and_Obese_Users_Make_and_Sustain_Lifestyle_Changes [doi: 10.1109/LISAT.2015.7160218]
34. Wang JB, Cadmus-Bertram LA, Natarajan L, White MM, Madanat H, Nichols JF, et al. Wearable sensor/device (Fitbit One) and SMS text-messaging prompts to increase physical activity in overweight and obese adults: a randomized controlled trial. Telemed J E Health 2015 Oct;21(10):782-792 [FREE Full text] [doi: 10.1089/tmj.2014.0176] [Medline: 26431257]
35. Ehn M, Eriksson LC, Åkerberg N, Johansson AC. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 01;6(2):e34 [FREE Full text] [doi: 10.2196/mhealth.8345] [Medline: 29391342]
36. Compernolle S, Vandelanotte C, Cardon G, De Bourdeaudhuij I, De Cocker K. Effectiveness of a web-based, computer-tailored, pedometer-based physical activity intervention for adults: a cluster randomized controlled trial. J Med Internet Res 2015 Feb 09;17(2):e38 [FREE Full text] [doi: 10.2196/jmir.3402] [Medline: 25665498]
37. Devereux-Fitzgerald A, Powell R, Dewhurst A, French DP. The acceptability of physical activity interventions to older adults: a systematic review and meta-synthesis. Soc Sci Med 2016 Dec;158:14-23. [doi: 10.1016/j.socscimed.2016.04.006] [Medline: 27104307]
38. Hermsen S, Moons J, Kerkhof P, Wiekens C, De Groot M. Determinants for sustained use of an activity tracker: observational study. JMIR Mhealth Uhealth 2017 Oct 30;5(10):e164 [FREE Full text] [doi: 10.2196/mhealth.7311] [Medline: 29084709]
39. Meyer J, Schnauber J, Heuten W, Wienbergen H, Hambrecht R, Appelrather HJ, et al. Exploring longitudinal use of activity trackers. 2016 Presented at: 2016 IEEE International Conference on Healthcare Informatics (ICHI); Oct 4-7, 2016; Chicago, IL. [doi: 10.1109/ICHI.2016.29]
40. Valerio MA, Rodriguez N, Winkler P, Lopez J, Dennison M, Liang Y, et al. Comparing two sampling methods to engage hard-to-reach communities in research priority setting. BMC Med Res Methodol 2016 Oct 28;16(1):146 [FREE Full text] [doi: 10.1186/s12874-016-0242-z] [Medline: 27793191]
41. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. BMC Med Res Methodol 2014 Mar 25;14:42 [FREE Full text] [doi: 10.1186/1471-2288-14-42] [Medline: 24669751]
42. Guetterman TC. Descriptions of sampling practices within five approaches to qualitative research in education and the health sciences. Forum Qual Soc Res 2015;16(2) [FREE Full text] [doi: 10.17169/fqs-16.2.2290]
43. Feng GC. Intercoder reliability indices: disuse, misuse, and abuse. Qual Quant Int J Methodol 2014 May;48(3):1803-1815. [doi: 10.1007/s11135-013-9956-8]
44. Davison WP. The third-person effect in communication. Public Opin Q 1983 Jan 01;47(1):1-15. [doi: 10.1086/268763]
45. Mata J, Silva MN, Vieira PM, 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]
46. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012 Jun 22;9:78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
47. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68] [Medline: 11392867]
48. Ryan RM, Deci EL. The darker and brighter sides of human existence: basic psychological needs as a unifying concept. Psychol Inq 2000 Oct;11(4):319-338. [doi: 10.1207/S15327965PLI1104_03]
49. Silva de Lima AL, Hahn T, de Vries NM, Cohen E, Bataille L, Little MA, et al. Large-scale wearable sensor deployment in Parkinson's patients: the Parkinson@home study protocol. JMIR Res Protoc 2016 Aug 26;5(3):e172 [FREE Full text] [doi: 10.2196/resprot.5990] [Medline: 27565186]
50. Deci E. Intrinsic Motivation. New York,NY: Plenum Press; 1975.
51. Deci E, Ryan R, Richard M. Intrinsic Motivation And Self-Determination In Human Behavior. New York, NY: Springer US; 1985.
52. Ryan RM, Plant RW, O'Malley S. Initial motivations for alcohol treatment: relations with patient characteristics, treatment involvement, and dropout. Addict Behav 1995;20(3):279-297. [doi: 10.1016/0306-4603(94)00072-7] [Medline: 7653312]
53. Williams GC, Freedman ZR, Deci EL. Supporting autonomy to motivate patients with diabetes for glucose control. Diabetes Care 1998 Oct;21(10):1644-1651. [doi: 10.2337/diacare.21.10.1644] [Medline: 9773724]
54. Williams GC, Grow VM, Freedman ZR, Ryan RM, Deci EL. Motivational predictors of weight loss and weight-loss maintenance. J Pers Soc Psychol 1996 Jan;70(1):115-126. [doi: 10.1037/0022-3514.70.1.115] [Medline: 8558405]
55. Williams GC, Rodin GC, Ryan RM, Grolnick WS, Deci EL. Autonomous regulation and long-term medication adherence in adult outpatients. Health Psychol 1998 May;17(3):269-276. [Medline: 9619477]

56. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older adults' experiences using a commercially available monitor to self-track their physical activity. JMIR Mhealth Uhealth 2016 Apr 13;4(2):e35 [FREE Full text] [doi: 10.2196/mhealth.5120] [Medline: 27076486]
57. Olanrewaju O, Kelly S, Cowan A, Brayne C, Lafortune L. Physical activity in community dwelling older people: a systematic review of reviews of interventions and context. PLoS One 2016;11(12):e0168614 [FREE Full text] [doi: 10.1371/journal.pone.0168614] [Medline: 27997604]
58. van Stralen MM, De Vries H, Mudde AN, Bolman C, Lechner L. Determinants of initiation and maintenance of physical activity among older adults: a literature review. Health Psychol Rev 2009 Sep;3(2):147-207. [doi: 10.1080/17437190903229462]
59. Mansi S, Milosavljevic S, Tumilty S, Hendrick P, Higgs C, Baxter DG. Investigating the effect of a 3-month workplace-based pedometer-driven walking programme on health-related quality of life in meat processing workers: a feasibility study within a randomized controlled trial. BMC Public Health 2015 Apr 22;15:410 [FREE Full text] [doi: 10.1186/s12889-015-1736-z] [Medline: 25895747]
60. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJ, Martin BW, Lancet Physical Activity Series Working Group. Correlates of physical activity: why are some people physically active and others not? Lancet 2012 Jul 21;380(9838):258-271. [doi: 10.1016/S0140-6736(12)60735-1] [Medline: 22818938]
61. Eyler AA, Brownson RC, Donatelle RJ, King AC, Brown D, Sallis JF. Physical activity social support and middle- and older-aged minority women: results from a US survey. Soc Sci Med 1999 Sep;49(6):781-789. [doi: 10.1016/S0277-9536(99)00137-9] [Medline: 10459889]
62. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
63. Rovniak LS, Kong L, Hovell MF, Ding D, Sallis JF, Ray CA, et al. Engineering online and in-person social networks for physical activity: a randomized trial. Ann Behav Med 2016 Dec;50(6):885-897 [FREE Full text] [doi: 10.1007/s12160-016-9814-8] [Medline: 27405724]
64. Fausset CB, Mitzner TL, Price CE, Jones BD, Fain BW, Rogers WA. Older adults' use of and attitudes toward activity monitoring technologies. Proc Hum Factors Ergon Soc Annu Meet 2013 Sep 30;57(1):1683-1687. [doi: 10.1177/1541931213571374]
65. Floegel TA, Giacobbi Jr PR, Dzierzewski JM, Aiken-Morgan AT, Roberts B, McCrae CS, et al. Intervention markers of physical activity maintenance in older adults. Am J Health Behav 2015 Jul;39(4):487-499 [FREE Full text] [doi: 10.5993/AJHB.39.4.5] [Medline: 26018097]
66. Sullivan AN, Lachman ME. Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. Front Public Health 2016 Jan 11;4:289 [FREE Full text] [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]
67. US Department of Health and Human Services, Administration for Community Living. Administration for Community Living. Profile of Older Americans URL: https://www.acl.gov/aging-and-disability-in-america/data-and-research/profile-older-americans [accessed 2018-06-08] [WebCite Cache ID 701lwIJfG]
68. Cotten SR, Yost EA, Berkowsky RW, Winstead V, Anderson WA. Designing Technology Training for Older Adults in Continuing Care Retirement Communities. United Kingdom: Routledge; 2016.

Abbreviations

RQ: research question
SDT: self-determination theory
TTM: transtheoretical model

Edited by G Eysenbach; submitted 12.01.18; peer-reviewed by E Chiauzzi, H Similä, M Cabrita; comments to author 25.03.18; revised version received 09.07.18; accepted 13.11.18; published 05.04.19.

Please cite as:
Kononova A, Li L, Kamp K, Bowen M, Rikard RV, Cotten S, Peng W
The Use of Wearable Activity Trackers Among Older Adults: Focus Group Study of Tracker Perceptions, Motivators, and Barriers in the Maintenance Stage of Behavior Change
JMIR Mhealth Uhealth 2019;7(4):e9832
URL: https://mhealth.jmir.org/2019/4/e9832/
doi:10.2196/mhealth.9832
PMID:30950807

Review

Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis

Katie-Jane Brickwood[1], BHM, BHSc (Hons); Greig Watson[1], PhD; Jane O'Brien[1], PhD; Andrew D Williams[1], PhD

School of Health Science, College of Health and Medicine, University of Tasmania, Newnham, Australia

Corresponding Author:
Katie-Jane Brickwood, BHM, BHSc (Hons)
School of Health Science
College of Health and Medicine
University of Tasmania
Newnham Drive
Newnham, 7250
Australia
Phone: 61 0363245487
Email: katiejane.brickwood@utas.edu.au

Abstract

Background: The range of benefits associated with regular physical activity participation is irrefutable. Despite the well-known benefits, physical inactivity remains one of the major contributing factors to ill-health throughout industrialized countries. Traditional lifestyle interventions such as group education or telephone counseling are effective at increasing physical activity participation; however, physical activity levels tend to decline over time. Consumer-based wearable activity trackers that allow users to objectively monitor activity levels are now widely available and may offer an alternative method for assisting individuals to remain physically active.

Objective: This review aimed to determine the effects of interventions utilizing consumer-based wearable activity trackers on physical activity participation and sedentary behavior when compared with interventions that do not utilize activity tracker feedback.

Methods: A systematic review was performed searching the following databases for studies that included the use of a consumer-based wearable activity tracker to improve physical activity participation: Cochrane Controlled Register of Trials, MEDLINE, PubMed, Scopus, Web of Science, Cumulative Index of Nursing and Allied Health Literature, SPORTDiscus, and Health Technology Assessments. Controlled trials of adults comparing the use of a consumer-based wearable activity tracker with other nonactivity tracker–based interventions were included. The main outcome measures were physical activity participation and sedentary behavior. All studies were assessed for risk of bias, and the Grades of Recommendation, Assessment, Development, and Evaluation system was used to rank the quality of evidence. The guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement were followed. A random-effects meta-analysis was completed on the included outcome measures to estimate the treatment effect of interventions that included an activity tracker compared with a control group.

Results: There was a significant increase in daily step count (standardized mean difference [SMD] 0.24; 95% CI 0.16 to 0.33; *P*<.001), moderate and vigorous physical activity (SMD 0.27; 95% CI 0.15 to 0.39; *P*<.001), and energy expenditure (SMD 0.28; 95% CI 0.03 to 0.54; *P*=.03) and a nonsignificant decrease in sedentary behavior (SMD −0.20; 95% CI −0.43 to 0.03; *P*=.08) following the intervention versus control comparator across all studies in the meta-analyses. In general, included studies were at low risk of bias, except for performance bias. Heterogeneity varied across the included meta-analyses ranging from low (I^2=3%) for daily step count through to high (I^2=67%) for sedentary behavior.

Conclusions: Utilizing a consumer-based wearable activity tracker as either the primary component of an intervention or as part of a broader physical activity intervention has the potential to increase physical activity participation. As the effects of physical activity interventions are often short term, the inclusion of a consumer-based wearable activity tracker may provide an effective tool to assist health professionals to provide ongoing monitoring and support.

(JMIR Mhealth Uhealth 2019;7(4):e11819) doi:10.2196/11819

KEYWORDS

exercise; fitness trackers; telemedicine; meta-analysis

Introduction

Physical Activity and Sedentary Behavior

There is significant evidence to support the varied physical and mental health benefits of participation in regular physical activity [1-4]. In addition, regular participation in physical activity plays an important role in maintaining functional independence into aging and decreases the risk of morbidity and mortality [5]. Despite the well-known benefits of physical activity participation, 31% of adults worldwide are insufficiently active as they do not meet the minimum recommendations of at least 30 min of moderate-intensity physical activity on at least 5 days every week, 20 min of vigorous-intensity physical activity on at least 3 days every week, or an equivalent combination achieving 600 metabolic equivalent min per week [6]. In addition, sedentary behavior, which is defined as any waking behavior while in a sitting, reclining, or lying position [7], is independently associated with poor health outcomes, including all-cause and cardiovascular disease mortality [8,9].

Traditional Interventions to Increase Physical Activity Participation

The use of structured lifestyle interventions is reported to be effective in increasing physical activity participation and reducing the progression of chronic diseases [10-12]. In addition, lifestyle interventions have shown to be effective in reducing sedentary behavior [13]. Traditionally, structured lifestyle interventions utilize group or individual education, behavior change techniques, self-monitoring, the provision of written information materials, and/or telephone counseling. Interventions utilizing these methods have shown to be effective at increasing physical activity participation in the short term [12,14]; however, evidence regarding their long-term effectiveness is limited [15-17]. In addition, these types of interventions are often labor and resource intensive [18].

Wearable Activity Trackers

Consumer-based wearable activity trackers are now readily available and can provide individuals with the ability to objectively monitor their physical activity levels. In addition, when combined with the use of smartphone and computer apps, they may assist users through a range of motivational and tracking tools to better manage their personal health [19]. In addition to providing real-time feedback relating to daily steps and energy expenditure, consumer-based wearable activity trackers have the potential to provide specific, tailored feedback through specifically designed algorithms or by health professionals. This type of emerging technology may provide an alternative means of providing ongoing support and motivation to individuals both looking to increase their activity levels or to maintain activity levels following a structured lifestyle intervention [20]. Moreover, consumer-based wearable activity trackers may assist in reducing the resource and time burden associated with traditional methods of providing ongoing support. Randomized controlled trials have shown that these devices have promise in relation to increasing physical activity participation [21,22]; however, participant numbers in individual studies tend to be low, making it difficult to adequately assess the benefits of these devices. Furthermore, there is limited research relating to their long-term adherence and effectiveness. This systematic review and meta-analysis aimed to review the effects of interventions that utilize consumer-based wearable activity trackers compared with a nonactivity tracker–based control group on physical activity participation in adults.

Methods

Search Method

The following Web-based databases were searched using a combination of Medical Subject Headings (MeSH) and free text terms: CENTRAL, MEDLINE, PubMed, Scopus, Web of Science, CINHAL, SPORTDiscus, and Health Technology Assessment. Search strategies were developed relating to the 2 primary concepts of the review: the use of a consumer-based wearable activity tracker and altering physical activity participation. To identify studies that included the use of a consumer-based wearable activity tracker, we used search terms including *Activity Tracker*, *Wearable device*, and *Fitness Tracker* (MeSH). Search terms used to identify studies that focused on altering physical activity participation included *Physical Fitness* (MeSH), *Sedentary Lifestyle* (MeSH), *Step Count*, and *Behaviour Change*. Each database was searched from inception to March 15, 2017, with no language restrictions. Search strategies were adapted for each database as necessary. A full search strategy is available in Multimedia Appendix 1. Reference lists of retrieved articles were checked, and citation searches were performed on key articles. Authors were contacted for additional information where necessary. The search was limited to human studies.

Inclusion and Exclusion Criteria

Published and unpublished controlled trials of adults (aged over 18 years) that utilized a consumer-based wearable activity tracker were included in this review. The effect of consumer-based wearable activity trackers on physical activity participation and sedentary behavior was assessed. For the purpose of this review, consumer-based wearable activity trackers were defined as an electronic device that monitors physical activity and provides automated real-time feedback and may also include interactive behavior change tools via a smartphone or Web-based platform. Consumer-based tracker refers to an activity tracker that is available for purchase to the general public and therefore excludes laboratory-based or research-specific devices. Wearable tracker refers to a device that is easily worn and removed and does not require specialized equipment such as a harness or adhesive dressings.

Studies that included the use of a consumer-based wearable activity tracker as either the basis of the intervention or as a component of a multifaceted intervention were included. Studies that included the use of established behavioral change techniques

such as group or individual counseling or information sessions, financial incentives, or telephone counseling were classified as multifaceted interventions. Interventions that included tools such as regular emails, text messages, online algorithms, or smartphone apps were classified as *wearable-based* interventions.

Studies that compared consumer-based wearable activity trackers with exercise-based interventions (eg, exercise groups), nonexercise interventions (eg, group education programs), and routine (usual) care were included. Control groups that also utilized a consumer-based wearable activity tracker were included; provided feedback from the activity tracker was blinded to the participant. Studies that utilized consumer-based wearable activity trackers for the entirety of the intervention or as a follow-up component to a structured lifestyle intervention were included as were studies examining the effect of consumer-based wearable activity trackers on sedentary behavior.

Review articles, validity, reliability studies, and conference abstracts were excluded. Acceptability and feasibility studies were included provided data relating to physical activity participation was included. The authors of identified ongoing studies were contacted to obtain study progress and request available results for inclusion in the meta-analysis.

Data Extraction

Titles and abstracts were screened in 4 steps: removal of duplicates, by title, by abstract, and by full text. Article titles and abstracts were systematically screened based on the predetermined exclusion criteria (Multimedia Appendix 2). Potentially eligible papers were retrieved by the primary author (KB). All manuscripts identified as requiring full-text review were reviewed independently by 2 authors (KB and AW) according to the exclusion criteria. A third reviewer (GW) resolved any conflicts. The data extraction tool in Covidence (Veritas Health Innovation Ltd, VIC 3000, Australia) online software [23] was used with data extraction performed by both authors (KB and AW) individually and differences resolved by consensus.

Risk of Bias and Quality Assessment

Two reviewers (KB and AW) assessed each study independently for risk of bias using Covidence online software [23] across 7 domains [24]. Each domain was scored as low, unclear, or high risk of bias. Disagreement was resolved by consensus. The following domains were assessed:

- Sequence generation: Was the method used to generate the allocation sequence appropriate to produce comparable groups? The risk of bias was rated as unclear if methods were not accurately described.
- Allocation sequence generation: Was the method used to conceal the allocation sequence appropriate to prevent allocation being known? The risk of bias was rated as unclear if methods were not accurately described.
- Blinding of participants and personnel: Were participants and study personnel blinded to the group allocation? Although this domain was included in the risk of bias assessment, it is important to note that because of the type of studies included, blinding of participants is not feasible, and therefore, all studies were assessed as high risk of bias for this domain.
- Blinding of outcome assessment: Was the outcome measure objective or subjective? If a subjective measure was used, the risk of bias was assessed as high. If an objective measure was used, the risk of bias was assessed as low as objective measures are less likely to be influenced by a lack of blinding.
- Incomplete outcome data: Were incomplete outcome data adequately addressed? Was the analysis an intention-to-treat analysis or were missing data imputed appropriately?
- Selective outcome reporting: Were outcomes prespecified in a study protocol or trial registration and reported as specified?
- Other sources of bias: Were there other sources of bias not previously mentioned, such as author conflicts of interest?

The Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) system was used to rank the quality of evidence for each study using GRADEprofiler Guideline Development Tool online software [25] in conjunction with Chapter 12.2 of the *Cochrane Handbook for Systematic Reviews of Interventions* [24]. The GRADE approach uses 5 considerations (study limitations, consistency of effect, imprecision, indirectness, and publication bias) to assess the quality of evidence for each outcome. The following criteria are used for assigning a grade of evidence:

- High: further research is very unlikely to change our confidence in the estimate of effect.
- Moderate: we are moderately confident the true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different.
- Low: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.
- Very low: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of the effect.

The grade of evidence was downgraded once if:

- More than 25% (n=7) of included studies were at high risk of bias in any criteria (study limitations)
- Heterogeneity was statistically significant and the I^2 value was more than 40% (inconsistency)
- There were differences between included studies in methodological factors such as intervention types and length, included age ranges, and included study populations (indirectness)
- Observed confidence intervals were wide because of small sample sizes (imprecision)
- There was direct evidence of publication bias.

Data Synthesis

Due to the variability of the included studies, random-effects meta-analyses [26] were performed on the following physical activity outcomes using Review Manager (RevMan) [27]: daily step count, minutes per week spent in moderate and vigorous physical activity (MVPA; min/week), energy expenditure

(kcal/week), and minutes per day spent in sedentary behavior (min/day). Data presented as minutes of MVPA per day were multiplied by 7 to calculate minutes of MVPA per week to allow for greater transferability to current physical activity guidelines. Where sedentary behavior data were presented as hours per day, values were divided by 60 to obtain minutes per day. Each meta-analysis compared the results of the intervention group(s) with the study-specific control group. Studies that included multiple intervention groups were entered multiple times, with each intervention group compared against the control group. Studies that included interventions that did not utilize a consumer-based wearable activity tracker were not included in the meta-analysis. Data presented as least-squares mean, SE, or 95% CI were converted to SD using the RevMan calculator. Due to the range of data presentation formats of included studies, all meta-analyses were presented as standardized mean difference (SMD) to accommodate for adjusted and unadjusted means [24]. Mean and SD were requested from authors where data were presented as median and interquartile ranges (IQR) based on the suggestion that estimated mean and SD can be used in a meta-analysis [28]. Authors of studies that presented data in a graphical format were contacted to obtain exact values. Heterogeneity was assessed using I^2 for each meta-analysis.

Where intervention effects were reported using SMD, the reported values were converted into a meaningful value using the pooled SD of studies that reported end point values. Studies that presented mean change data or in which SD was estimated based on IQR were not included in the pooled SD calculation.

Results

Study Selection

The database search was completed during March 2017, with article collection and screening conducted in April to June 2017. A total of 3739 studies were retrieved from the search strategy, with a further 6 studies identified through reference checks. No non-English papers were identified. A total of 1148 duplicates were removed; 2597 studies were screened by title and abstract, with 2484 studies removed as they did not meet the inclusion criteria. The remaining 113 studies were assessed for full-text eligibility, with 89 studies excluded. A total of 28 randomized controlled trials were included in this systematic review [21,22,29-54]. Authors of study protocols were contacted to obtain study progress and results if available. As of December 2017, 3 authors had since published study results [22,36,39] and 1 author [33] had provided unpublished results, which were therefore included in the current systematic review and meta-analysis. Two studies [49,54] were excluded from the meta-analysis. One study [54] was excluded because of all the data being presented in a graphical format. The other study [49] was not included in the meta-analysis as physical activity data were reported in *activity units*. One additional study [38] was excluded from the MVPA meta-analysis because of the graphical representation of data but was not excluded from all meta-analyses as other data were presented in a tabular format. Figure 1 outlines the screening process, including the status of ongoing studies.

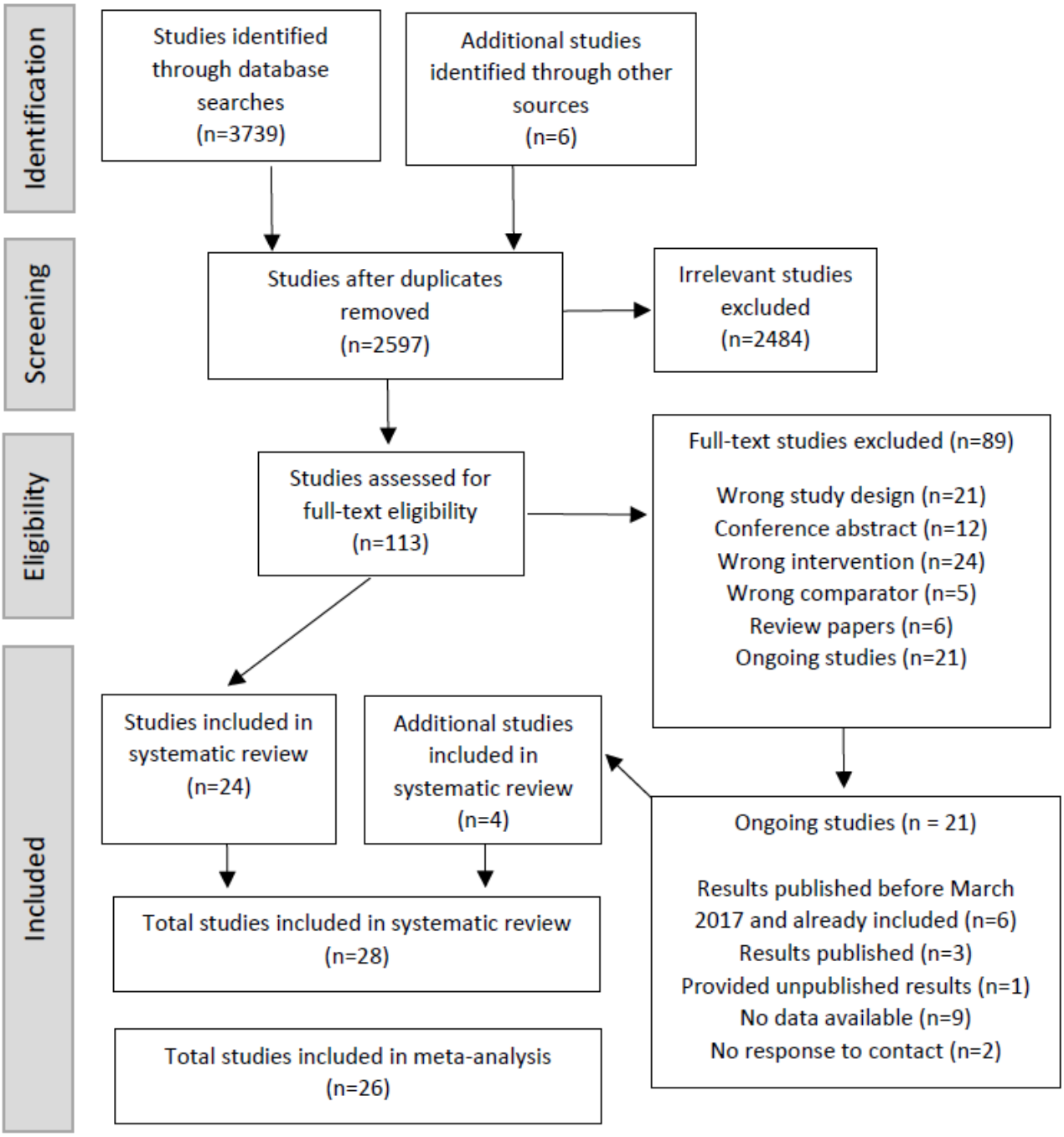
Study Characteristics

A total of 3646 participants across 9 countries were included, with a mean age ranging from 17.9 years to 79.5 years. Included studies were all published between 2007 and 2018. One study was published only as a protocol paper with unpublished results provided by the author [33]. Thirteen studies included young adults (≤39.9 years) [30-34,37,38,42,44,46-48,50], 15 included middle-aged adults (40-64.9 years) [21,22,29,35,36,39-41,43,45,51-54], and 1 study included adults aged over 65 years [49]. Furthermore, 17 studies specified that the participant must have regular access to the internet, a computer, tablet, and/or smartphone [21,30,33,35-37,39,41-47,51-53], with 3 studies requiring participants to be proficient at using the internet and/or smartphones [21,22,45].

Comparator interventions differed across the studies and ranged from maintenance of usual lifestyle [31,44,53], waitlist [30,36,39,40,52], wearing of a consumer-based wearable activity tracker but blinded to feedback [22,38,41,49,50], use of a standard pedometer [21], standard behavioral group-based interventions [29,37,43,45,46,51], telephone counseling [35,37], use of a smartphone app [33,42], and the provision of education materials through mixed media (emails, text message, and written) [21,32,34,47,48,54].

The way in which a consumer-based wearable activity tracker was incorporated into the interventions of included studies ranged from forming the basis of the intervention (wearable-based) [22,31,33,34,38,41,42,44,47,48,54] to being used as a monitoring tool as part of a broader intervention (multifaceted) [21,29,30,32,34-37,39,40,43,45-47,49-54]. Overall, 8 studies included more than 1 intervention group [34,43,45-47,52-54]. In addition, 4 studies included either an unstructured follow-up phase [34,42] or additional intervention phase utilizing a nonblinded activity tracker for all participants [41,50]. Data from these phases were not included in this review as they did not meet the eligibility criteria. Moreover, 17 studies reported activity tracker adherence data [21,22,30,32,34,35,37,38,40,41,43-45,50-52,54], with 13 of these studies [21,22,30,34,35,40,41,43,44,50-52,54] reporting activity tracker wear on at least 50% of the study intervention days. Furthermore, 4 studies [32,37,38,45] reported low activity tracker wear time, with 1 study [32] reporting that all participants had ceased wearing their activity tracker by the end of the intervention. All studies included some form of behavioral change techniques ranging from basic techniques such as the provision of feedback and goal setting to interventions based on the Coventry, Aberdeen and London-Refined taxonomy [55], social cognitive theory [56], and social determination theory [57].

Figure 1. Flowchart of study selection. Studies with results published and authors that provided unpublished results were included in the systematic review.



The initial search resulted in 21 protocol publications being identified [33,58-77]. Of the identified protocol publications, 6 had published results available, which were also identified as part of the initial search strategy [59,71,72,74,75,77]. Results from 3 protocol publications [61,65,68] were published after the initial search and were subsequently included [22,36,39]. One author provided unpublished results [33], 2 protocol authors did not return contact [66,76], and 9 authors were still collecting data or preparing manuscripts [58,60,62-64,67,69,70,73].

Accuracy and Reliability of Included Consumer-Based Wearable Activity Trackers

Included studies utilized a range of consumer-based wearable activity trackers, including various Fitbit models [21,29,33-36,49,50] (Fitbit, San Francisco, CA), Jawbone UP 24 [22,30,40,42] (Jawbone, San Francisco, CA), Gruve [31] (Gruve Solution MUVE, Inc, USA), LumoBack [32] (Lumo BodyTech, Inc, Palo Alto, California, USA), various BodyMedia models [37,43,45,46,51,54] (BodyMedia, Pittsburgh PA), Polar Active [38] (Polar Electro, Finland), Fitbug [41] (Chicago IL), Pebble+ [44] (Fitlinxx Inc), Fitmeter [47] (FitLife, Suwon, Korea), Personal Activity Monitor [48] (PAM BV, Doorwerth, the Netherlands), and Withings Pulse [52] (Cambridge, MA). One study did not specify the brand of consumer-based wearable activity tracker utilized [53]. Fitbit One, Zip and Charge HR, Jawbone Up, LumoBack, and Withings Pulse have all demonstrated excellent test-retest reliability for step count (Intraclass Correlation Coefficient [ICC] >.90) [78-80];

however, a recent review into the use of Fitbit activity trackers suggests that steps are overestimated in free-living conditions [81]. The Polar Active has shown to correlate (r^2=0.74) with the doubly labeled water technique for assessing energy expenditure during military training, which related to the study setting [82]. The Bodymedia Sensewear showed good reliability during outdoor walking (ICC=0.82); however, poor reliability was observed during various treadmill walking speeds (ICC=0.18 to 0.27) [80]. An earlier version of the Pebble+, the ActiPed by FitLinxx has demonstrated a high level of accuracy for step count (−1.30%) and good reliability (ICC=0.85) [83]; however, currently, there are no validity or reliability data for the Pebble+. The PAM has shown similar validity to the ActiGraph accelerometer (r^2=0.95) and good reliability (ICC=0.80) [84]. There is currently limited reliability and validity data relating to the Fitmeter; however, it has been reported that the Fitmeter does correlate with gas analyzer measures for energy expenditure (r^2=0.82) [85]. The Gruve monitor has been shown to be accurate in measuring sedentary and walking activities when compared with a gold standard system (r^2=0.98) [86] and is recommended for use in interventions aiming to reduce sedentary behavior [87]. No validation or reliability data could be found relating to the Fitbug Orb.

Study Outcomes

Physical activity behavior measures included number of steps taken per day [21,22,29,30,32,34,40-42,44,50,54], minutes spent in MVPA per week [21,29,30,33-37,39,48,51,53], and energy expenditure [43,45-47,52]. Sedentary behavior was reported by 8 studies with data presented as hours per day [31,37,38], minutes per day [33,39,40], minutes per 16 hours [32], or percentage of total day spent sitting [29,37]. A summary of outcome measures for all included studies is included in Table 1. A detailed summary of included studies is available in Multimedia Appendix 3.

Risk of Bias

Risk of bias judgments for each included study are presented in Figure 2. One study [53] was assessed as high risk of selection bias because of the randomization of practices rather than individuals. All studies were assessed as high risk of bias for performance bias because of the nature of the intervention and control conditions making blinding impossible. Blinding of outcome assessors (detection bias) was assessed as high risk for 7 studies [31,43,45-48,52] because of the use of subjective outcome measures. One further study [54] was also assessed as a high risk of detection bias as participants were provided with activity level feedback at each assessment with comparisons with previous results. The management of incomplete outcome data was assessed as high risk in 2 studies [29,38]. In addition, 3 studies [36,48,54] were assessed as high risk for selective outcome reporting, and 1 study [22] was judged as high risk for other sources of bias because of conflicts of interest declared by the authors. Publication bias was assessed for daily step count and MVPA with no bias identified. Publication bias was unable to be assessed for other outcome measures because of less than 10 included studies.

Meta-Analysis Results

A total of 26 studies were included across all meta-analyses [21,22,29-48,50-53]. Results were primarily presented as mean and SD or 95% CI, or as mean change and SD, SE, or 95% CI. Two studies [48,52] presented data as median and IQRs, suggesting that the data were not normally distributed. Authors of both the studies were contacted, and they provided mean and SD data, and the studies were subsequently included in the meta-analyses [28]. One study [54] was excluded from the meta-analyses as all data were presented in a graphical format. The authors were contacted for results, but no return contact was received. An additional study was excluded [49] because of data being presented as *activity units*.

Physical Activity Participation

Steps

Overall, 12 studies reported changes in the number of steps taken by participants [21,22,29,30,32,34,40-42,44,50,54]. A random-effects meta-analysis using SMD was performed on 11 studies as 1 study [54] was excluded because of graphical presentation of data. Step data were objectively measured using a range of accelerometers or pedometers. There was a significant increase in step count following the intervention versus control comparator (SMD 0.23; 95% CI 0.15 to 0.32; *P*<.001; Figure 3) across all studies in the meta-analysis, representing an approximate increase of 627 steps (95% CI 417 to 862 steps) per day. Heterogeneity was low [88] and nonsignificant (I^2=3%; *P*=.42). We judged the quality of evidence for consumer-based wearable activity trackers to increase the daily number of steps as low. The quality of the evidence was rated as being low based on being downgraded twice, once because of the high risk of bias identified in the included studies and once because of the level of indirectness associated with the broad range of included interventions, comparators, populations, and settings. The summary of findings table for all outcome measures is available in Multimedia Appendix 4.

Further subgroup analysis was completed, separating the included studies into interventions that were wearable-based [22,34,41,42,44] and those that were multifaceted [21,29,30,32,34,40,50]. A significant increase in daily step count following the intervention versus control comparator was observed in both wearable-based (SMD 0.20; 95% CI 0.08 to 0.33; *P*=.002; Figure 3) and multifaceted (SMD 0.26; 95% CI 0.12 to 0.41; *P*>.001; Figure 3) meta-analyses. This is representative of an approximate increase of 475 steps (95% CI 190 to 784 steps) per day and 685 steps (95% CI 316 to 1080 steps) per day, respectively. Low and nonsignificant heterogeneity was observed in both subgroup analyses (I^2=0%; *P*=.61 and I^2=25%; *P*=.23, respectively).

Table 1. Outcome measures of physical activity participation and sedentary behavior.

Study	Outcome measures	Outcome measurement instrument	Objective/subjective
Ashe et al, 2015 [29]	Steps/day, MVPA[a] (min/day), and sitting time (%)	ActiGraph GT3X accelerometer	Objective
Ashton et al, 2017 [30]	Steps/day and MVPA (min/week)	Yamax Digiwalker SW200 and Godin Leisure-Time Exercise Questionnaire	Objective and subjective
Barwais et al, 2013 [31]	Sitting time (hours/day)	7-day Sedentary and Light Intensity Physical Activity Log	Subjective
Brakeridge et al, 2016 [32]	Steps/day and sitting time (min/16 hours)	ActivPal activity monitor	Objective
Cadmus-Bertram et al, 2015 [21]	Steps/day and MVPA (min/week)	ActiGraph GT3X accelerometer	Objective
Duncan et al, 2016 [33]	MVPA (min/week) and sitting time (min/day)	Geneactiv accelerometer	Objective
Finkelstein et al, 2016 [34]	Steps/day and MVPA (min/week)	ActiGraph GT3X accelerometer	Objective
Hartman et al, 2016 [35]	MVPA (min/day)	ActiGraph GT3X accelerometer	Objective
Hartman et al, 2018 [36]	MVPA (min/day)	ActiGraph GT3X accelerometer	Objective
Jakicic et al, 2016 [37]	MVPA (min/week) and sitting time (hours/day)	Sensewear Pro Armband	Objective
Jauho et al, 2015 [38]	MVPA (min/day)[a] and sitting time (hours/day)	Polar Active (as used in intervention)	Objective
Li et al, 2017 [39]	MVPA ≥3 metabolic equivalents (min/day) and sitting time (min/day)	Sensewear Mini Armband	Objective
Lyons et al, 2017 [40]	Steps/day and sitting time (min/day)	ActivPal activity monitor	Objective
Martin et al, 2015 [41]	Steps/day	Fitbug Orb accelerometer (as used in intervention)	Objective
Melton et al, 2016 [42]	Steps/day	ActiGraph GT3X accelerometer	Objective
Pellegrini et al, 2012 [43]	Energy expenditure (kcal/week)	PPAQ[b]	Subjective
Poirier et al, 2016 [44]	Steps/day	Pebble+ (as used intervention)	Objective
Polzien et al, 2007 [45]	Energy expenditure (kcal/week)	PPAQ	Subjective
Rogers et al, 2016 [46]	Energy expenditure (kcal/week)	PPAQ	Subjective
Shin et al, 2017 [47]	Energy expenditure (kcal/week)	International Physical Activity Questionnaire Short-Form	Subjective
Skrepnik et al, 2017 [22]	Steps/day	Jawbone UP 24 (as used in intervention)	Objective
Slootmaker et al, 2009 [48]	MVPA (min/week)	The Activity Questionnaire for Adolescents and Adults	Subjective
Thompson et al[c], 2014 [49]	Activity units	Research-grade triaxial accelerometer	Objective
Thorndike et al, 2014 [50]	Steps/day	Fitbit (as used in intervention)	Objective
Unick et al, 2012 [51]	MVPA (min/week)	Sensewear Armband	Objective
Valle et al, 2017 [52]	Energy expenditure (kcal/week)	PPAQ	Subjective
van der Weegen et al, 2015 [53]	MVPA (min/day)	Personal Activity Monitor accelerometer	Objective
Van Hoye[c] et al, 2015 [54]	Steps/day	Sensewear Armband (as used in intervention)	Objective

[a]MVPA: moderate and vigorous physical activity.

[b]PPAQ: Paffenbarger Physical Activity Questionnaire.

[c]Study not included in meta-analysis.

Figure 2. Risk of bias summary: review authors' judgments about each risk of bias item for each included study. Green symbols represent a low risk of bias, yellow symbols represent an unclear risk of bias, and red symbols represent a high risk of bias.

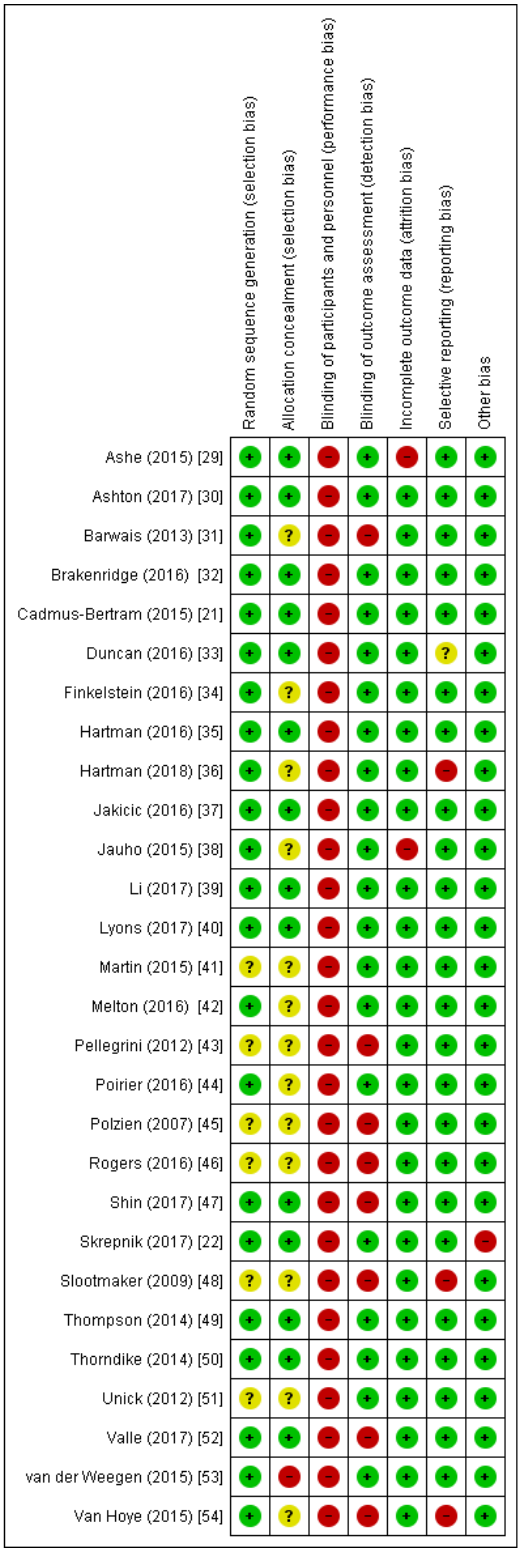
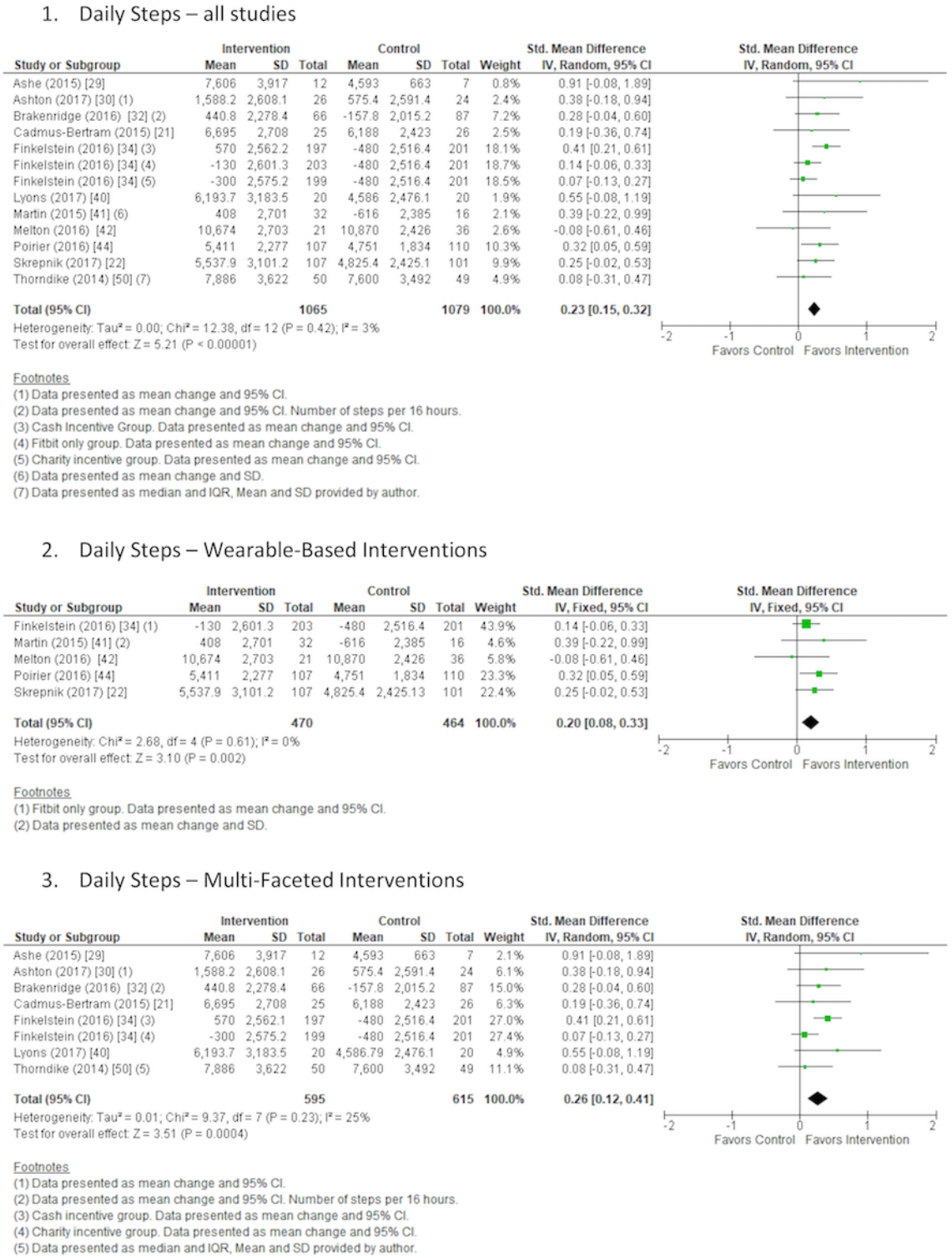


Figure 3. Forest plot of standardized mean difference of steps per day in studies comparing an intervention that included a consumer-based wearable activity tracker with a control group that did not utilize a consumer-based wearable activity tracker. Subgroup analysis was completed on studies that included wearable-based interventions compared with control and multifaceted interventions compared with control. Green square indicates the standardized mean difference for each individual study. Black square indicates the overall standardized mean difference for all studies.



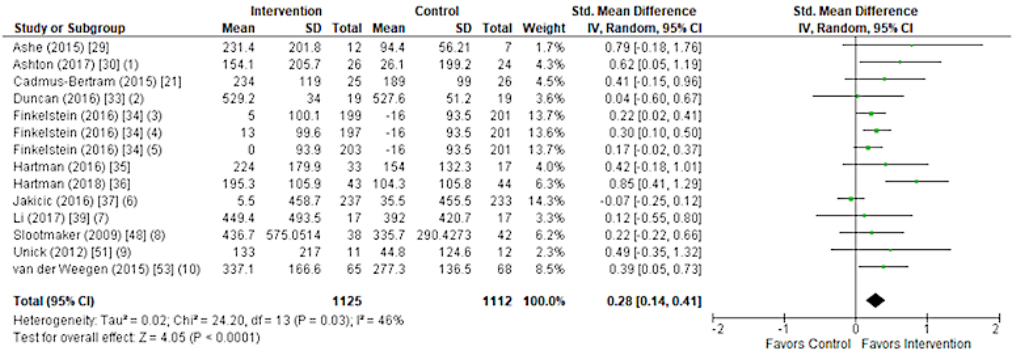
Moderate and Vigorous Physical Activity

A total of 15 studies measured levels of MVPA [21,29-31,33-39,48,51,53,54]. Moreover, 2 studies [38,54] presented data in a graphical format and 1 study [31] presented results for MVPA, separately. A random-effects meta-analysis using SMD was performed on the 12 remaining studies [21,29,30,33-37,39,48,51,53]. Of the included studies, 10 measured MVPA objectively through a range of accelerometers, and 2 studies [30,48] used subjective self-reported measures of MVPA. There was a significant increase in minutes per day spent in MVPA following the intervention versus control comparator (SMD 0.28; 95% CI 0.14 to 0.41; *P*<.001; Figure 4) across all studies in the meta-analysis with moderate and

significant (I^2=46%, *P*=.03) heterogeneity observed. These findings represent an approximate increase of 75 min (95% CI 42 to 109 min) per day of MVPA. The quality of the evidence was rated as very low based on being downgraded 3 times, once because of the high risk of bias identified, once because of the level of inconsistency associated with the observed heterogeneity, and once because of the level of indirectness associated with the broad range of included interventions, comparators, populations, and settings.

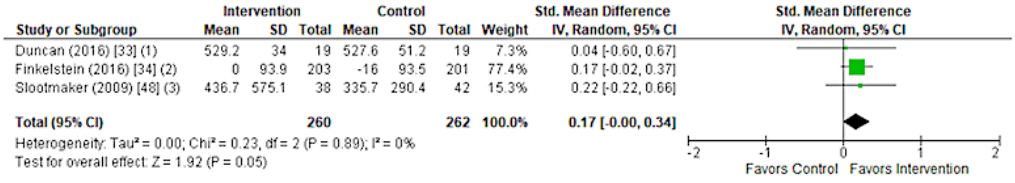
Figure 4. Forest plot of standardized mean difference of time spent in moderate and vigorous physical activity per week in studies comparing an intervention that included a consumer-based wearable activity tracker with a control group that did not utilize a consumer-based wearable activity tracker. Subgroup analysis was completed on studies that included wearable-based interventions compared with control and multifaceted interventions compared with control. Green square indicates the standardized mean difference for each individual study. Black square indicates the overall standardized mean difference for all studies.

1. Minute per week of moderate and vigorous physical activity – all studies



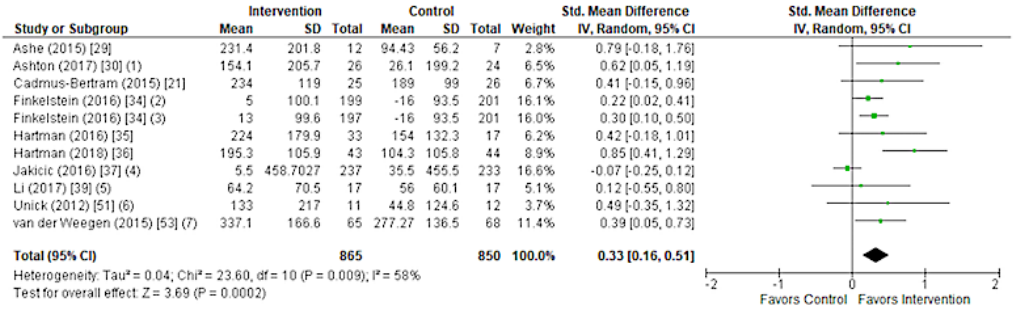
Footnotes
(1) Data presented as mean change and 95% CI. Moderate and vigorous physical activity subjectively measured.
(2) Unpublished data.
(3) Charity incentive group. Data presented as mean change and 95% CI.
(4) Cash incentive group. Data presented as mean change and 95% CI.
(5) Fitbit only group. Data presented as mean change and 95% CI.
(6) Data presented as least squares mean and standard error.
(7) Data presented as bouted moderate and vigorous physical activity greater than 3 metabolic equivalents.
(8) Data presented as median and IQR, Mean and SD provided by author.
(9) Data presented as mean change and SD. Only intervention completers included in analysis.
(10) Self-management support program + tool intervention group.

2. Minutes per week of moderate and vigorous physical activity – Wearable-Based Interventions



Footnotes
(1) Unpublished data.
(2) Fitbit only group. Data presented as mean change and 95% CI
(3) Data presented as median and IQR, Mean and SD provided by author.

3. Minues per week of moderate and vigorous physical activity – Multi-Faceted Interventions



Footnotes
(1) Data presented as mean change and 95% CI.
(2) Charity incentive group. Data presented as mean change and 95% CI.
(3) Cash incentive group. Data presented as mean change and 95% CI.
(4) Data presented as lease squares mean and SE
(5) Data presented as bouted moderate and vigorous physical activity greater than 3 metabolic equivalents.
(6) Data presented as mean change and SD. Only intervention completers included in analysis.
(7) Self-management support program + tool intervention group.

Further subgroup analysis was completed on studies that included wearable-based interventions [33,34,48] and multifaceted interventions [21,29,30,34-37,39,51,53]. There was a nonsignificant increase in minutes per day spent in MVPA following the intervention versus control comparator (SMD 0.17; 95% CI −0.00 to 0.34; *P*=.05; Figure 4) across studies included in the wearable-based meta-analysis, representing an approximate increase of 40 min (95% CI 0 to 80 min) per day of MVPA. Nonsignificant and low heterogeneity was observed (I^2=0%; *P*=.89). A significant increase was observed in minutes per day spent in MVPA following the intervention versus control comparator (SMD 0.33; 95% CI 0.16 to 0.51; *P*<.001; Figure 4) across studies included in the multifaceted meta-analysis. This represents an approximate increase of 92 min (95% CI 45 to 142 min) per day of MVPA. Observed heterogeneity for studies included in the multifaceted intervention was high and significant (I^2=58%; *P*=.009).

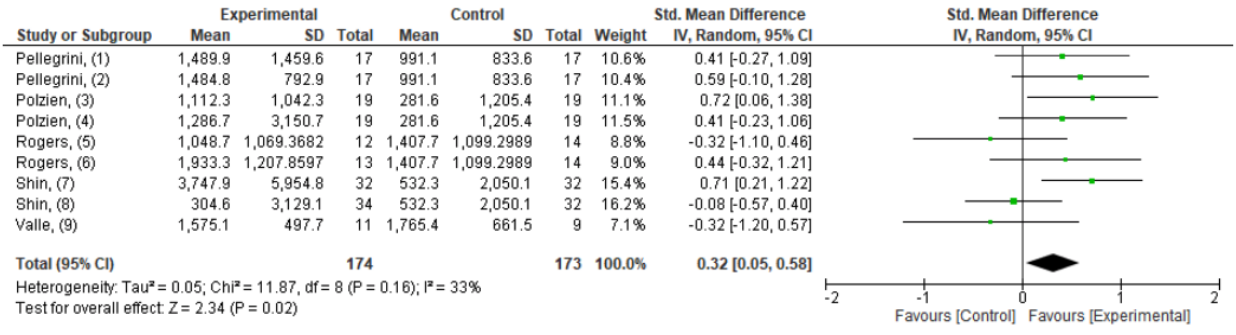
Energy Expenditure

Overall, 5 studies [43,45-47,52] reported physical activity levels in terms of energy expenditure, expressed as kcal per week. The Paffenbarger Physical Activity Questionnaire or the International Physical Activity Questionnaire were utilized in the included studies to obtain self-reported physical activity levels. A random effects meta-analysis using SMD performed on the five included studies showed a significant increase in energy expenditure following the intervention versus control comparator (SMD 0.32; 95%CI 0.05 to 0.58; *P*=.02; Figure 5) across all studies in the meta-analysis. Heterogeneity was low and non-significant (I^2=33%, *P*=.16). These findings represent an approximate increase of 300 kcal (95% CI 32 to 579) in energy expenditure per week. The quality of evidence was rated as being low based on being downgraded twice, once because of the high risk of bias identified and once because of the level of indirectness associated with the broad range of included interventions, comparators, populations, and settings.

Sedentary Behavior

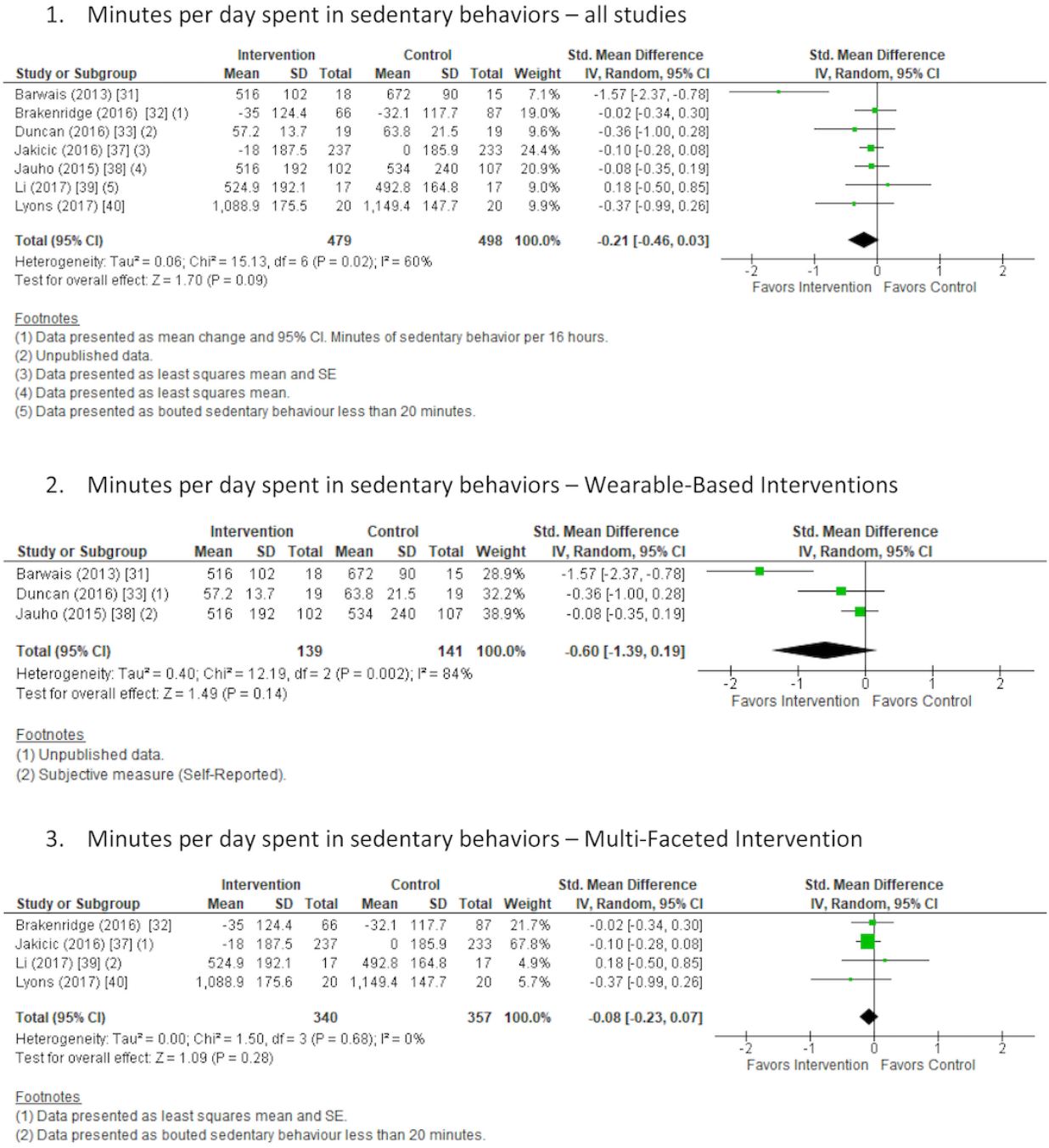
Overall, 8 studies [29,31-33,37-40] reported changes in sedentary behavior. Furthermore, 2 studies [29,37] reported sedentary behavior as percentage of the day spent sitting, and 7 studies [31-33,37-40] reported minutes or hours of sedentary behavior per day. One study [39] reported sedentary behavior as bouts of greater than 20 min, and 1 study [32] reported minutes of sedentary behavior per 16 hours. Sedentary behavior was objectively measured except for 1 study [38], which utilized a self-reported questionnaire to obtain daily sitting time. A random-effects meta-analysis was completed on 7 studies that reported changes in sedentary behavior using SMD. One study [29] that reported sedentary behavior as percentage only was not included in the meta-analysis. For the 1 study [37] that reported sedentary behavior as percentage and hours per day, only hours per day were included in the analysis. There was a nonsignificant decrease in sedentary behavior following the intervention versus control comparator (SMD −0.21; 95% CI -0.46 to 0.03; *P*=.09; Figure 6) across all studies in the meta-analysis with a moderate and significant level of heterogeneity (I^2=60%, *P*=.02). This finding represents approximately 37 min (95% CI −81 to 5 min) less spent in sedentary behavior. The quality of evidence was rated as very low based on being downgraded 3 times, once because of the high risk of bias identified, once because of the level of inconsistency associated with the observed heterogeneity, and once because of the level of indirectness associated with the broad range of included interventions, comparators, populations, and settings.

Figure 5. Forest plot of standardized mean difference of energy expenditure in studies comparing an intervention that included a consumer-based wearable activity tracker with a control group that did not utilize a consumer-based wearable activity tracker. Green square indicates the standardized mean difference for each individual study. Black square indicates the overall standardized mean difference for all studies.



Footnotes
(1) Tech Only Group
(2) SBWL + Tech Group
(3) Con-Tech Group. Data presented as mean change +/- SD.
(4) Int-Tech Group. Data presented as mean change +/- SD.
(5) Tech Group. Data presented as least squares mean & SE.
(6) En-Tech Group. Data presented as least squares mean & SE.
(7) Smartcare + financial incentives group. Data presented as mean change +/- SD.
(8) Smartcare Group. Data presented as mean change +/- SD.
(9) Intervention + Group. Data presented as Mean & IQR. Author provided Mean +/- SD.

Figure 6. Forest plot of standardized mean difference of time spent in sedentary behaviors in studies comparing an intervention that included a consumer-based wearable activity tracker with a control group that did not utilize a consumer-based wearable activity tracker. Subgroup analysis was completed on studies that included wearable-based interventions compared with control and multifaceted interventions compared with control. Green square indicates the standardized mean difference for each individual study. Black square indicates the overall standardized mean difference for all studies.



Further subgroup analysis was completed on studies that included a wearable-based intervention [31,33,38] compared with control and multifaceted interventions [32,37,39,40] compared with control. There was a nonsignificant decrease in sedentary behavior following both wearable-based and multifaceted interventions versus control comparator (SMD −0.60; 95% CI −1.40 to 0.19; *P*=.14 and SMD −0.08; 95% CI −0.23 to 0.07; *P*=.28, respectively; Figure 6) across the meta-analyses. These findings represent a decrease of approximately 115 min (95% CI −269 to 36 min) per day spent in sedentary behaviors for wearable-based interventions and a decrease of 13 min (95% CI −39 to 12 min) spent in sedentary behaviors for multifaceted interventions. High and significant heterogeneity (I^2=84%; *P*=.002) was observed for the wearable-based analysis and low and nonsignificant (I^2=0%; *P*=.68) for the multifaceted analysis.

Discussion

Principal Findings

This systematic review and meta-analysis summarizes the results of interventions that utilize a consumer-based wearable activity tracker to assist in the improvement of physical activity participation [21,22,29-54]. The results show a significant improvement in all measures of physical activity participation when compared with control groups, even when interventions were separated into wearable-based and multifaceted. However, intervention groups that were multifaceted in nature appeared to have a greater effect on physical activity participation when compared with control groups than those that included just the use of a consumer-based wearable activity tracker compared with control groups. No real differences in sedentary behavior were observed for either wearable-based or multifaceted interventions compared with control groups.

Physical Activity Participation

Participants who received an intervention including a consumer-based wearable activity tracker demonstrated a significant improvement in daily steps, MVPA, and energy expenditure when compared with control groups. The quality of evidence was low for daily steps and energy expenditure and very low for MVPA. MVPA was the only physical activity outcome measure with a significant level of heterogeneity. This is most likely because of the inclusion of a greater range of intervention types when compared with other outcome measures and the method by which MVPA was measured, including 2 studies [30,48] that subjectively measured MVPA and another study [39] that presented MVPA data in bouts. Despite the low level of certainty, it is encouraging to see a significant positive intervention effect across all included measures of physical activity participation. A recent review [89] into the use of an electronic activity monitor system as an intervention modality also concluded that activity monitors have the potential to increase physical activity levels as did a review into the inclusion of an activity monitor in addition to a behavioral physical activity intervention for overweight and obese adults [90]. A small-to-moderate effect on physical activity participation was also reported by a review that looked at the effect of wearables and smartphone apps as an intervention modality [91].

An overall positive effect for consumer-based wearable activity trackers as an intervention tool was observed, even when interventions were separated into wearable-based only and multifaceted interventions. Subgroup analysis determined a larger effect size for interventions that were multifaceted (vs control) than for wearable-based (vs control) interventions. This suggests that consumer-based wearable activity trackers can be effective on their own, but when combined with other behavior change techniques, such as telephone counseling or group-based education, the improvement in physical activity participation is greater. However, the magnitude of the additive effect of each individual component of multifaceted interventions cannot be determined by meta-analysis. Only 3 of the included studies [34,47,54] directly compared wearable-based and multifaceted interventions. Two of these studies [34,47] offered financial incentives, with both studies reporting a significant improvement in daily step count for those receiving financial incentives compared with either control groups or those not receiving incentives. Financial incentives have previously been shown to be effective at increasing exercise adherence [92]. The third study [54] reported that adding personalized coaching to the use of a consumer-based wearable activity tracker resulted in sustained increases in physical activity behaviors over the 4-week intervention period. For MVPA, the observed heterogeneity was low and nonsignificant for the wearable-based analysis but high and significant for the multifaceted analysis. This indicates that the different types of included interventions contribute to the observed heterogeneity.

Two of the included studies [34,44] reported a significant increase in daily steps, 4 studies [30,34,36,53] reported a significant increase in MVPA, and 2 studies [45,47] reported a significant increase in energy expenditure for the intervention group compared with control. Most of the remaining studies reported a nonsignificant increase in physical activity participation. One study reported a reduction in daily steps [42], 1 study reported a reduction in MVPA participation [37], and 3 studies reported a reduction in energy expenditure [46,47,52] in at least one of the included intervention groups. Potential reasons why other included studies did not find a significant intervention effect or observed a reduction in physical activity participation may be because of several of the studies being pilot studies [29,30,40,47,52], which often have insufficient power because of small sample sizes [93]; a moderate-to-high loss to follow-up in some studies [29,32,33,37,42,43,46,48]; the type of intervention provided to the control group; the actual wear time of the activity tracker; the length of the interventions; and the use of subjective outcome measures. In addition, 8 studies [29,32,33,37,42,43,46,48] that reported physical activity outcome measures had greater than 20% loss to follow-up, with 3 of these studies reporting a reduction in physical activity participation. One study [32] reported a 56% loss to follow-up in the intervention group and 65% in the control group, and another study [43] had a large disparity in loss to follow-up between groups, with 47% loss in the control group compared with 11.7% loss in the intervention group. Intention-to-treat analysis was used in the statistical analysis of these studies, which is a cautious approach and minimizes the influence of loss to follow-up. Despite this, this style of analysis is less likely to show a positive treatment effect [94]. A variety of comparator interventions were utilized within the included studies, ranging from no contact in the form of usual care or waitlist comparators to participants being provided with behavioral group-based interventions or telephone counseling. Previously, both group behavioral interventions and telephone counseling have been demonstrated to be effective in increasing physical activity participation [14,95,96]. Actual wear time of the activity tracker varied, ranging from over 90% wear time [21,22] to all participants ceasing to wear the device by the end of the intervention [32] as the study design allowed self-directed wear of the activity tracker. The 2 studies that were 12 months or longer [32,37] reported lower adherence rates compared with shorter duration studies. Issues with long-term adherence to lifestyle and behavioral change interventions are well recognized [16,17,97]. Subjective, self-reported questionnaires were used

to obtain participants' daily energy expenditure in all included studies and MVPA in 2 included studies. Although the use of these types of questionnaires is widely accepted, self-reported questionnaires have been shown to be less robust in measuring energy expenditure and MVPA when compared with objective measures [98,99].

Sedentary Behavior

Less time was spent in sedentary behavior for those receiving a consumer-based wearable activity tracker intervention compared with control groups; however, the finding was not significant, and the quality of the evidence was graded as very low [29,31-33,37-40]. Similar to the studies included in the physical activity participation analyses, a high risk of bias and serious level of indirectness was identified in the included studies. In addition, because of the observed level of heterogeneity, the quality of evidence was further downgraded. The included studies reported sedentary behavior in a range of formats, including percentage, minutes per day and per 16 hours, bouts of sitting time, and objective and subjective measurements, which could contribute to the observed level of heterogeneity. Only 1 included study [32] specifically aimed to reduce sitting time. One other study [40] utilized the idle alert feature of the consumer-based wearable activity tracker. Although a reduction of sitting time is often observed in interventions targeting physical activity promotion, interventions that specifically target sedentary behaviors are more effective [100-102]. The lack of specific focus on reducing sedentary behavior in 6 of the 8 included studies may have contributed to the nonsignificant finding. In contrast to the reported findings relating to physical activity participation, the subgroup analysis showed that the wearable-based interventions had a greater effect on reducing sedentary behaviors compared with control groups than the multifaceted interventions. This finding may be because none of the interventions included in the wearable-based meta-analysis exceeded 12 weeks in duration, whereas the multifaceted analysis included 2 studies that were 12 months or more in duration. Interestingly, the wearable-based interventions had a high and significant level of heterogeneity, whereas the heterogeneity was nonsignificant and low for the multifaceted interventions. This result was potentially because of the inclusion of 1 study [31] in the wearable-based meta-analysis that reported a 21% reduction in sitting time.

Strengths and Limitations of Review

The current analysis incorporates a wide range of participant populations, ranging from younger to older adults as well as individuals that are apparently healthy to individuals with diagnosed chronic conditions. Previous reviews have focused on specific chronic condition populations [103] and patient subgroups such as overweight and obese adults [90]. A thorough systematic methodology was followed, and the inclusion of a meta-analysis allows for interpretation of the combined effects of including a consumer-based wearable activity tracker as part of a physical activity intervention. A wide range of physical activity interventions that differ in the way in which a consumer-based wearable activity tracker was utilized were included in the review. The benefits of using a consumer-based wearable activity tracker in addition to behavioral interventions have previously been demonstrated [90], meaning this review adds further support for the use of consumer-based wearable activity trackers in a range of different settings. In addition, the use of a consumer-based wearable activity tracker as a stand-alone intervention was examined, with the results indicating that even without supporting behavior change techniques, the use of a consumer-based wearable activity tracker could be effective in increasing physical activity participation. This may have clinical relevance as increased physical activity participation may lead to improvements in overall health.

Although the inclusion of a wide range of interventions and study populations had advantages in terms of general applications, the heterogeneity of the included study designs makes the comparison and synthesis of results difficult and lowers the overall quality of the evidence. The interventions used as comparators also differ greatly between included studies, once again making the comparison of results difficult.

Conclusions and Practical Implications

Utilizing a consumer-based wearable activity tracker either as the primary component of an intervention or as part of a broader physical activity intervention has the potential to increase physical activity participation. Although the quality of evidence is low to very low, the included studies encompass a large age range and include males and females and a range of healthy and chronic condition populations. Although findings were not significant in all studies, short-term interventions utilizing a consumer-based wearable activity tracker generally resulted in increased physical activity participation. This suggests that consumer-based wearable activity trackers may be complementary to traditional intervention modalities such as group-based education and telephone counseling. The effects of physical activity interventions are generally short term, with ongoing contact from health professionals increasing long-term adherence to physical activity participation. Therefore, consumer-based wearable activity trackers have the potential to be included as an effective tool to assist health professionals to provide ongoing monitoring and support to patients with minimal resource expenditure. Further research to determine the effect of consumer-based wearable activity tracker independent of other traditional physical activity interventions would be beneficial as would investigations of the cost-effectiveness of consumer-based wearable activity tracker interventions. Given the potential novelty factor associated with the use of consumer-based wearable activity trackers, further investigation into their long-term usage and effectiveness would be useful to guide potential clinical applications and future recommendations.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Full search strategy.

[PDF File (Adobe PDF File), 197KB - mhealth_v7i4e11819_app1.pdf]

Multimedia Appendix 2

Exclusion criteria.

[PDF File (Adobe PDF File), 180KB - mhealth_v7i4e11819_app2.pdf]

Multimedia Appendix 3

Detailed summary of study design and baseline characteristics for all included studies.

[PDF File (Adobe PDF File), 205KB - mhealth_v7i4e11819_app3.pdf]

Multimedia Appendix 4

GradePRO summary of findings table.

[PDF File (Adobe PDF File), 29KB - mhealth_v7i4e11819_app4.pdf]

References

1. Hupin D, Roche F, Gremeaux V, Chatard J, Oriol M, Gaspoz J, et al. Even a low-dose of moderate-to-vigorous physical activity reduces mortality by 22% in adults aged ≥60 years: a systematic review and meta-analysis. Br J Sports Med 2015 Oct;49(19):1262-1267. [doi: 10.1136/bjsports-2014-094306] [Medline: 26238869]
2. Echouffo-Tcheugui JB, Butler J, Yancy CW, Fonarow GC. Association of physical activity or fitness with incident heart failure: a systematic review and meta-analysis. Circ Heart Fail 2015 Sep;8(5):853-861 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.115.002070] [Medline: 26175539]
3. Liu L, Shi Y, Li T, Qin Q, Yin J, Pang S, et al. Leisure time physical activity and cancer risk: evaluation of the WHO's recommendation based on 126 high-quality epidemiological studies. Br J Sports Med 2016 Mar;50(6):372-378. [doi: 10.1136/bjsports-2015-094728] [Medline: 26500336]
4. Rosenbaum S, Tiedemann A, Sherrington C, Curtis J, Ward PB. Physical activity interventions for people with mental illness: a systematic review and meta-analysis. J Clin Psychiatry 2014 Sep;75(9):964-974. [doi: 10.4088/JCP.13r08765] [Medline: 24813261]
5. Paterson DH, Warburton DE. Physical activity and functional limitations in older adults: a systematic review related to Canada's Physical Activity Guidelines. Int J Behav Nutr Phys Act 2010 May 11;7:38 [FREE Full text] [doi: 10.1186/1479-5868-7-38] [Medline: 20459782]
6. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
7. 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]
8. 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]
9. Proper KI, Singh AS, van Mechelen W, Chinapaw MJ. Sedentary behaviors and health outcomes among adults: a systematic review of prospective studies. Am J Prev Med 2011 Feb;40(2):174-182. [doi: 10.1016/j.amepre.2010.10.015] [Medline: 21238866]
10. Bock C, Jarczok MN, Litaker D. Community-based efforts to promote physical activity: a systematic review of interventions considering mode of delivery, study quality and population subgroups. J Sci Med Sport 2014 May;17(3):276-282. [doi: 10.1016/j.jsams.2013.04.009] [Medline: 23693030]
11. Patnode CD, Evans CV, Senger CA, Redmond N, Lin JS. Behavioral counseling to promote a healthful diet and physical activity for cardiovascular disease prevention in adults without known cardiovascular disease risk factors: updated evidence report and systematic review for the US Preventive Services Task Force. J Am Med Assoc 2017 Jul 11;318(2):175-193. [doi: 10.1001/jama.2017.3303] [Medline: 28697259]
12. Cradock KA, ÓLaighin G, Finucane FM, Gainforth HL, Quinlan LR, Ginis KA. Behaviour change techniques targeting both diet and physical activity in type 2 diabetes: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017 Feb 8;14(1):18 [FREE Full text] [doi: 10.1186/s12966-016-0436-0] [Medline: 28178985]

13. Howlett N, Trivedi D, Troop NA, Chater AM. Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med 2018 Feb 28;9(1):147-157. [doi: 10.1093/tbm/iby010] [Medline: 29506209]
14. Goode AD, Reeves MM, Eakin EG. Telephone-delivered interventions for physical activity and dietary behavior change: an updated systematic review. Am J Prev Med 2012 Jan;42(1):81-88. [doi: 10.1016/j.amepre.2011.08.025] [Medline: 22176852]
15. Hillsdon M, Foster C, Thorogood M. Interventions for promoting physical activity. Cochrane Database Syst Rev 2005(1):CD003180 [FREE Full text] [doi: 10.1002/14651858.CD003180.pub2] [Medline: 15674903]
16. O'Halloran PD, Blackstock F, Shields N, Holland A, Iles R, Kingsley M, et al. Motivational interviewing to increase physical activity in people with chronic health conditions: a systematic review and meta-analysis. Clin Rehabil 2014 Dec;28(12):1159-1171. [doi: 10.1177/0269215514536210] [Medline: 24942478]
17. van der Bij AK, Laurant MG, Wensing M. Effectiveness of physical activity interventions for older adults: a review. Am J Prev Med 2002 Feb;22(2):120-133. [Medline: 11818183]
18. Coughlin SS, Stewart J. Use of consumer wearable devices to promote physical activity: a review of health intervention studies. J Environ Health Sci 2016 Nov;2(6) [FREE Full text] [Medline: 28428979]
19. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
20. Preusse KC, Mitzner TL, Fausset CB, Rogers WA. Older adults' acceptance of activity trackers. J Appl Gerontol 2017 Feb;36(2):127-155 [FREE Full text] [doi: 10.1177/0733464815624151] [Medline: 26753803]
21. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a Fitbit-based physical activity intervention for women. Am J Prev Med 2015 Sep;49(3):414-418. [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]
22. Skrepnik N, Spitzer A, Altman R, Hoekstra J, Stewart J, Toselli R. Assessing the impact of a novel smartphone application compared with standard follow-up on mobility of patients with knee osteoarthritis following treatment with Hylan G-F 20: a randomized controlled trial. JMIR Mhealth Uhealth 2017 May 9;5(5):e64 [FREE Full text] [doi: 10.2196/mhealth.7179] [Medline: 28487266]
23. Covidence. URL: https://www.covidence.org/home [accessed 2018-07-29] [WebCite Cache ID 71GTSit9s]
24. Higgins JT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. London, United Kingdom: The Cochrane Collaboration; 2011.
25. GRADEpro. URL: https://gradepro.org/ [accessed 2018-07-29] [WebCite Cache ID 71GTlWHkd]
26. Hunter JE, Schmidt FL. Fixed effects vs random effects meta-analysis models: implications for cumulative reesearch knowledg. Int J Select Assess 2000 Dec;8(4):275-292. [doi: 10.1111/1468-2389.00156]
27. Cochrane Community. RevMan 5 URL: https://community.cochrane.org/help/tools-and-software/revman-5/revman-5-download [accessed 2019-03-05] [WebCite Cache ID 76euXSQhL]
28. Fu R, Vandermeer BW, Shamliyan TA, O'Neil ME, Yazdi F, Fox SH, et al. Handling continuous outcomes in quantitative synthesis. Methods guide for comparative effectiveness reviews. Rockville, MD: Agency for Healthcare Research and Quality; 2013.
29. Ashe MC, Winters M, Hoppmann CA, Dawes MG, Gardiner PA, Giangregorio LM, et al. "Not just another walking program": Everyday Activity Supports You (EASY) model-a randomized pilot study for a parallel randomized controlled trial. Pilot Feasibility Stud 2015;1:4 [FREE Full text] [doi: 10.1186/2055-5784-1-4] [Medline: 27175291]
30. Ashton LM, Morgan PJ, Hutchesson MJ, Rollo ME, Collins CE. Feasibility and preliminary efficacy of the 'HEYMAN' healthy lifestyle program for young men: a pilot randomised controlled trial. Nutr J 2017 Jan 13;16(1):2 [FREE Full text] [doi: 10.1186/s12937-017-0227-8] [Medline: 28086890]
31. Barwais FA, Cuddihy TF, Tomson LM. Physical activity, sedentary behavior and total wellness changes among sedentary adults: a 4-week randomized controlled trial. Health Qual Life Outcomes 2013;11:183 [FREE Full text] [doi: 10.1186/1477-7525-11-183] [Medline: 24168638]
32. Brakenridge CL, Fjeldsoe BS, Young DC, Winkler EA, Dunstan DW, Straker LM, et al. Evaluating the effectiveness of organisational-level strategies with or without an activity tracker to reduce office workers' sitting time: a cluster-randomised trial. Int J Behav Nutr Phys Act 2016 Nov 4;13(1):115 [FREE Full text] [doi: 10.1186/s12966-016-0441-3] [Medline: 27814738]
33. Duncan MJ, Vandelanotte C, Trost SG, Rebar AL, Rogers N, Burton NW, et al. Balanced: a randomised trial examining the efficacy of two self-monitoring methods for an app-based multi-behaviour intervention to improve physical activity, sitting and sleep in adults. BMC Public Health 2016 Dec 30;16:670 [FREE Full text] [doi: 10.1186/s12889-016-3256-x] [Medline: 27473327]
34. Finkelstein EA, Haaland BA, Bilger M, Sahasranaman A, Sloan RA, Nang EE, et al. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. Lancet Diabetes Endocrinol 2016 Dec;4(12):983-995. [doi: 10.1016/S2213-8587(16)30284-4] [Medline: 27717766]

35. Hartman SJ, Nelson SH, Cadmus-Bertram LA, Patterson RE, Parker BA, Pierce JP. Technology-and phone-based weight loss intervention: pilot RCT in women at elevated breast cancer risk. Am J Prev Med 2016 Nov;51(5):714-721. [doi: 10.1016/j.amepre.2016.06.024] [Medline: 27593420]
36. Hartman SJ, Nelson SH, Myers E, Natarajan L, Sears DD, Palmer BW, et al. Randomized controlled trial of increasing physical activity on objectively measured and self-reported cognitive functioning among breast cancer survivors: the memory & motion study. Cancer 2018 Jan 1;124(1):192-202. [doi: 10.1002/cncr.30987] [Medline: 28926676]
37. Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, et al. Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. J Am Med Assoc 2016 Sep 20;316(11):1161-1171. [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
38. Jauho A, Pyky R, Ahola R, Kangas M, Virtanen P, Korpelainen R, et al. Effect of wrist-worn activity monitor feedback on physical activity behavior: a randomized controlled trial in Finnish young men. Prev Med Rep 2015;2:628-634 [FREE Full text] [doi: 10.1016/j.pmedr.2015.07.005] [Medline: 26844128]
39. Li LC, Sayre EC, Xie H, Clayton C, Feehan LM. A community-based physical activity counselling program for people with knee osteoarthritis: feasibility and preliminary efficacy of the track-OA study. JMIR Mhealth Uhealth 2017 Jun 26;5(6):e86 [FREE Full text] [doi: 10.2196/mhealth.7863] [Medline: 28652228]
40. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: a randomized controlled pilot trial. JMIR Mhealth Uhealth 2017 Mar 6;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]
41. Martin SS, Feldman DI, Blumenthal RS, Jones SR, Post WS, McKibben RA, et al. mActive: a randomized clinical trial of an automated mHealth intervention for physical activity promotion. J Am Heart Assoc 2015 Nov 9;4(11) [FREE Full text] [doi: 10.1161/JAHA.115.002239] [Medline: 26553211]
42. Melton BF, Buman MP, Vogel RL, Harris BS, Bigham LE. Wearable devices to improve physical activity and sleep. J Black Stud 2016 Jul 27;47(6):610-625. [doi: 10.1177/0021934716653349]
43. Pellegrini CA, Verba SD, Otto AD, Helsel DL, Davis KK, Jakicic JM. The comparison of a technology-based system and an in-person behavioral weight loss intervention. Obesity (Silver Spring) 2012 Feb;20(2):356-363 [FREE Full text] [doi: 10.1038/oby.2011.13] [Medline: 21311506]
44. Poirier J, Bennett WL, Jerome GJ, Shah NG, Lazo M, Yeh H, et al. Effectiveness of an activity tracker-and internet-based adaptive walking program for adults: a randomized controlled trial. J Med Internet Res 2016;18(2):e34 [FREE Full text] [doi: 10.2196/jmir.5295] [Medline: 26860434]
45. Polzien KM, Jakicic JM, Tate DF, Otto AD. The efficacy of a technology-based system in a short-term behavioral weight loss intervention. Obesity (Silver Spring) 2007 Apr;15(4):825-830. [doi: 10.1038/oby.2007.584] [Medline: 17426316]
46. Rogers RJ, Lang W, Gibbs BB, Davis KK, Burke LE, Kovacs SJ, et al. Applying a technology-based system for weight loss in adults with obesity. Obes Sci Pract 2016 Mar;2(1):3-12 [FREE Full text] [doi: 10.1002/osp4.18] [Medline: 27812375]
47. Shin DW, Yun JM, Shin J, Kwon H, Min HY, Joh H, et al. Enhancing physical activity and reducing obesity through smartcare and financial incentives: a pilot randomized trial. Obesity (Silver Spring) 2017 Feb;25(2):302-310. [doi: 10.1002/oby.21731] [Medline: 28063226]
48. Slootmaker SM, Chinapaw MJ, Schuit AJ, Seidell JC, Van Mechelen W. Feasibility and effectiveness of online physical activity advice based on a personal activity monitor: randomized controlled trial. J Med Internet Res 2009;11(3):e27 [FREE Full text] [doi: 10.2196/jmir.1139] [Medline: 19674956]
49. Thompson WG, Kuhle CL, Koepp GA, McCrady-Spitzer SK, Levine JA. "Go4Life" exercise counseling, accelerometer feedback, and activity levels in older people. Arch Gerontol Geriatr 2014;58(3):314-319. [doi: 10.1016/j.archger.2014.01.004] [Medline: 24485546]
50. Thorndike AN, Mills S, Sonnenberg L, Palakshappa D, Gao T, Pau CT, et al. Activity monitor intervention to promote physical activity of physicians-in-training: randomized controlled trial. PLoS One 2014;9(6):e100251 [FREE Full text] [doi: 10.1371/journal.pone.0100251] [Medline: 24950218]
51. Unick JL, O'Leary KC, Bond DS, Wing RR. Physical activity enhancement to a behavioral weight loss program for severely obese individuals: a preliminary investigation. ISRN Obes 2012 Sep 5;2012 [FREE Full text] [doi: 10.5402/2012/465158] [Medline: 24379985]
52. Valle CG, Deal AM, Tate DF. Preventing weight gain in African American breast cancer survivors using smart scales and activity trackers: a randomized controlled pilot study. J Cancer Surviv 2017 Feb;11(1):133-148. [doi: 10.1007/s11764-016-0571-2] [Medline: 27631874]
53. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Wittel L. It's LiFe! Mobile and web-based monitoring and feedback tool embedded in primary care increases physical activity: a cluster randomized controlled trial. J Med Internet Res 2015 Jul 24;17(7):e184 [FREE Full text] [doi: 10.2196/jmir.4579] [Medline: 26209025]
54. Van Hoye K, Boen F, Lefevre J. The impact of different degrees of feedback on physical activity levels: a 4-Week intervention study. Int J Environ Res Public Health 2015 Jun;12(6):6561-6581 [FREE Full text] [doi: 10.3390/ijerph120606561] [Medline: 26067990]

55. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
56. Bandura A. Social Cognitive Theory, in The International Encyclopedia of Communication. Int Encycl Commun 2008. [doi: 10.1002/9781405186407.wbiecs053]
57. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68] [Medline: 11392867]
58. Amorim AB, Pappas E, Simic M, Ferreira ML, Tiedemann A, Jennings M, et al. Integrating mobile health and physical activity to reduce the burden of Chronic low back pain Trial (IMPACT): a pilot trial protocol. BMC Musculoskelet Disord 2016;17:36 [FREE Full text] [doi: 10.1186/s12891-015-0852-3] [Medline: 26787469]
59. Barry VW, McClain AC, Shuger S, Sui X, Hardin JW, Hand GA, et al. Using a technology-based intervention to promote weight loss in sedentary overweight or obese adults: a randomized controlled trial study design. Diabetes Metab Syndr Obes 2011 Feb 4;4:67-77 [FREE Full text] [doi: 10.2147/DMSO.S14526] [Medline: 21448324]
60. Brickwood K, Smith ST, Watson G, Williams AD. The effect of ongoing feedback on physical activity levels following an exercise intervention in older adults: a randomised controlled trial protocol. BMC Sports Sci Med Rehabil 2017;9:1 [FREE Full text] [doi: 10.1186/s13102-016-0066-5] [Medline: 28078090]
61. Clayton C, Feehan L, Goldsmith CH, Miller WC, Grewal N, Ye J, et al. Feasibility and preliminary efficacy of a physical activity counseling intervention using Fitbit in people with knee osteoarthritis: the TRACK-OA study protocol. Pilot Feasibility Stud 2015;1:30 [FREE Full text] [doi: 10.1186/s40814-015-0027-x] [Medline: 27965808]
62. Coffeng JK, van der Ploeg HP, Castellano JM, Fernández-Alvira JM, Ibáñez B, García-Lunar I, et al. A 30-month worksite-based lifestyle program to promote cardiovascular health in middle-aged bank employees: design of the TANSNIP-PESA randomized controlled trial. Am Heart J 2017 Feb;184:121-132 [FREE Full text] [doi: 10.1016/j.ahj.2016.11.002] [Medline: 28224926]
63. Gnagnarella P, Dragà D, Baggi F, Simoncini MC, Sabbatini A, Mazzocco K, et al. Promoting weight loss through diet and exercise in overweight or obese breast cancer survivors (InForma): study protocol for a randomized controlled trial. Trials 2016 Dec 28;17:363 [FREE Full text] [doi: 10.1186/s13063-016-1487-x] [Medline: 27464488]
64. Granado-Font E, Flores-Mateo G, Sorlí-Aguilar M, Montaña-Carreras X, Ferre-Grau C, Barrera-Uriarte M, et al. Effectiveness of a Smartphone application and wearable device for weight loss in overweight or obese primary care patients: protocol for a randomised controlled trial. BMC Public Health 2015 Jun 4;15:531 [FREE Full text] [doi: 10.1186/s12889-015-1845-8] [Medline: 26041131]
65. Hartman SJ, Natarajan L, Palmer BW, Parker B, Patterson RE, Sears DD. Impact of increasing physical activity on cognitive functioning in breast cancer survivors: rationale and study design of memory & motion. Contemp Clin Trials 2015 Nov;45(Pt B):371-376 [FREE Full text] [doi: 10.1016/j.cct.2015.09.021] [Medline: 26427563]
66. Hasselmann V, Oesch P, Fernandez-Luque L, Bachmann S. Are exergames promoting mobility an attractive alternative to conventional self-regulated exercises for elderly people in a rehabilitation setting? Study protocol of a randomized controlled trial. BMC Geriatr 2015 Sep 7;15:108 [FREE Full text] [doi: 10.1186/s12877-015-0106-0] [Medline: 26346751]
67. Hassett L, van den Berg M, Lindley RI, Crotty M, McCluskey A, van der Ploeg HP, et al. Effect of affordable technology on physical activity levels and mobility outcomes in rehabilitation: a protocol for the Activity and MObility UsiNg Technology (AMOUNT) rehabilitation trial. BMJ Open 2016 Dec 6;6(6):e012074 [FREE Full text] [doi: 10.1136/bmjopen-2016-012074] [Medline: 27266776]
68. Jones D, Skrepnik N, Toselli RM, Leroy B. Incorporating novel mobile health technologies into management of knee osteoarthritis in patients treated with intra-articular hyaluronic acid: rationale and protocol of a randomized controlled trial. JMIR Res Protoc 2016 Aug 9;5(3):e164 [FREE Full text] [doi: 10.2196/resprot.5940] [Medline: 27506148]
69. Norton MC, Clark CJ, Tschanz JT, Hartin P, Fauth EB, Gast JA, et al. The design and progress of a multidomain lifestyle intervention to improve brain health in middle-aged persons to reduce later Alzheimer's disease risk: the Gray Matters randomized trial. Alzheimers Dement (N Y) 2015 Jun;1(1):53-62 [FREE Full text] [doi: 10.1016/j.trci.2015.05.001] [Medline: 29854925]
70. Peacock OJ, Western MJ, Batterham AM, Stathi A, Standage M, Tapp A, et al. Multidimensional individualised Physical ACTivity (Mi-PACT)--a technology-enabled intervention to promote physical activity in primary care: study protocol for a randomised controlled trial. Trials 2015 Aug 28;16:381 [FREE Full text] [doi: 10.1186/s13063-015-0892-x] [Medline: 26314577]
71. Shin DW, Joh H, Yun JM, Kwon HT, Lee H, Min H, et al. Design and baseline characteristics of participants in the Enhancing Physical Activity and Reducing Obesity through Smartcare and Financial Incentives (EPAROSFI): a pilot randomized controlled trial. Contemp Clin Trials 2016 Mar;47:115-122. [doi: 10.1016/j.cct.2015.12.019] [Medline: 26744232]
72. Slootmaker SM, Paw MJ, Schuit AJ, Seidell JC, van Mechelen W. Promoting physical activity using an activity monitor and a tailored web-based advice: design of a randomized controlled trial [ISRCTN93896459]. BMC Public Health 2005 Dec 15;5:134 [FREE Full text] [doi: 10.1186/1471-2458-5-134] [Medline: 16356182]
73. van der Meij E, Huirne JA, Bouwsma EV, van Dongen JM, Terwee CB, van de Ven PM, et al. Substitution of usual perioperative care by eHealth to enhance postoperative recovery in patients undergoing general surgical or gynecological

procedures: study protocol of a randomized controlled trial. JMIR Res Protoc 2016 Dec 21;5(4):e245 [FREE Full text] [doi: 10.2196/resprot.6580] [Medline: 28003177]
74. van Hoye K, Boen F, Lefevre J. The effects of physical activity feedback on behavior and awareness in employees: study protocol for a randomized controlled trial. Int J Telemed Appl 2012;2012:460712 [FREE Full text] [doi: 10.1155/2012/460712] [Medline: 23056040]
75. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Wiejden T, de Wittel L. A monitoring and feedback tool embedded in a counselling protocol to increase physical activity of patients with COPD or type 2 diabetes in primary care: study protocol of a three-arm cluster randomised controlled trial. BMC Fam Pract 2014;15:93 [FREE Full text] [doi: 10.1186/1471-2296-15-93] [Medline: 24885096]
76. Wolk S, Meißner T, Linke S, Müssle B, Wierick A, Bogner A, et al. Use of activity tracking in major visceral surgery-the Enhanced Perioperative Mobilization (EPM) trial: study protocol for a randomized controlled trial. Trials 2017 Dec 21;18(1):77 [FREE Full text] [doi: 10.1186/s13063-017-1782-1] [Medline: 28222805]
77. Brakenridge CL, Fjeldsoe BS, Young DC, Winkler EA, Dunstan DW, Straker LM, et al. Organizational-level strategies with or without an activity tracker to reduce office workers' sitting time: rationale and study design of a pilot cluster-randomized trial. JMIR Res Protoc 2016 May 25;5(2):e73 [FREE Full text] [doi: 10.2196/resprot.5438] [Medline: 27226457]
78. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015;12:42 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]
79. Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. 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]
80. 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]
81. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018 Aug 9;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
82. Kinnunen H, Tanskanen M, Kyröläinen H, Westerterp KR. Wrist-worn accelerometers in assessment of energy expenditure during intensive training. Physiol Meas 2012 Nov;33(11):1841-1854. [doi: 10.1088/0967-3334/33/11/1841] [Medline: 23110981]
83. Brown DK, Grimwade D, Martinez-Bussion D, Taylor MJ, Gladwell VF. The validity of the ActiPed for physical activity monitoring. Int J Sports Med 2013 May;34(5):431-437. [doi: 10.1055/s-0032-1323723] [Medline: 23184480]
84. Slootmaker SM, Chin AP, Schuit AJ, van Mechelen W, Koppes LL. Concurrent validity of the PAM accelerometer relative to the MTI Actigraph using oxygen consumption as a reference. Scand J Med Sci Sports 2009 Feb;19(1):36-43. [doi: 10.1111/j.1600-0838.2007.00740.x] [Medline: 18266793]
85. Kim DY, Jung Y, Park R, Joo N. Different location of triaxial accelerometer and different energy expenditures. Yonsei Med J 2014 Jul;55(4):1145-1151 [FREE Full text] [doi: 10.3349/ymj.2014.55.4.1145] [Medline: 24954349]
86. Manohar CU, Koepp GA, McCrady-Spitzer SK, Levine JA. A stand-alone accelerometer system for free-living individuals to measure and promote physical activity. Infant Child Adolesc Nutr 2012 Jun 11;4(4):222-229. [doi: 10.1177/1941406412450644]
87. Wilmot EG, Davies MJ, Edwardson CL, Gorely T, Khunti K, Nimmo M, et al. Rationale and study design for a randomised controlled trial to reduce sedentary time in adults at risk of type 2 diabetes mellitus: project stand (Sedentary Time ANd diabetes). BMC Public Health 2011 Dec 8;11:908 [FREE Full text] [doi: 10.1186/1471-2458-11-908] [Medline: 22151909]
88. Schunemann H, Brozek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. In: The GRADE Working Group. Hamilton, Canada: McMaster University; 2015:2013.
89. Lewis ZH, Lyons EJ, Jarvis JM, Baillargeon J. Using an electronic activity monitor system as an intervention modality: a systematic review. BMC Public Health 2015;15:585 [FREE Full text] [doi: 10.1186/s12889-015-1947-3] [Medline: 26104189]
90. de Vries HJ, Kooiman TJ, van Ittersum MW, van Brussel M, de Groot M. Do activity monitors increase physical activity in adults with overweight or obesity? A systematic review and meta-analysis. Obesity (Silver Spring) 2016 Oct;24(10):2078-2091 [FREE Full text] [doi: 10.1002/oby.21619] [Medline: 27670401]
91. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. Sports Med Open 2018 Sep 3;4(1):42 [FREE Full text] [doi: 10.1186/s40798-018-0157-9] [Medline: 30178072]
92. Mitchell MS, Goodman JM, Alter DA, John LK, Oh PI, Pakosh MT, et al. Financial incentives for exercise adherence in adults: systematic review and meta-analysis. Am J Prev Med 2013 Nov;45(5):658-667. [doi: 10.1016/j.amepre.2013.06.017] [Medline: 24139781]

93. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
94. Fergusson D, Aaron SD, Guyatt G, Hébert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. Br Med J 2002 Sep 21;325(7365):652-654 [FREE Full text] [Medline: 12242181]
95. Samdal GB, Eide GE, Barth T, Williams G, Meland E. Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. Int J Behav Nutr Phys Act 2017 Dec 28;14(1):42 [FREE Full text] [doi: 10.1186/s12966-017-0494-y] [Medline: 28351367]
96. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF. Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis. Soc Sci Med 2017 Nov;192:125-133. [doi: 10.1016/j.socscimed.2017.09.021] [Medline: 28965003]
97. Lindahl B, Nilssön TK, Borch-Johnsen K, Røder ME, Söderberg S, Widman L, et al. A randomized lifestyle intervention with 5-year follow-up in subjects with impaired glucose tolerance: pronounced short-term impact but long-term adherence problems. Scand J Public Health 2009 Jun;37(4):434-442. [doi: 10.1177/1403494808101373] [Medline: 19181821]
98. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011;8:115 [FREE Full text] [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
99. Sylvia LG, Bernstein EE, Hubbard JL, Keating L, Anderson EJ. Practical guide to measuring physical activity. J Acad Nutr Diet 2014 Feb;114(2):199-208 [FREE Full text] [doi: 10.1016/j.jand.2013.09.018] [Medline: 24290836]
100. Prince SA, Saunders TJ, Gresty K, Reid RD. A comparison of the effectiveness of physical activity and sedentary behaviour interventions in reducing sedentary time in adults: a systematic review and meta-analysis of controlled trials. Obes Rev 2014 Nov;15(11):905-919 [FREE Full text] [doi: 10.1111/obr.12215] [Medline: 25112481]
101. Commissaris DA, Huysmans MA, Mathiassen SE, Srinivasan D, Koppes LL, Hendriksen IJ. Interventions to reduce sedentary behavior and increase physical activity during productive work: a systematic review. Scand J Work Environ Health 2016 May 1;42(3):181-191. [doi: 10.5271/sjweh.3544] [Medline: 26683116]
102. Martin A, Fitzsimons C, Jepson R, Saunders DH, van der Ploeg HP, Teixeira PJ, et al. Interventions with potential to reduce sedentary time in adults: systematic review and meta-analysis. Br J Sports Med 2015 Aug;49(16):1056-1063. [doi: 10.1136/bjsports-2014-094524] [Medline: 25907181]
103. Vaes AW, Cheung A, Atakhorrami M, Groenen MT, Amft O, Franssen FM, et al. Effect of 'activity monitor-based' counseling on physical activity and health-related outcomes in patients with chronic diseases: a systematic review and meta-analysis. Ann Med 2013 Sep;45(5-6):397-412. [doi: 10.3109/07853890.2013.810891] [Medline: 23952917]

Abbreviations

GRADE: Grades of Recommendation, Assessment, Development, and Evaluation
IQR: interquartile range
MeSH: Medical Subject Headings
MVPA: moderate and vigorous physical activity
PAM: Personal Activity Monitor
SMD: standardized mean difference

Edited by G Eysenbach; submitted 02.08.18; peer-reviewed by C Friedrich, L Li, K Tamura, J Volmerg; comments to author 12.10.18; revised version received 11.11.18; accepted 23.01.19; published 12.04.19.

Please cite as:
Brickwood KJ, 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;7(4):e11819
URL: https://mhealth.jmir.org/2019/4/e11819/
doi:10.2196/11819
PMID:30977740

Review

Augmented Reality in Medicine: Systematic and Bibliographic Review

Martin Eckert[1*], BSc; Julia S Volmerg[1*], BSc; Christoph M Friedrich[1,2], PhD

[1]Department of Computer Science, University of Applied Sciences and Arts Dortmund, Dortmund, Germany

[2]Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, Essen, Germany

[*]these authors contributed equally

Corresponding Author:
Christoph M Friedrich, PhD
Department of Computer Science
University of Applied Sciences and Arts Dortmund
Emil-Figge Straße 42
Dortmund, 44227
Germany
Phone: 49 2319112 ext 6796
Email: christoph.friedrich@fh-dortmund.de

Abstract

Background: Augmented reality (AR) is a technology that integrates digital information into the user's real-world environment. It offers a new approach for treatments and education in medicine. AR aids in surgery planning and patient treatment and helps explain complex medical situations to patients and their relatives.

Objective: This systematic and bibliographic review offers an overview of the development of apps in AR with a medical use case from March 2012 to June 2017. This work can aid as a guide to the literature and categorizes the publications in the field of AR research.

Methods: From March 2012 to June 2017, a total of 1309 publications from PubMed and Scopus databases were manually analyzed and categorized based on a predefined taxonomy. Of the total, 340 duplicates were removed and 631 publications were excluded due to incorrect classification or unavailable technical data. The remaining 338 publications were original research studies on AR. An assessment of the maturity of the projects was conducted on these publications by using the technology readiness level. To provide a comprehensive process of inclusion and exclusion, the authors adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.

Results: The results showed an increasing trend in the number of publications on AR in medicine. There were no relevant clinical trials on the effect of AR in medicine. Domains that used display technologies seemed to be researched more than other medical fields. The technology readiness level showed that AR technology is following a rough bell curve from levels 4 to 7. Current AR technology is more often applied to treatment scenarios than training scenarios.

Conclusions: This work discusses the applicability and future development of augmented- and mixed-reality technologies such as wearable computers and AR devices. It offers an overview of current technology and a base for researchers interested in developing AR apps in medicine. The field of AR is well researched, and there is a positive trend in its application, but its use is still in the early stages in the field of medicine and it is not widely adopted in clinical practice. Clinical studies proving the effectiveness of applied AR technologies are still lacking.

(JMIR Mhealth Uhealth 2019;7(4):e10967) doi:10.2196/10967

KEYWORDS

mixed/augmented reality; medicine; mobile computing; systematic review; mobile phone

Introduction

Background

Augmented reality (AR) is a technology that extends the user's reality using digital information. It has become a publicly discussed topic in our society and a prime field for new kinds of apps in the medical sector. AR can be seen in many aspects of medicine; for example, Kamphuis et al reported that AR technologies have started maturing in the field of anatomical and physiological education [1]. There is a high demand for assisting systems due to increased stress in public health systems, which is one of the reasons for the fast development within the field of AR and virtual reality (VR).

The focus of this systematic and bibliographic review is to provide insight into the research conducted in the field of AR. The scope lies in its advances in the medical field, centering on the medical specialty, technical impact, maturity of projects, and publications on the topic of AR in medicine.

Recently Chen et al [2] published a review covering the development of AR technology in the medical field. Their study thematically overlaps with this review paper, but in their paper [2], only Scopus [3] was used as a data source and a text mining approach was used to analyze the retrieved data. In contrast, this review uses manual checking, broader categorization, and the PubMed database in addition to Scopus.

Within the review presented here, the publications have been manually categorized and structured according to medical branches, applied technologies in hardware and software, and strong indicators of the maturity of the developed technology.

For the classification of AR apps, the taxonomy proposed in the Handbook of Augmented Reality by Hugues et al [4] was used. The taxonomies by Schmalstieg et al [5] and Aukstakalnis [6] were used to categorize AR displays and the technique of tracking. For the assessment of maturity, the technology readiness level [7] was used, which is a method to rank and analyze the demonstrated technologies.

The aim of this review paper was to build a foundation and guide to the literature and to be used as a motivation for scientists, researchers, and developers in the field of AR in medical settings. The main objectives were to assess the current state of research, identify possible future trends, and provide an overview. Another outcome of this research was an interactive table (Multimedia Appendix 1) of the research conducted in AR in the field of medicine from March 2012 to June 2017.

Overview of Augmented Reality Technology

This section offers a brief overview of AR and its definitions, differentiating it from VR. Additionally, a short overview of its technical development and its applications will be given.

There are several definitions of AR, depending on how it enhances our environment with artificially added information and whether one can interact with this information [8].

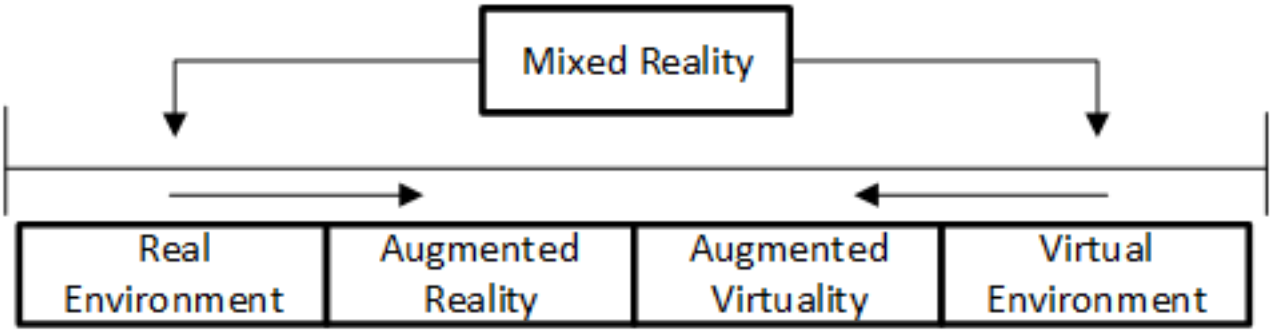
AR is a variation of VR [9]. In contrast to VR, a user of an AR system always experiences their own reality in real-time. A VR system always has a synthetic feature, and it imitates reality rather "than supplements the real world" [10], although there are VR systems that imitate AR by using cameras showing the user his/her surroundings augmented with additional information. Azuma et al [9] reported that a VR environment is a completely synthetic environment that separates the user from reality.

Another term used in this context is mixed reality (MR), which could be explained by the "reality-virtuality continuum" explained by Milgram et al [11]. MR shows the reality at one end and VR at the other end, with AR and augmented virtuality (AV) lying between the two ends (Figure 1).

Milgram et al used the term MR to distinguish different MR displays and design a taxonomy for categorization of MR systems [11]. In addition, augmentation through senses other than vision is important, but not as common. The auditory sense or haptic sense is an additional source of information. Each origin of additional information can be classified as AR/MR.

From a historical viewpoint, the first approach to AR was the Sensorama, a machine that was supposed to provide a cinematic experience with all senses [8]. It was developed by Heilig in the 1950s and was the first documented reference to AR, although at that time, there was no distinction between AR and VR. In 1968, Sutherland [12] developed a head-mounted display (HMD), which made it possible to experience AR and VR environments for the first time. The first reference to AR as a term was made by Caudell, a researcher at Boeing who coined the term in 1990 [13]. Two years later, Caudell and Mizell [14] developed an early prototype, which enabled technicians to project blueprints onto a surface.

Figure 1. "Reality-Virtuality Continuum" by Milgram et al [11].



With the invention of handheld devices like smartphones and tablets, there were opportunities to advance to a bigger audience. In 2013, Google presented Google Glass, an HMD that provides hands-free interaction via a voice interface and enables the user to call, send texts, or search the internet. In 2015, Microsoft presented the HoloLens; this device allows one to see and interact with holographic 3D virtual objects via voice, gaze, and gestures.

Methods

This section provides an overview of the methods used to acquire relevant publications in addition to the restrictions and inclusion and exclusion criteria used. Thereafter, an introduction to the principles used to analyze the publications and the relevant variables are presented.

Acquisition of Publications

To gather the initial data, a query was designed to collect publications from the PubMed [15] and Scopus [3] databases, containing the keywords “augmented” and “reality” restricted to the period between May 15, 2012, and June 30, 2017. The search for Scopus was additionally restricted by the subject area medicine,” and the timeframe was manually readjusted to the timeframe used for PubMed. The term “mixed reality” was applied synonymously to AR, but was not explicitly searched for.

In three stages, the results of the query were filtered using a set of factors to determine the validity and relevance for this review. After each exclusion stage, the remaining papers were analyzed, and the corresponding variables were entered into the results file. We then processed each batch individually. To ensure reliability during the classification phase, a subset of 85 (25%) of 338 publications were cross-checked by the authors, and conflicts were resolved by consensus.

In the next step, data were analyzed and prepared for visualization, where applicable. Analysis and visualization were conducted using R (version 3.4.2; “Short Summer”) [16].

Data Analysis

Exclusion and Inclusion Criteria

The initial queries returned 1309 results (Multimedia Appendix 2), from which 340 duplicates were removed, leaving 969 eligible publications. The authors of these publications were contacted if the publications were not accessible. Subsequently, mismatched publications were excluded in three iterations, yielding a total of 338 publications. The last iteration was an additional categorization of review publications (Multimedia Appendix 2). A visual overview of the filtering process is shown in Figure 2, which follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart [17]. The complete dataset analyzed, along with the results in the form of an interactive table, is available in Multimedia Appendix 1. For a complete reference list, see Multimedia Appendix 3.

Table 1 shows the exclusion criteria. The main reason for exclusion of a publication was the absence of a concrete connection to medical treatment or training of health professionals, or “No Treatment or Training,” such as papers about the influence of Pokémon Go on college students [18].

The second main reason for exclusion was a false positive result, a criterion that includes publications without any connection to AR because they either focus entirely on VR or separately contain the tags “augmented” and “reality.” This included the paper by Yoo et al about the effect of training alone or training with VR, both augmented by electromyography, for children with cerebral palsy [19].

The exclusion criteria “Other” refers to papers where only the abstract could be found, papers where the authors did not reply, conference posters, collections of papers, and books. One additional criterion for exclusion was veterinary publications, for example, the paper by Sutton et al about the glass knife-fish and the way it uses “electrosensory feedback” to hold a position in a moving environment [20].

Figure 2. Filtering Process of the initial query and remaining results. NTT = No Treatment or Training, FP = False Positive.

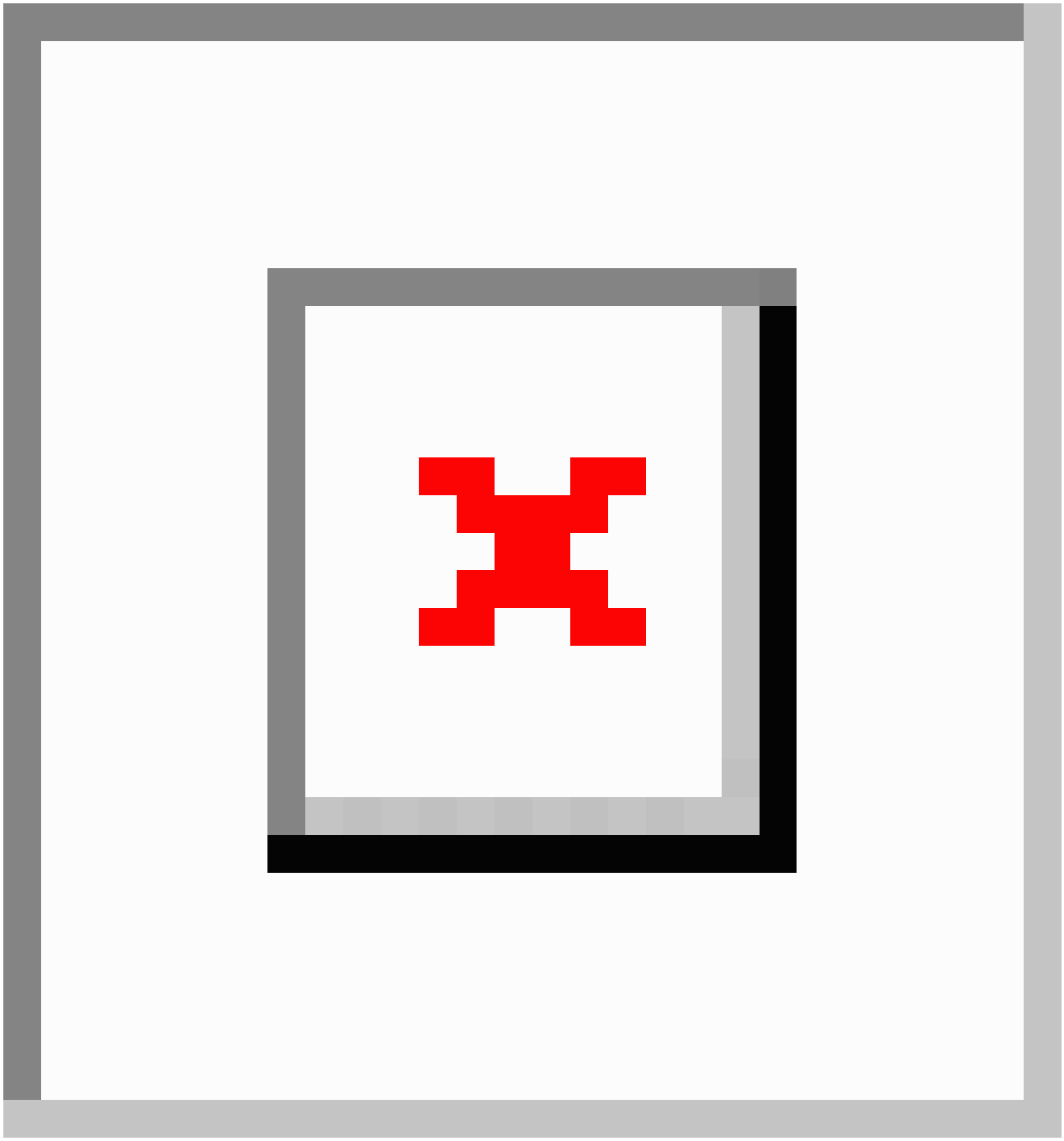


Table 1. Overview of exclusion criteria and number of excluded publications.

Exclusion criteria	Number of excluded publications
No treatment or training	257
False positive	173
Paper not in English	13
Veterinary medicine	3
Other	81
Duplicates	340
Review	104
Total	971

Table 2. Overview of the investigated variables and the related classification method.

Variable	Classification method
Year of publication	Metadata
Geolocation	Metadata
MeSH[a] data	MeSH [21]
Medical scope	Manual
Interactive or haptic	Manual
Collaboration	Manual, binary
Clinical trial	Clinicaltrials.gov [22] and manual, binary
Augmented reality taxonomy	Hugues et al [4] and manual
Technology readiness level	US Department of Defense [7] and manual
Display	Schmalstieg et al [5] and manual
Tracking	Schmalstieg et al [5] and manual

[a]MeSH: Medical Subject Headings.

Summary of Variables

A short overview of the variables used is presented in Table 2.

Classifications

To classify the publications, several factors were chosen to create a comparable dataset within the research of AR applied to medicine. Besides the classification methods from the metadata, such as the year of publication and geographical location of the first author, several other classification factors were manually introduced into the dataset.

Medical Subject Headings

For medical classification, the Medical Subject Headings (MeSH) terms were used. These terms, only applicable to the PubMed database, offer a terminology for categorization of biomedical apps and can be used, for example, to categorize publications in PubMed. For this review, the 2017 version of MeSH terms was used.

Clinical Trial

For additional classification, data from ClinicalTrials.gov [22] were used. The database assesses if the publication, in case it is a prototype or product, was tested in a clinical environment during a study and if it was registered on ClinicalTrials.gov.

Collaboration

This variable was used to assess how collaboration using the AR device was realized. This could be done via a remote connection. For example, this category included studies in which two medical professionals present at different locations see the same AR environment and are allowed to interact with it.

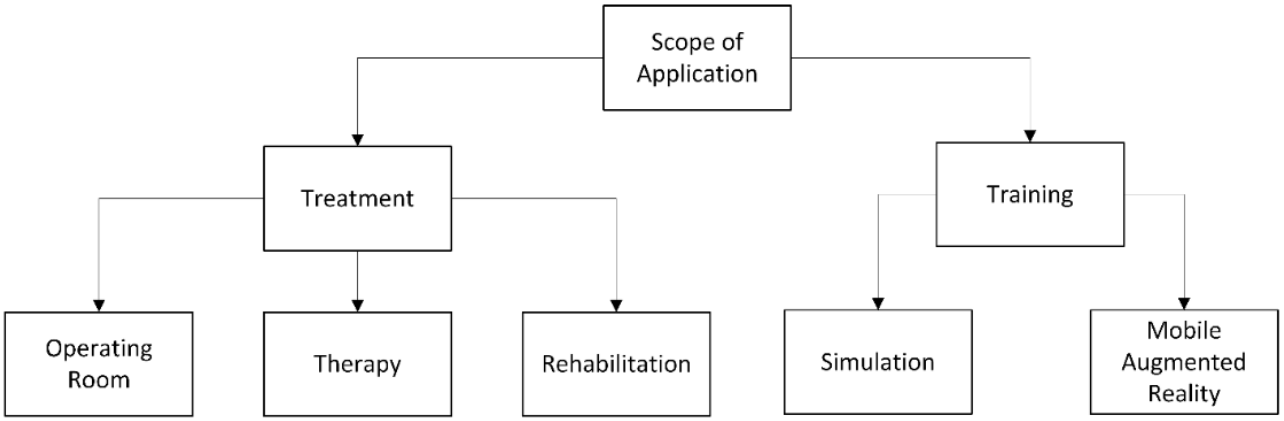
Research Maturity

The technology readiness level assessment is a method to estimate the maturity of technology (Table 3). Usually, the level ranges from 1 to 9, where 1 is the least and 9 is the most matured technology. It was introduced by the US Department of Defense to rate technologies. The US Department of Defense definition from 2011 [7] was used to determine and quantify the state of research projects in this paper.

Table 3. Technology readiness levels according to the US Department of Defense [7].

Technology readiness level	Description
1	Basic principles observed and reported
2	Technology concept or application formulated
3	Analytical and experimental critical function or characteristic proof of concept
4	Component or breadboard validation in a laboratory environment
5	Component or breadboard validation in a relevant environment
6	System/subsystem model or prototype demonstration in a relevant environment
7	System prototype demonstration in an operational environment
8	Actual system completed and qualified through test and demonstration
9	Actual system proved through successful mission operations

Figure 3. Medical scope criteria for application.



Medical Scope

The scope of the published apps was manually split into two subgroups—treatment and training—as shown in Figure 3. The treatment group was further subdivided into three groups: operating room, therapy, and rehabilitation. The training group was subdivided into simulation and mobile AR. The subgroups originated after the short initial overview of the publications.

Augmented Reality Display

In this section, we present a short introduction to the AR display technology. Schmalstieg et al [5] divided the placement of the display technologies into three spaces: “head,” “body,” and “world.” The placement can be identified by looking at where the display is stationed. “Desktop displays, Virtual Mirrors, Virtual Showcase, Window and Portal Displays and Projector-based displays” [5] are located in a fixed place in the world; “Handheld displays” [5] are stationed on the body; and a “Near-Eye Display” [5] is placed on the head of the user.

Head Location

The HMD is the most well-known display technology used for AR. These displays are further classified into “Optical-See-Through Head-Mounted Display” and “Video-See-Through Head-Mounted Display” [5]. The Optical-See-Through enables a user to see the real world augmented with information via see-through lenses such as the Microsoft HoloLens. Videos see-through displays use an additional camera to provide the user with surrounding reality. Thus, it is not seen directly by the user (eg, HTC Vive). An example of this technology in a medical context is the digital microscope in an operating room.

Body Location

The commonly known smartphone and tablet are best examples of handheld AR devices and are mostly known for games like Pokémon Go. The back camera of such devices is used to provide the user with a video see-through image. In a medical context, this could be an app to visualize anatomical structures in a book.

World Location

A display located in the world has many advantages for AR. These displays can be divided into desktop display, virtual mirror, projector-based display, and stationary display.

A desktop display can be used as an AR display by adding a webcam, which then provides the necessary input of reality. The display then shows both reality and additional information within it. A virtual mirror uses the front camera of the device and shows a picture of what is in front of the camera. Another kind of an augmented mirror is the visual showcase, but it does not show the user; it is a stationary variant of the optical see-through, which allows the user to see through alongside additional information. The last stationary displays are projector-based displays. They can be dependent or independent from the view.

As another option, one can simply mount the projectors onto an HMD, creating a personal projection with retro-reflection screens. The location “World” can be found in any medical context; a known example here would be in the operating room, where images acquired prior to surgeries are either shown on a display or projected directly onto the patient.

Augmented Reality Tracking

Tracking in an AR environment can be categorized by the different technologies used to track objects. In this paper, the applied categorization refers to whether the sensors are stationary or mobile and which type of optical tracking was used [5].

Today, many systems use more than one tracking mechanism; the HTC Vive system uses stationary tracking in the form of stationary lighthouses in the room with beam detectors in the tracked devices and mobile tracking in the form of gyro sensors.

Stationary Tracking Systems

Stationary tracking is a technique that uses either a mechanical device or an electromagnetic field, infrared light, or ultrasound to obtain position data. Commonly known stationary tracking systems are the SteamVR trackers, which use a time-based location estimation for the HTC Vive or other AR/VR solutions. Another option is Ultrasonic Tracking [5], which uses an ultrasonic pulse as a time-of-flight source. Therefore, it is possible to track a position by measuring the time a pulse needs from the source to the sensor.

Mobile Tracking Systems

While stationary tracking does not allow the user to move around much, mobile sensors allow tracking outdoors. The most popular mobile sensor is the Global Positioning System, which

determines the position of an object by the time of flight of signals emitted by a satellite. Inside an already established wireless network, it is also possible to track an object simply by the base station used to connect to the wireless network.

Optical Tracking

There are two ways to classify tracking in AR: "model-based versus model-free tracking" and "markers/fiducial versus natural features" [5]. Model-based tracking uses an existing model that is created beforehand. The model-free tracking is an on-the-fly technique, with a temporary model. This allows for more flexibility, especially if there is a combination of 3D tracking and 3D scanning. Marker tracking, also known as fiducial tracking, offers a possibility for more robust algorithms because markers are previously known patterns that are more easily recognized. Natural feature markers often require higher image quality to detect the object.

Interest points, also known as key points, are mostly used to track an object, but the key points should be easily detected and stable from all angles. An alternative for interest points is edge features, but here, it is necessary to distinguish the edge from the background.

Augmented Reality Taxonomy

In this section, the parts of the "Functional Taxonomy" [4], introduced by Hugues et al, that are used for the classification will be explained. The functionality "Artificial Environment" was excluded, since it defines AR in the context of time. An overview of the part of the taxonomy used is shown in Figure 4.

Six subfunctionalities of augmented perception can be separated by considering the ability to assist in decision making for AR. The first is functionality 0, where real images and virtual entities are shown on the screen but have no relation. This conditioning indicates that there is one screen but two different boxes: One box shows real images and one shows a virtual entity.

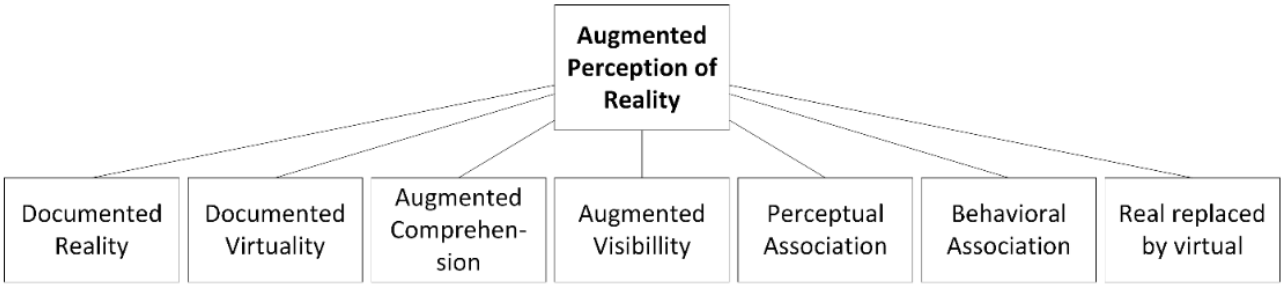
Subfunctionality 1: Documented Reality and Documented Virtuality

The functionality "Documented Reality and Documented Virtuality" [4] is a minimal function of AR. Augmentation only consists of two boxes: One displays the real images and one displays the virtual entity. However, on distinguishing functionality 1 from functionality 0, real images and virtual entities are related to each other in functionality 1 and therefore provide additional information within the context of reality.

Subfunctionality 2: Reality With Augmented Perception or Understanding

In "Reality with Augmented Perception or Understanding" [4] environments, there is only one box left, which is shared by real images and virtual entities. The subfunctionality can be categorized into two levels—"Augmented Understanding" and "Augmented Visibility." "Augmented Understanding" means that the virtual entities show alignment with real images but are not always close to each other, and "Augmented Visibility" means that the virtual entities cover the real images completely.

Figure 4. Functional taxonomy according to Hugues et al [4].



Subfunctionality 3: Perceptual Association of the Real and Virtual Images

The subfunctionality "Perceptual Association of the Real and Virtual" [4] is divided further into the levels of "Incrustation" [4] and "Integration" [4] of virtual entities on real images; therefore, this functionality has the ability to differentiate between a projection of a tumor on top of an organ and the 3D image of the tumor onto the whole organ.

Subfunctionality 4: Behavioral Association of the Real and Virtual Images

The "Behavioural [sic] Association of the Real and Virtual" [4] is a step further from the "Perceptual Association" functionality, as it adds physical properties to virtual objects, specifically the properties of the real object.

Subfunctionality 5: Substituting the Real by the Virtual or Virtualized Reality

"Substituting the Real by the Virtual or Virtualised [sic] reality" [4] is a subfunctionality that allows the real scene to be replaced by an artificial image and vice versa. Therefore, it is possible to change the angle of the view, which makes it possible to see not only reality through a camera but also an artificial image from another point of view.

Results

Review Publications

The query results contained 104 review publications. Most of the reviews assessed the use of AR in a surgery setting. The review publications can be found in Multimedia Appendix 2.

Khor et al [23] provided insight into the use of AR and VR, with an emphasis on the surgical workplace. Thomas et al [24] highlighted a computer-aided medicine revolution. Moglia et

al [25] provided a review of VR and AR simulators for robot-assisted surgery. Robotic surgery is also the topic of another publication from 2015 [26]. Slade et al provided an assessment [27] of the use of wearable technology in a surgical setting. Bluemel et al [28] gave an overview of freehand single-photon emission computed tomography (SPECT) for navigation and radio-guided surgical procedures like sentinel lymph node biopsies.

These publications do not include neurosurgical reviews. However, Pelargos et al [29] refined a review that covers the historical development, current use, and emerging applications of AR and VR in the field of neurosurgery. A recent study provided a systematic overview of technologies using AR in the field of neurosurgery [30]. Marcus et al [31] published an overview of robotics in keyhole transcranial endoscope-assisted microsurgery.

In total, two publications assessed the use of AR technologies in Urology. One study [32] provided an overview of the history and current state of pediatric robotic surgery, mainly in India. It covered the topic of AR in the outlook, assessing its potential to support pediatric robotic urology. Hamacher et al [33] published a paper on the development of VR, AR, and MR in existing consumer products, in which the main emphasis was on VR. This review paper covers the influence of these technologies used in urology.

Several endoscopic devices are equipped with AR technologies. Mahmud et al [34] provide a prognosis to integrate AR technology into an endoscopic device. They emphasized the importance of collaboration between computer scientists and physicians. Feussner et al [35] stressed upon the importance of a close collaboration between programmers and physicians. They compiled an extensive overview of available technologies and identified associated technological problems. They also estimated the time to bring the technologies to a broadly applicable system.

Smith et al described the use of AR in education in the fields of Obstetrics and Gynecology [36]. There is an emphasis on the change in clinical education, to empower students to use new technologies such as VR and AR devices as well as holograms while teaching in a focused and comprehensive manner.

Year of Publication

Table 4 shows the distribution of publications over the time. We observed an increase in the number of papers published from 2012 to 2014, a decrease in 2015, and a peak in 2016. The fact that only three papers were electronically pre-published in 2011 can be explained by the rise of electronic pre-publishing in 2011.

Geolocation

As shown in Table 5, in the majority of publications analyzed, the first author was located in the United States, followed by Germany, Japan, and France. For brevity, this table is cutoff after the top 10 locations.

Table 4. Evaluated publications according to publication date from March 2012 to June 2017 (N=338).

Year of publication	Number of publications
2011	3
2012	25
2013	51
2014	81
2015	59
2016	71
2017	48

Table 5. Distribution of the top 10 geolocations of the first author (N=264).

Country	Number of publications
United States	73
Germany	40
Japan	27
France	25
China	25
Canada	25
Switzerland	14
Spain	13
United Kingdom	11
Italy	11

Table 6. Analysis of the top 10 Medical Subject Headings terms sorted and shown by frequency (N=191).

Medical Subject Headings terms	Number of publications
[Surgery, Computer-Assisted]	77
[Imaging, Three-Dimensional]	65
[Tomography, X-Ray Computed]	39
[Laparoscopy]	24
[Feasibility Studies]	23
[Neurosurgical Procedures]	15
[Pilot Projects]	14
[Computer-Assisted Instruction]	12
[Image Interpretation, Computer-Assisted]	11
[fluoroscopy]	11

Table 7. Overview of the results from the variables Clinical Trial and Collaboration (N=338 for each).

Variable	Proposed app tested in a clinical trial	
	Yes	No
Clinical Trial	3	335
Collaboration	16	322

Medical Subject Headings Terms

The publications from PubMed feature the title and abstract and are annotated with MeSH terms [21]. These publications were not included in the Scopus database. To gain more insight into the nature of the publications, the MeSH terms were analyzed. Of the 267 publications from PubMed, at the time of analysis, only 191 featured MeSH-annotated terms. These terms were added by experts, and therefore, not all recent publications were annotated. The term frequency was calculated, and the results of the top 10 terms are presented in Table 6.

Clinical Trial

The factor clinical trials was included to analyze if the proposed app was tested in a clinical trial registered in ClinicalTrial.gov [22]. As shown in Table 7, this occurred in three publications, including a study by Ortiz-Catalan [37] who registered their study in ClinicalTrial.gov; this study was about phantom limb pain and aimed to show that treatment with AR decreases pain.

Collaboration

Table 7 shows the results of a possible collaboration through the app. As shown in most cases, this was not possible; only 16 projects enabled the user to share their AR space simultaneously with other users. This can occur through a remote connection or on a local level, as shown by Vera at al [38] who presented an AR telemonitoring platform to support students learning laparoscopic techniques by overlaying the students' view with the view of the mentor. Shared open-world displays were not considered to be shareable, even if every computer monitor represents a collaboration.

Research Maturity

Research maturity, categorized by the technology readiness level index [7], shows a trend in levels 6 and 7 (Table 8). There are also publications that contain research on the effectiveness of AR products; these are represented in level 9. The most common technologies featured in these publications are laparoscopic tools.

Medical Scope

The results of the Medical scope are presented in Figure 5. Of the 338 publications, 84.3% (n=285) were identified as AR projects that dealt with the actual treatment of patients and 15.7% (n=53) dealt with training scenarios, for example, a clinical simulation feasibility study with Google Glasses [39].

Of the 285 publications identified to be in the treatment category, 69.5% (n=198) projects dealt with scenarios in the operating room and 9.1% (n=26) projects were set in a rehabilitation setting. In addition, 21.4% (n=61) of the presented projects included direct involvement in the therapy of the patient.

Classification by Schmalstieg et al

The results of the Display classification are shown in Figure 6. The Sankey plot shows that most of the displays are classified as "World," and most of those have a common "Display." For example, Kranzfelder [40] presented a system to add information from computerized tomography onto the images of gastrointestinal endoscopy. In the stream "Head," one exemplary study by Carenzo et al [41] used Google Glass in a disaster medicine scenario as a tool to provide a triage algorithm in order to help assign triage codes to those injured in a disaster situation.

Table 8. Projection of technology readiness level distribution for all publications (N=338).

Technology readiness level	Number of publications
2	3
3	29
4	80
5	45
6	94
7	71
8	6
9	10

Figure 5. Categorization of medical scope and count of examined augmented reality projects including reviews (N=338).

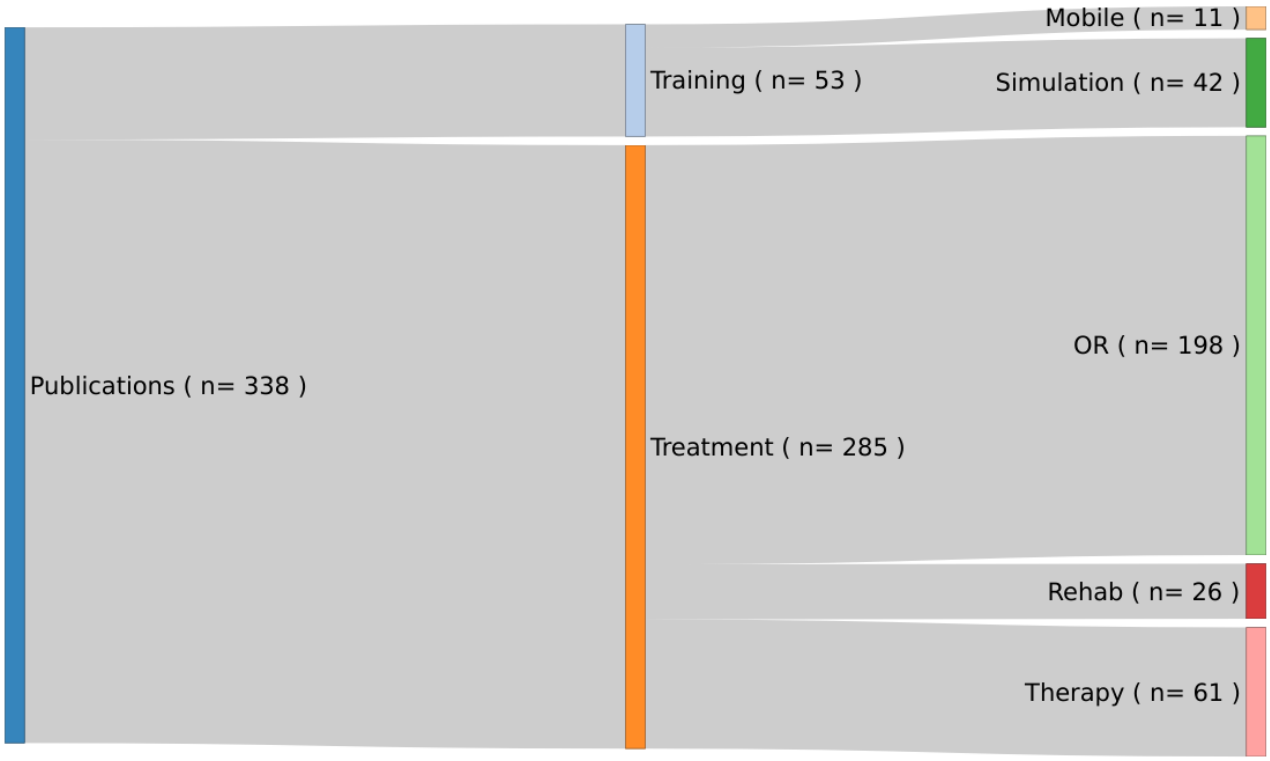


Figure 6. Categorization of display according to Schmalstieg et al [5] (N=338).

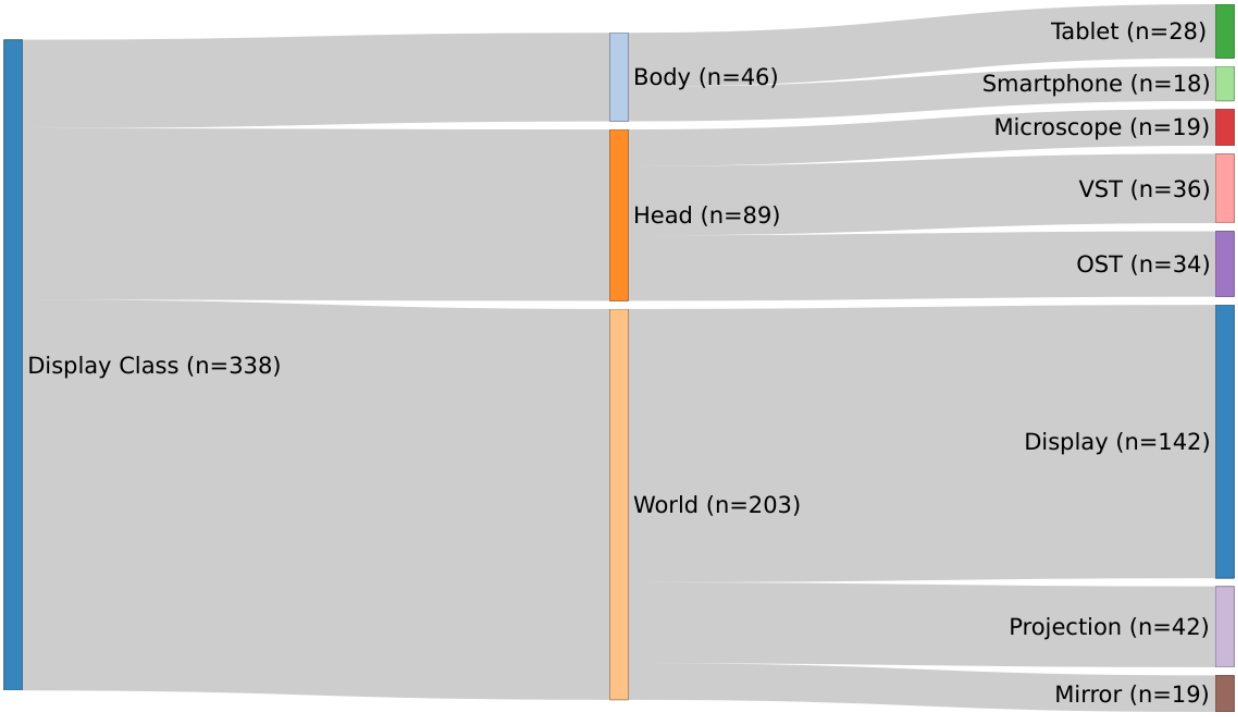


Table 9. Overview of the variables Mobile and Stationary Movement Tracking (N=338).

Movement tracking	Use of the tracking system		
	Yes	No	None
Mobile	247	63	28
Stationary	134	176	28

Within the stream "Body," some of the applications used a "Tablet" to visualize 3D information from a medical imaging system. For example, a system used a tablet showing a previously virtually reconstructed tumor during a neurosurgical operation, to assist the surgeon while drilling [42]. Another example is an app for medical students to learn about gunshot wounds [43].

Of all analyzed (N=338) publications, most optical tracking mechanisms are accomplished via a marker (n=223), mostly through commonly known mechanisms such as color variation (ie, chessboard pattern). This pattern was used by Edgcumbe et al [44], who applied the pattern of triangulation in a laparoscopic scenario. Natural Markers were used in 73 publications and a mixture of natural and marker-based optical tracking was used in 3 publications; no optical tracking was used in 39 papers.

The results for the stationary and mobile tracking mechanisms are presented in Table 5.

To avoid doubling numbers, for any device using more than one tracking mechanism, the stationary device was chosen. This occurred mostly in the Head Location or Video-See-Through scenarios, since there are devices like the HTC Vive or the Oculus Rift that depend on tracking with infrared, which are found in a stationary device or involve smaller movements that often rely on gyroscopes.

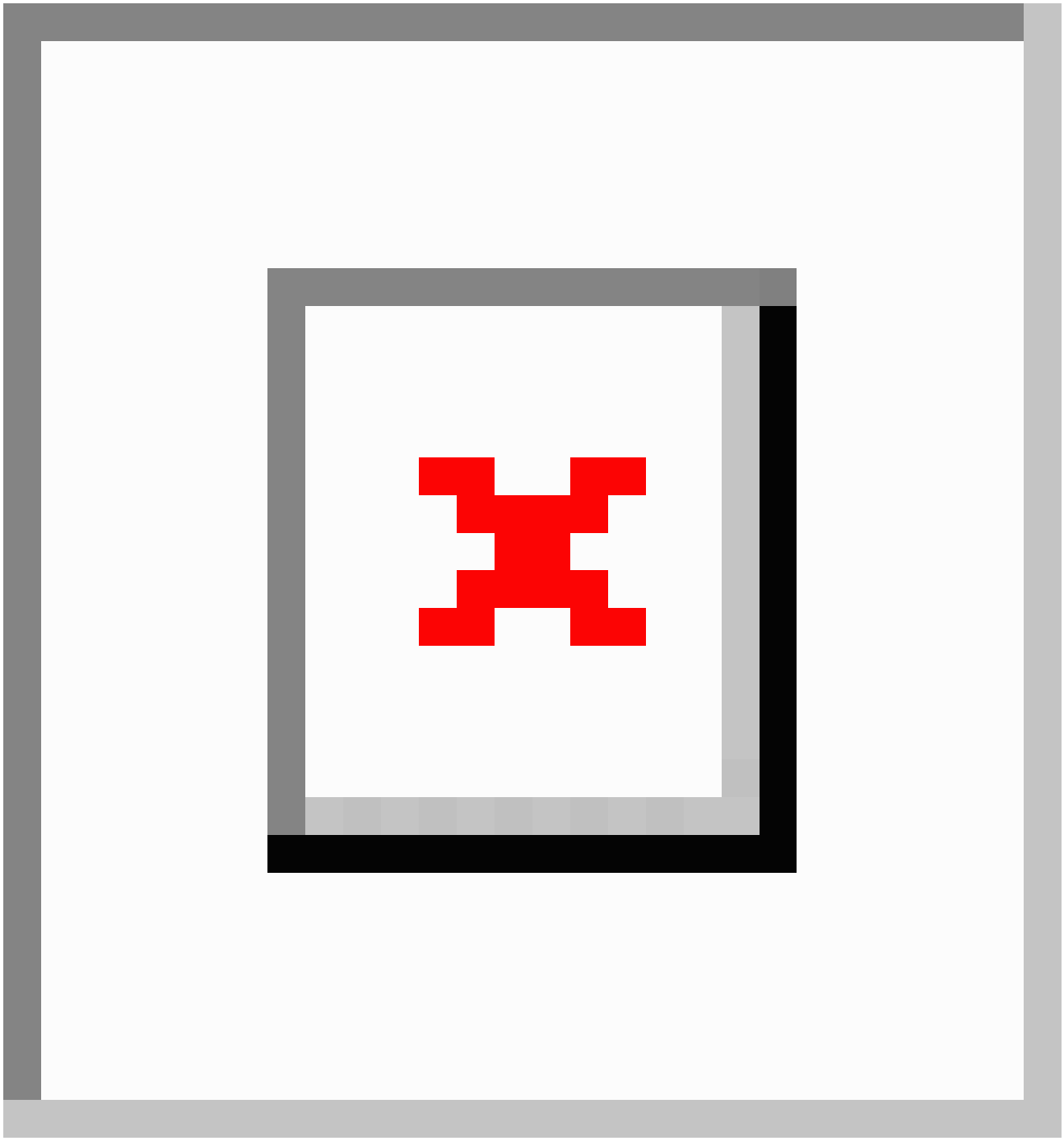
Table 9 shows that 134 of the analyzed publications used a stationary tracking system, while 28 used no tracking system and 176 did not use a stationary system. It is possible to track movement in a mobile manner, and this maneuver can be applied in a medical field. Mobile Tracking was provided in 247 of the presented apps.

In some cases, there was no tracking of movement, like the app for rehabilitation proposed in the study by Chinthammit et al [45], who overlaid an image produced by a trainer onto the reality from a patient, so that the patient could mimic the movement.

Classification by Hugues et al

Figure 7 shows that of the 338 analyzed papers, in most cases (n=191, 56.5%), a "Perceptual Association" between the real object and the virtual object can be examined. For example, KleinJan et al [46] developed an app for the declipseSPECT, a device that can present preoperative data from digital imaging and communications in medicine files in 3D. This app can combine the AR aspect of the device with fluorescence imaging by adding a fluorescence layer into the scenario, making it easier for surgeons to track and extract sentinel nodes in surgeries.

Figure 7. Categorization according to the taxonomy of Hugues et al [4] (N=338).



In the second most classified functionality—"Perceptual Understanding"—there is a connection between the real and virtual environment, similar to the app presented by Marker et al [47]. They implemented a method to add information to a screen while it is being visualized in a magnetic resonance imaging system. This approach supports the surgeon to target the paravertebral space.

The functionalities "Behavioral Association" and "Substitution" are addressed in 35 of the analyzed papers. One example is the paper by Olivieri et al [48] about an app that monitors the user via electroencephalography and gives feedback based on the concentration of the participant while training in a surgical simulation. Eleven published apps show only additional information that is unrelated to the reality that surrounds the user. This can be seen, for example, in the study by Wilson et al [49], who developed an app to treat a tension pneumothorax by decompressing it. The device showed the steps necessary to fulfill this treatment, which did not depend on where the user was looking.

Discussion

Principal Findings

Evaluating the recent publications in the analysis timeframe, we detected an increasing trend in the number of papers on AR. This is most likely because the technology is developing, and affordable systems are partially available in daily life. In most

cases, the first author was located in the United States, followed by Germany, Japan, and France.

Inspecting only the PubMed database and their MeSH annotations, the dominating terms were found to be Computer-Assisted Surgery, Three-Dimensional Imaging, Computed X-Ray Tomography, and Laparoscopy. All terms were connected to technology-intensive environments that make use of screens and advanced visualization technology. This finding was also observed in the analysis of the Medical Scope, as the results show that the main scope was Treatment, and within this area, primarily, the operating room.

With only three clinical trials registered in ClinicalTrials.gov, the scope for clinical trials was unexpected. One possible explanation for the low number of registered trials could be that it is not mandatory to register a medical invention prior to a clinical trial. Furthermore, most techniques were at a technology readiness level of 6 to 7; thus, they were mostly prototypes but needed additional completion and proving, a step that would lead to a clinical trial. We often observed that trials were conducted, but with a low number of subjects (N<9), or that the trials conducted were animal trials, which are not registered.

Only a very small subset of explored AR technologies featured a collaboration aspect. This could be due to the lack of widely available network coverage in sensitive environments or simply the size of AR technologies. Wearable technologies, in particular, are still too bulky to be used in a professional setting.

The maturity analysis of the research projects showed that the majority of publications were distributed in a rough bell-curve shape around level 6 and the use of technology was demonstrated in a relevant environment. This indicates that either early research is not published or basic research is no longer necessary, because AR technology is underway to production readiness.

From a medical perspective, the majority of papers were categorized under “Treatment.” This is due to the fact that the use of displays is already established in operating rooms, and this is a small step toward adding information to the images shown.

This finding is additionally supported by the next categorization; the technological perspective showed that the majority of publications were written under the stream “World,” and in these cases, the majority were in the class “Display.”

The training scope is dominated by surgery simulations. Since the technology is easy to apply in a training environment, this is an expected finding.

The second most important role after the presentation of the augmented world is the tracking of real-life objects that are used in the augmented space. Most of the tracking mechanisms used “marker-based” optical tracking, a technique that, in most cases, is more robust and accurate than “markerless” optical tracking. Only 39 publications did not use any kind of optical tracking, as it was not necessary for their apps. The majority of projects used a mobile tracking system. This can be explained by the broader range of movements provided by mobile tracking systems, which is essential in most cases, especially in surroundings like a hospital room.

Defining the different levels of augmentation based on the taxonomy by Hugues et al, majority of the apps were assigned to the class “Perceptual Association”; thus, most apps seek to make associations between the real and the virtual world, for example, the association between an organ seen in endoscopy and the visual image projected onto it. The second most observed definition was “Perceptual Understanding”; the virtual object draws a connection to what is seen and explains or shows certain objects in the real world.

Strength and Limitations

The strength of manually analyzing the publications is that data could be deducted, which is not explicitly written, and were therefore not detectable by automated text mining. For example, some articles stated which device was used but did not disclose the tracking technology or which display was used explicitly. Since specifications for most devices are known, the basic tracking and display technologies were added to the result, as derived from context, for example, the publication by Shi et al [50] who presented an AR app in combination with a robot. The robot used fuzzy controlled mechanisms to help surgeons control a drill during a mandible plastic surgery; the surgeons can see additional information about the bone structure through AR glasses.

An additional strength of manual analysis is that whenever ambiguous terms were used, it could be compensated. If a system was introduced as AR but was, by the Milgram definition, a VR system, the paper was excluded. In contrast, publications were included even if they involved ambiguous use of the term “virtual,” as seen in the study by Lozano-Quilis [51], who proposed a “virtual rehabilitation” scenario, showing and explaining an AR environment.

One strength of this review, in contrast to that by Chen et al [2] who used only Scopus to retrieve published papers, is that this paper used two databases (PubMed and Scopus) to find the initial set of publications.

The number of analyzed publications is, in contrast to the study by Chen et al [2], a dataset limitation. Another limitation is that publications without an abstract and not containing AR in the title were not identified in our search. Because the publications were analyzed manually, there is a possibility of an individual bias of the analyzing author. To prevent this bias, the authors cross-checked 25% of the publications.

Conclusions

This work provides a detailed view into 5 years of AR research in medicine. AR in medicine is an emerging technology that can benefit medical practitioners, health care professionals, and patients. The assessment of the technology readiness level shows that the AR technology is beyond the testing phase, and practical applications are becoming more common.

MeSH term analysis showed that the fields of Computer-Assisted Surgery, Three-Dimensional Imaging, and Computed X-Ray Tomography are the most explored. These

fields already make use of advanced display technologies, and it is easy to integrate AR technologies into their workflow.

There was also a clear trend in technologies assisting actual treatment of patients in comparison to technologies in a training environment. In the treatment/training aspect, a more balanced distribution of publications was expected.

AR technology is an upcoming technology that will impact the treatment of patients in the future. With shrinking and more powerful hardware, the technology will be able to merge better into existing workflows and create opportunities for patients, doctors, and health care professionals.

The aim of this review was to offer a foundation to researchers, which was met by offering a categorized list. A certain scope of research can be searched from our list by using the interactive table in Multimedia Appendix 1, which can, for example, be filtered to fit a certain category or criteria.

In contrast to the study by Chen et al [2], the authors expanded the data source. Chen et al [2] only accessed the Scopus database, whereas we used PubMed and Scopus databases as a basis for analysis. Therefore, a more comprehensive dataset was created. Another advantage of manual analysis is the ability to analyze graphics and information that are not directly contained in the text. An analysis using text mining is limited to recognizing synonyms and structures and displaying them.

Acknowledgments

The authors would like to thank the reviewers for their constructive comments.

Authors' Contributions

JV and ME equally searched, filtered, and analyzed the publications. The study design was developed in collaboration with CF. CF critically revised the paper. All authors approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interactive Excel table dataset.

[XLSX File (Microsoft Excel File), 51KB - mhealth_v7i4e10967_app1.xlsx]

Multimedia Appendix 2

Publications before first exclusion, including reviews.

[XLSX File (Microsoft Excel File), 101KB - mhealth_v7i4e10967_app2.xlsx]

Multimedia Appendix 3

Dataset of the complete reference list.

[TXT File, 605KB - mhealth_v7i4e10967_app3.txt]

References

1. Kamphuis C, Barsom E, Schijven M, Christoph N. Augmented reality in medical education? Perspect Med Educ 2014 Sep;3(4):300-311 [FREE Full text] [doi: 10.1007/s40037-013-0107-7] [Medline: 24464832]
2. Chen L, Day TW, Tang W, John NW. Recent Developments and Future Challenges in Medical Mixed Reality. : IEEE; 2017 Presented at: IEEE International Symposium on Mixed and Augmented Reality (ISMAR); 2017-10-09; Nantes, France p. 123-135.
3. Scopus. 2017. URL: https://www.scopus.com [accessed 2019-03-22] [WebCite Cache ID 773raipHC]
4. Hugues O, Fuchs P, Nannipieri O. New Augmented Reality Taxonomy: Technologies Features of Augmented Environment. In: Handbook of augmented reality. New York: Springer; 2011:47-63.
5. Schmalstieg D, Hollerer T. Augmented reality: Principles and practice, 1st edition. Boston: Addison-Wesley; 2016.
6. Aukstakalnis S. Practical Augmented Reality: A Guide to the Technologies, Applications, and Human Factors for AR and VR, 1st edition. Boston: Addison Wesley Professional; 2016.
7. Department of Defense. Technology Readiness Assessment (TRA) Guidance URL: http://www.webcitation.org/6uPYmlBWo [accessed 2017-10-20] [WebCite Cache ID 6uPYmlBWo]
8. Carmigniani J, Furht B. Augmented Reality: An Overview. In: Handbook of augmented reality. New York: Springer; 2011:3-46.

9. Azuma RT. A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments 1997 Aug;6(4):355-385. [doi: 10.1162/pres.1997.6.4.355]
10. Boud AC, Haniff DJ, Baber C, Steiner SJ. Virtual reality and augmented reality as a training tool for assembly tasks. 1999 Presented at: IEEE International Conference on Information Visualization, an International Conference on Computer Visualization & Graphics; July 14, 1999; London, England p. 32-36.
11. Milgram P, Takemura H, Utsumi A, Kishino F. Augmented reality: a class of displays on the reality-virtuality continuum. In: Telemanipulator and Telepresence Technologies. Bellingham, WA: SPIE; 1995:282-292.
12. Sutherland IE. A head-mounted three dimensional display. New York, USA: ACM Press; 1968 Presented at: Proceedings of the Fall Joint Computer Conference, Part I; Dec 9, 1968; New York p. 757. [doi: 10.1145/1476589.1476686]
13. Lee K. Augmented Reality in Education and Training. TechTrends 2012 Feb 7;56(2):13-21. [doi: 10.1007/s11528-012-0559-3]
14. Caudell TP, Mizell DW. Augmented reality: An application of heads-up display technology to manual manufacturing processes. 1992 Presented at: International Conference on System Sciences; 1992; Hawai p. 659-669. [doi: 10.1109/HICSS.1992.183317]
15. PubMed. 2017. URL: https://www.pubmed.gov [accessed 2017-10-20] [WebCite Cache ID 6uPZ3W1PK]
16. R Core Team. R. Vienna, Austria: R Foundation for Statistical Computing; 2011. The R Project for Statistical Computing URL: https://www.r-project.org [accessed 2017-10-20] [WebCite Cache ID 76dw9RZNh]
17. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
18. Barkley JE, Lepp A, Glickman EL. Pokèmon Go! May Promote Walking, Discourage Sedentary Behavior in College Students. Games Health J 2017 Jun;6(3):165-170. [doi: 10.1089/g4h.2017.0009] [Medline: 28628384]
19. Yoo JW, Lee DR, Cha YJ, You SH. Augmented effects of EMG biofeedback interfaced with virtual reality on neuromuscular control and movement coordination during reaching in children with cerebral palsy. NeuroRehabilitation 2017;40(2):175-185. [doi: 10.3233/NRE-161402] [Medline: 28222541]
20. Sutton EE, Demir A, Stamper SA, Fortune ES, Cowan NJ. Dynamic modulation of visual and electrosensory gains for locomotor control. J R Soc Interface 2016 Dec;13(118) [FREE Full text] [doi: 10.1098/rsif.2016.0057] [Medline: 27170650]
21. NIH - U.S. National Library of Medicine. Medical Subject Headings URL: https://www.nlm.nih.gov/mesh/ [accessed 2018-04-27] [WebCite Cache ID 6yzK1T0ds]
22. NIH - U.S. National Library of Medicine - ClinicalTrials.gov. 2017. URL: https://www.clinicaltrials.gov [accessed 2017-10-22] [WebCite Cache ID 6uPX7DjzK]
23. Khor WS, Baker B, Amin K, Chan A, Patel K, Wong J. Augmented and virtual reality in surgery-the digital surgical environment: applications, limitations and legal pitfalls. Ann Transl Med 2016 Dec;4(23):454 [FREE Full text] [doi: 10.21037/atm.2016.12.23] [Medline: 28090510]
24. Thomas DJ. Augmented reality in surgery: The Computer-Aided Medicine revolution. Int J Surg 2016 Dec;36(Pt A):25. [doi: 10.1016/j.ijsu.2016.10.003] [Medline: 27741424]
25. Moglia A, Ferrari V, Morelli L, Ferrari M, Mosca F, Cuschieri A. A Systematic Review of Virtual Reality Simulators for Robot-assisted Surgery. Eur Urol 2016 Dec;69(6):1065-1080. [doi: 10.1016/j.eururo.2015.09.021] [Medline: 26433570]
26. Diana M, Marescaux J. Robotic surgery. Br J Surg 2015 Jan;102(2):e15-e28. [doi: 10.1002/bjs.9711] [Medline: 25627128]
27. Slade SJA, Veillette CJH. The application of wearable technology in surgery: ensuring the positive impact of the wearable revolution on surgical patients. Front Surg 2014;1:39 [FREE Full text] [doi: 10.3389/fsurg.2014.00039] [Medline: 25593963]
28. Bluemel C, Matthies P, Herrmann K, Povoski SP. 3D scintigraphic imaging and navigation in radioguided surgery: freehand SPECT technology and its clinical applications. Expert Rev Med Devices 2016;13(4):339-351. [doi: 10.1586/17434440.2016.1154456] [Medline: 26878667]
29. Pelargos PE, Nagasawa DT, Lagman C, Tenn S, Demos JV, Lee SJ, et al. Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. J Clin Neurosci 2017 Jan;35:1-4. [doi: 10.1016/j.jocn.2016.09.002] [Medline: 28137372]
30. Meola A, Cutolo F, Carbone M, Cagnazzo F, Ferrari M, Ferrari V. Augmented reality in neurosurgery: a systematic review. Neurosurg Rev 2017 Oct;40(4):537-548. [doi: 10.1007/s10143-016-0732-9] [Medline: 27154018]
31. Marcus HJ, Pratt P, Hughes-Hallett A, Cundy TP, Marcus AP, Yang G, et al. Comparative effectiveness and safety of image guidance systems in surgery: a preclinical randomised study. Lancet 2015 Feb 26;385 Suppl 1:S64. [doi: 10.1016/S0140-6736(15)60379-8] [Medline: 26312886]
32. Kearns JT, Gundeti MS. Pediatric robotic urologic surgery-2014. J Indian Assoc Pediatr Surg 2014 Jul;19(3):123-128 [FREE Full text] [doi: 10.4103/0971-9261.136456] [Medline: 25197187]
33. Hamacher A, Kim SJ, Cho ST, Pardeshi S, Lee SH, Eun S, et al. Application of Virtual, Augmented, and Mixed Reality to Urology. Int Neurourol J 2016 Sep;20(3):172-181 [FREE Full text] [doi: 10.5213/inj.1632714.357] [Medline: 27706017]
34. Mahmud N, Cohen J, Tsourides K, Berzin TM. Computer vision and augmented reality in gastrointestinal endoscopy. Gastroenterol Rep (Oxf) 2015 Aug;3(3):179-184 [FREE Full text] [doi: 10.1093/gastro/gov027] [Medline: 26133175]

35. Feussner H, Fiolka A, Schneider A, Cuntz T, Coy J, von Tiesenhausen C, et al. The "Iceberg Phenomenon": As Soon as One Technological Problem in NOTES Is Solved, the Next One Appears!. Surg Innov 2015 Dec;22(6):643-650. [doi: 10.1177/1553350615573578] [Medline: 25733547]
36. Smith ML, Foley MR. Transforming Clinical Education in Obstetrics and Gynecology: Gone Is the Day of the Sage on the Stage. Obstet Gynecol 2016 Apr;127(4):763-767. [doi: 10.1097/AOG.0000000000001356] [Medline: 26959215]
37. Ortiz-Catalan M, Guðmundsdóttir RA, Kristoffersen MB, Zepeda-Echavarria A, Caine-Winterberger K, Kulbacka-Ortiz K, et al. Phantom motor execution facilitated by machine learning and augmented reality as treatment for phantom limb pain: a single group, clinical trial in patients with chronic intractable phantom limb pain. Lancet 2016 Dec 10;388(10062):2885-2894. [doi: 10.1016/S0140-6736(16)31598-7] [Medline: 27916234]
38. Vera AM, Russo M, Mohsin A, Tsuda S. Augmented reality telementoring (ART) platform: a randomized controlled trial to assess the efficacy of a new surgical education technology. Surg Endosc 2014 Dec;28(12):3467-3472. [doi: 10.1007/s00464-014-3625-4] [Medline: 24962856]
39. Chaballout B, Molloy M, Vaughn J, Brisson IR, Shaw R. Feasibility of Augmented Reality in Clinical Simulations: Using Google Glass With Manikins. JMIR Med Educ 2016 Mar 07;2(1):e2 [FREE Full text] [doi: 10.2196/mededu.5159] [Medline: 27731862]
40. Kranzfelder M, Wilhelm D, Doundoulakis M, Schneider A, Bauer M, Reiser S, et al. A probe-based electromagnetic navigation system to integrate computed tomography during upper gastrointestinal endoscopy. Endoscopy 2014 Apr;46(4):302-305. [doi: 10.1055/s-0033-1358814] [Medline: 24254384]
41. Carenzo L, Barra FL, Ingrassia PL, Colombo D, Costa A, Della CF. Disaster medicine through Google Glass. Eur J Emerg Med 2015 Jun;22(3):222-225. [doi: 10.1097/MEJ.0000000000000229] [Medline: 25460812]
42. Deng W, Li F, Wang M, Song Z. Easy-to-use augmented reality neuronavigation using a wireless tablet PC. Stereotact Funct Neurosurg 2014;92(1):17-24. [doi: 10.1159/000354816] [Medline: 24216673]
43. Albrecht U, Folta-Schoofs K, Behrends M, von Jan U. Effects of mobile augmented reality learning compared to textbook learning on medical students: randomized controlled pilot study. J Med Internet Res 2013;15(8):e182 [FREE Full text] [doi: 10.2196/jmir.2497] [Medline: 23963306]
44. Edgcumbe P, Pratt P, Yang G, Nguan C, Rohling R. Pico Lantern: Surface reconstruction and augmented reality in laparoscopic surgery using a pick-up laser projector. Med Image Anal 2015 Oct;25(1):95-102. [doi: 10.1016/j.media.2015.04.008] [Medline: 26024818]
45. Chinthammit W, Merritt T, Pedersen S, Williams A, Visentin D, Rowe R, et al. Ghostman: augmented reality application for telerehabilitation and remote instruction of a novel motor skill. Biomed Res Int 2014;2014:646347 [FREE Full text] [doi: 10.1155/2014/646347] [Medline: 24829910]
46. KleinJan GH, van den Berg NS, van Oosterom MN, Wendler T, Miwa M, Bex A, et al. Toward (Hybrid) Navigation of a Fluorescence Camera in an Open Surgery Setting. J Nucl Med 2016 Oct;57(10):1650-1653 [FREE Full text] [doi: 10.2967/jnumed.115.171645] [Medline: 27230927]
47. Marker DR, Thainual PU, Ungi T, Flammang AJ, Fichtinger G, Iordachita II, et al. 1.5 T augmented reality navigated interventional MRI: paravertebral sympathetic plexus injections. Diagn Interv Radiol 2017;23(3):227-232 [FREE Full text] [doi: 10.5152/dir.2017.16323] [Medline: 28420598]
48. Olivieri E, Barresi G, Mattos LS. BCI-based user training in surgical robotics. : IEEE; 2015 Aug Presented at: Proc IEEE Eng Med Biol Soc; Aug 25, 2015; Milan, Italy p. 4918-4921. [doi: 10.1109/EMBC.2015.7319495]
49. Wilson KL, Doswell JT, Fashola OS, Debeatham W, Darko N, Walker TM, et al. Using augmented reality as a clinical support tool to assist combat medics in the treatment of tension pneumothoraces. Mil Med 2013 Sep;178(9):981-985. [doi: 10.7205/MILMED-D-13-00074] [Medline: 24005547]
50. Shi Y, Lin L, Zhou C, Zhu M, Xie L, Chai G. A study of an assisting robot for mandible plastic surgery based on augmented reality. Minim Invasive Ther Allied Technol 2017 Feb;26(1):23-30. [doi: 10.1080/13645706.2016.1216864] [Medline: 27564567]
51. Lozano-Quilis JA, Gil-Gómez H, Gil-Gómez JA, Albiol-Pérez S, Palacios-Navarro G, Fardoun HM, et al. Virtual rehabilitation for multiple sclerosis using a kinect-based system: randomized controlled trial. JMIR Serious Games 2014;2(2):e12 [FREE Full text] [doi: 10.2196/games.2933] [Medline: 25654242]

Abbreviations

AR: augmented reality
AV: augmented virtuality
HMD: head-mounted display
MeSH: Medical Subject Headings
MR: mixed reality
SPECT: single-photon emission computed tomography
VR: virtual reality

Edited by G Eysenbach; submitted 04.05.18; peer-reviewed by N Ribeiro, W Chinthammit, A Ahmadvand, A Davoudi, D Nault; comments to author 07.08.18; revised version received 02.01.19; accepted 04.03.19; published 26.04.19.

Please cite as:
Eckert M, Volmerg JS, Friedrich CM
Augmented Reality in Medicine: Systematic and Bibliographic Review
JMIR Mhealth Uhealth 2019;7(4):e10967
URL: http://mhealth.jmir.org/2019/4/e10967/
doi:10.2196/10967
PMID:31025950

Original Paper

Increasing Completion Rate and Benefits of Checklists: Prospective Evaluation of Surgical Safety Checklists With Smart Glasses

Thomas Boillat[1*], PhD; Peter Grantcharov[2*], BSc; Homero Rivas[3], MBA, MD

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

[2]Data Science Institute, Columbia University, New York, NY, United States

[3]Design Lab, College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates

[*]these authors contributed equally

Corresponding Author:
Homero Rivas, MBA, MD
Design Lab
College of Medicine
Mohammed Bin Rashid University of Medicine and Health Sciences
PO Box 505055
Dubai,
United Arab Emirates
Phone: 971 4 383 8746
Email: homero.rivas@mbru.ac.ae

Abstract

Background: Studies have demonstrated that surgical safety checklists (SSCs) can significantly reduce surgical complications and mortality rates. Such lists rely on traditional posters or paper, and their contents are generic regarding the type of surgery being performed. SSC completion rates and uniformity of content have been reported as modest and widely variable.

Objective: This study aimed to investigate the feasibility and potential of using smart glasses in the operating room to increase the benefits of SSCs by improving usability through contextualized content and, ideally, resulting in improved completion rates.

Methods: We prospectively evaluated and compared 80 preoperative time-out events with SSCs at a major academic medical center between June 2016 and February 2017. Participants were assigned to either a conventional checklist approach (poster, memory, or both) or a smart glasses app running on Google Glass.

Results: Four different surgeons conducted 41 checklists using conventional methods (ie, memory or poster) and 39 using the smart glasses app. The average checklist completion rate using conventional methods was 76%. Smart glasses allowed a completion rate of up to 100% with a decrease in average checklist duration of 18%.

Conclusions: Compared with alternatives such as posters, paper, and memory, smart glasses checklists are easier to use and follow. The glasses allowed surgeons to use contextualized time-out checklists, which increased the completion rate to 100% and reduced the checklist execution time and time required to prepare the equipment during surgical cases.

(JMIR Mhealth Uhealth 2019;7(4):e13447) doi:10.2196/13447

KEYWORDS

smart glasses; surgical safety checklists; surgery; usability; time-out event

Introduction

Background

Reducing complications and deaths in operating rooms (ORs) due to human error is a big challenge for hospitals. To explain variability in surgical outcomes, studies have primarily focused on patient pathophysiological risk factors and surgeon skills [1]. Hence, when the patient did not account for the surgical complication, the error was then attributed to the surgeon's aptitudes and capabilities [2]. However, more recent research shows that errors "arise not from the solitary actions of individuals but from conflicting, incomplete, or suboptimal systems [3]." These systems notably refer to the people involved

in surgical cases as well as the tasks, tools, or technologies; environment; and organization (eg, hospitals or clinics).

To analyze and improve patient safety, the US Institute of Medicine and the National Academy of Engineering have promoted the use of human factor techniques. Human factor techniques investigate factors and develop tools that facilitate the achievement of goals (eg, reduce errors, increase productivity, improve safety) [4]. In this view, among many different initiatives, the World Health Organization (WHO) has collected scientific evidence and published guidelines to address part of the problems with safety of surgical patients [5]. These guidelines are summarized as a 19-item checklist that aims to establish systematic verifications before anesthesia, before surgery, and after surgery [6]. Existing research demonstrates that when systematically applied, surgical safety checklists (SSCs) can reduce complications and mortality from 19.9% to 11.5% and 1.6% to 1.0%, respectively [7,8].

However, at the expense of patient safety, SSCs have not been entirely adopted by hospitals. In the United Kingdom and France, where the use of the SSC is mandatory, the average SSC completion rate is 60% [9,10]. Such low completion rates are notably explained by the OR's constraints and bad design of the SSC implementations (eg, the checklists are often formatted by administrative staff who do not have the required skills; as a result, checklists are often difficult to read due to inappropriate fonts and colors) [4].

This is particularly the case with the time-out checklist. Because it takes place right before surgery commences and surgeons are not supposed to leave the 30 cm (12-inch) sterile area around the surgical table [11], they have to rely on a poster on the OR wall, often far from their field of view [9]. Additionally, surgeons complain that time-out checklists are not specific enough and often require them to spend time verifying irrelevant things. Finally, checklists are sometimes obsolete, and checklist completion is often not documented in the patient's medical record. Consequently, time-out checklists often do not bring enough benefits to surgeons and, thus, are not systematically used [12,13].

In other fields where checklists are heavily used, information technologies such as mobile devices are very often solicited to improve checklist execution [14]. Existing literature shows that information technology can enhance checklist support by reducing human error and increasing safety [15]. Given that surgeons must not touch nonsterile equipment, the use of mobile devices in not optimal. However, smart glasses, which have recently become available on the market, present an interesting alternative, and surgeons have already begun to investigate their potential. While recent research has demonstrated the benefits of using smart glasses in ORs [16-18], no study has empirically investigated their use to execute checklists.

With this study, we investigate the following: *Are smart glasses a potential technology to use to execute SSCs?* Given that surgeons complain about the rigidity of traditional checklists, we also aim to evaluate the following: *How can smart glasses bring more benefit to SSCs?* To answer these two questions, the authors designed and evaluated a checklist app for smart glasses that was implemented over 6 months. Our results demonstrate that time-out checklists executed on smart glasses are easy to read, follow, and execute. When contextualized for a specific surgery, smart glasses can increase the time-out checklist completion rate to 100% while saving time in execution and preparation.

Existing Research: Smart Glasses as Candidates for Use During Surgical Safety Checklists

Smart glasses are a wearable technology that uses spectacle frames to display contextualized information in a person's field of view [19]. The main piece of hardware is a head-mounted display that allows the user to access texts, pictures, and videos. The glasses are also equipped with a high-definition front-end camera, touchpad, and microphone as well as a series of sensors (eg, accelerometer, gyroscope, Global Positioning System) [20]. Smart glasses are either connected to a mobile device (eg, mobile phone, tablet) or Wi-Fi network that enables access to the internet or a company's information system. Although the use of a head-mounted display was evaluated by anesthesiologists in ORs to display vital signs more than 20 years ago [21], smart glasses eventually drew the attention of surgeons again in 2013 when Google released Google Glass. Due to the position of the camera next to the surgeon's eye, it captures what the surgeon sees [22,23]. Within weeks of its distribution, Google Glass was being evaluated by surgeons in live-stream surgeries and was used to obtain advice from experts several thousand miles away. This device not only presents opportunities for medical students to visualize surgeries comfortably [16] but also for supervisors to evaluate junior surgeons [24]. Built on two recent literature reviews on the use of Google Glass in medicine [18,25], Table 1 summarizes and classifies a list of studies that took place in nonsimulated operative surgical settings.

Since 2013, many studies have investigated and demonstrated the ability of smart glasses to support surgeons in enabling remote diagnosis and assistance, documenting cases via photos or videos, and accessing patient information such as x-rays and vital signs, among other uses. However, while smart glasses are often cited in medical articles and scientific studies as potential candidates to address existing shortcomings with traditional checklist executions (ie, mostly the completion rate) [38,39], to our knowledge no research has empirically evaluated the use of smart glasses to execute checklists in a live OR setting.

Table 1. Existing research on the uses of smart glasses in health care.

Reference	Study design	Purpose of smart glasses use	Remote diagnosis and assistance	Video documentation	Photo documentation	Access patient information	(Self)-guidance
[26]	Pilot	Evaluate capacity of smart glasses to enhance communication, document cases, and access patient information	X		X	X	
[27]	Feasibility	Evaluate capacity of smart glasses to enhance communication and document cases	X	X	X		
[28]	Pilot	Use of head-mounted display on smart glasses to display vital sign parameters				X	
[29]	Pilot	List of smart glasses opportunities (eg, access to online medical encyclopedia, patient information, documentation, remote assistance)	X	X	X		
[30]	Case	Document cases and analyze pictures			X		
[31]	Pilot	Record first-person point-of-view video and photos and use as search engine		X	X	X	
[32]	Pilot	Live-stream video during surgery and facilitate remote telementoring between 2 surgeons, allowing real-time guidance of the operating surgeon	X				X
[33]	Feasibility	Assess the safety of using Google Glass by assessing the video quality of a telementoring session		X			
[34]	Pilot	Facilitate real-time observation and proctoring by mentoring surgeon experts in remote locations around the world	X				
[35]	Pilot	Enhance neuronavigation by projecting images directly on the Google Glass screen instead of traditional screens				X	X[a]
[36]	Pilot	Evaluate the use of Google Glass to document airway assessment and tracheal intubation			X		
[37]	Randomized controlled	Evaluate whether Google Glass can be used to perform an ultrasound-guided procedure					X

[a]Self-guidance was the secondary and not primary goal in surgery.

Methods

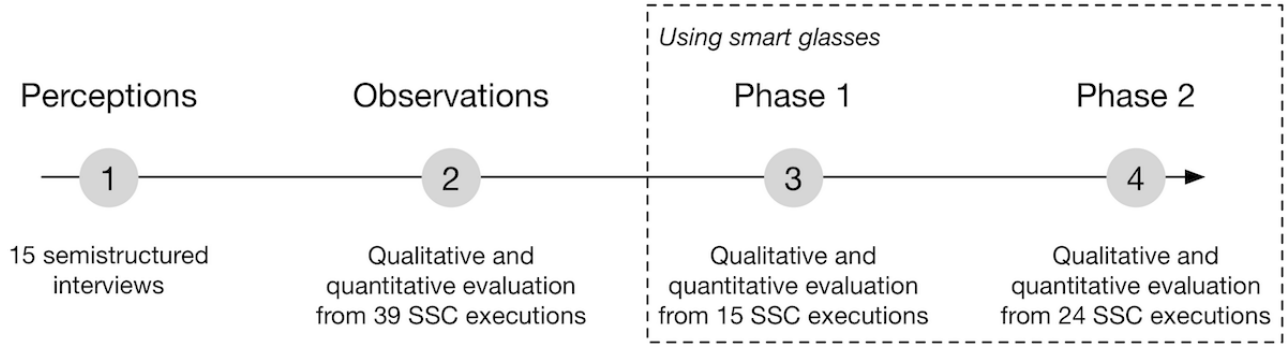
Context

We prospectively evaluated and compared 80 preoperative time-out events with SSCs at a major academic medical center between June 2016 and February 2017. Participants were assigned to either a conventional checklist approach (poster, memory, or both) or a smart glasses app (Google Glass). All surgical cases investigated were elective and gastrointestinal in nature. The hospital implemented the WHO SSC in 2009. This time-out checklist counts 13 items that must be verified by the surgeon responsible for the case; 5 additional items are used only when blood transfusion is part of the surgical procedure (which did not take place in this study). The checklist was customized according to the needs of the hospital and the different surgical departments. To ensure appropriate execution of the checklists, hospital checklist reporters attend random checklist executions (the checklist execution of a surgeon is monitored 1 to 4 times a month) and report their observations to the hospital's administrators. To support surgeons in their checklist executions, the hospital has equipped its ORs with wall posters of the customized WHO checklist. Alternatively, surgeons can use a paper-based version of the checklist with the help of a circulating nurse.

Checklist App for Google Glass

To answer our research questions, we iteratively designed a smart glasses app following the action design research methodology [40], which promotes the involvement of end users in the design of the solution to ensure its efficiency and usability. As described below and shown in Figure 1, end users were involved from the needs definition to the evaluation of the smart glasses checklist app.

Figure 1. Data collection throughout the research process. SSC: surgical safety checklist.



Our checklist app contained two screens: one to select the checklist and one to display the items of the checklist selected. In addition, we developed a checklist engine that allows for creating and maintaining individualized checklists. While checklist apps for smart glasses exist on the market, none of them provides the flexibility required for this study.

The app ran on smart glasses developed by Google, chosen for their popularity among surgeons and their versatile characteristics, such as a small screen that does not obstruct the surgeon's field of view and, thus, does not disrupt communication among the OR staff. Additionally, they are very light and can work offline without any network connection.

Approach

The research began with interviews of 15 surgeons to gauge their checklist execution experience and understand their challenges with traditional mediums. Data from the interviews also allowed us to compare their outcomes with the existing literature [9,12]. We then observed 41 conventional checklist executions in the OR, focusing on the way surgeons executed the checklists, completion rate, level of interaction with the OR staff, and duration. Leveraging the interviews and observation outcomes, we iteratively designed and refined a smart glasses app. When a new functionality of the app was ready, two surgeons were asked to evaluate it in a simulation room and make suggestions for refinements if the usability was too low (eg, difficult to read the items or navigate). Once the app met the requirements of the surgeons, it was evaluated in the ORs in two phases (Figure 1).

In phase 1, we evaluated the capacity of smart glasses to assist in executing the checklists, which involved only small wording adaptations in comparison with the official checklist (ie, poster). For phase 2, we collaborated with surgeons to customize surgery-specific checklists, following the recommendations by Weiser et al [41], in order to evaluate the benefits of having all the checklist items relevant for a set of surgeries. While it was clear which checklist items would be removed, it was not clear which new contextualized checklist items were appropriate for each checklist. Thus, it required several surgical cases to determine the relevant level of abstraction before including new checklist items. Surgeons also added items to ensure the readiness of specific equipment that is more likely not to be ready. Prior to the implementation of the smart glasses checklists, a 5-minute training was provided (only one surgeon had previous experience with Google Glass).

To evaluate the usability of smart glasses, we opted for a quantitative and qualitative research design. Using semistructured interviews, we were able to gather details about the ease of reading, following, and navigating through the checklist using smart glasses and modify the app based on surgeon feedback. We used descriptive statistics methods to evaluate the efficiency of the app.

Data Collection

Data were collected at four distinct points throughout the study as shown in Figure 1. First, data from semistructured interviews were used to understand surgeon perceptions of checklist roles, benefits, drawbacks, and requirements. In total, 9 surgeons and 6 residents (at least postgraduate year 3) took part in the interviews. Second, during the observation phase, the authors attended and documented the execution of 41 conventional time-out checklists in ORs performed by 3 surgeons and 3 residents who also took part in the interviews. The documentation included the sequence and items verified by the surgeons, medium used for the checklist execution, level of interaction with OR staff (ie, low, medium, high), and duration. Third, in phase 1, the smart glasses app was used by 2 surgeons and 2 residents in the OR for a total of 15 surgical cases. After each case, the authors conducted a semistructured interview with the surgeon that focused on the ease of reading, following, and validating time-out items on the smart glasses, as well as the usefulness of the glasses to document the checklist executions and ensure their completeness. Each question was rated on a 4-point Likert scale and included a free-text field used to document additional comments. Last, in phase 2, surgeons used contextualized verifications by means of the smart glasses app in the OR for 24 surgical cases. The same evaluation process was used as in phase 1.

Results

Interview Outcomes (Perceptions) and Observation Results

The 15 interviewed surgeons agreed on the importance of time-out checklists to unite the OR staff and establish a common ground regarding the surgical procedure. However, as highlighted by existing studies, the surgeons found the current time-out mediums (ie, poster and paper) to be too generic, thus limiting their benefits. Similarly, the majority of interviewees found the mediums used to support the time-out execution difficult to read and follow. Therefore, surgeons often created

their own and amended checklists that they could execute from memory. Alternatively, some surgeons started the time-out execution by memory and then used the poster to ensure they did not forget any verification.

Usability of Smart Glasses to Execute Surgical Safety Checklists (Phase 1)

To answer the primary research question of whether smart glasses are a viable technology to execute time-out checklists, we analyzed the usability of smart glasses. We focused on the ease of reading, following, and navigating through a time-out checklist.

Ease of Reading Surgical Safety Checklist Content

During the interviews, the majority of surgeons mentioned that the checklist wall poster is difficult to read and follow. With smart glasses, the checklist items appear in front of the surgeon's eye. Although the Google Glass screen appears small, it renders a picture equivalent to a 25-inch high-definition television sitting 2.4 meters (8 feet) away. To enhance readability, only one checklist item is displayed in white on a black background, as shown in Figure 2. Participants strongly agreed that items on smart glasses are easy to read.

Ease of Following and Navigating Through the Surgical Safety Checklist

We observed on multiple occasions that surgeons lost their place during their time-out execution and had to cease the execution in order to find the next item on the poster. This issue was even more severe when surgeons commenced the checklist by memory and then forget their place. With the smart glasses, time-out items are displayed sequentially, requiring surgeons to go through all of the verifications. To mitigate this inflexibility, we offered surgeons the ability to create customized checklists by adding, removing, or editing any steps from those included in the basic checklist.

With regard to navigating within the checklist, voice commands were used in phase 1. The word "next" would trigger the next checklist item while "back" would return to the preceding item. The evaluation revealed 12 false negative events (ie, saying "next" with no change occurring). Tests revealed that the quality and sensitivity of the microphone were responsible for the low voice command recognition rate. In response, the interaction mode was changed to head gestures during the same phase 1 surgeries. The head gestures were the following: nodding up to down displayed the next checklist item and nodding down to up displayed the previous item. Finally, we decoupled the checklist items that originally contained multiple verifications. This was notably the case in the verification of the patient's identity, position, and procedure. During the observation phase, surgeons forgot to verify at least one of the three items on 13 occasions. Surgeons either strongly agreed or agreed that checklists on smart glasses are easy to follow and navigate.

Benefits of Smart Glasses to Execute Time-Out Checklists (Phase 2)

To investigate the benefits of smart glasses, we looked into their effect on the completion rate of time-outs and their ability to document checklist executions and contextualize the checklist contents.

Checklist Completion Rate

Existing studies largely demonstrate that incomplete checklist executions represent a critical problem that leads to complications in the OR [10,13]. As shown in Table 2, memory is the most common of all checklist execution means; however, it provides the lowest average completion rate (72%) and greatest average standard deviation (16%) across the execution (minimum completed items = 6; maximum = 13). It is interesting to note that surgeons executing time-outs by memory often missed different critical verifications despite the type of surgery being the same. When asked, no checklist executors realized they had forgotten items. By strictly following the poster, surgeons reached an average completion rate of 83% with the lowest average standard deviation (8%). However, critical verifications such as the identity and position of the patient and the use of deep vein thrombosis prophylaxis were sometimes forgotten. When surgeons relied on their memory and the poster, we observed they had significant difficulties in identifying the items they had already verified and those that remained. For this reason, they only performed slightly better than with memory, with an average completion rate of 77% and an average standard deviation of 10% (minimum completed items = 8; maximum = 13).

When surgeons used smart glasses in phase 1, the average completion rate increased to 98% with an average standard deviation of 3% (minimum completed items = 13; maximum = 14). Given that verifications appeared sequentially, the app forced surgeons to go through all the verifications. Surgeons did not complain about this when asked in the following interviews. We noticed that surgeons introduced themselves with their names but not their roles on three occasions in phase 1 (the only checklist item not performed at 100%). This behavior was also observed in the other mediums. When asked, surgeons said that they know the team and they do not want to repeat useless information.

Usefulness of Smart Glasses to Document Time-Out Executions

In our observation phase, we realized that time-out executions are not systematically documented. In our smart glasses app, we automatically document the time and checklist item each time the surgeon validates an item. Therefore, the time-out execution can theoretically be paired with the patient's medical record. Given that this technique does not account for additional items that could be verified during the checklist execution, the entire time-out execution is recorded via the smart glasses microphone and stored in the device. All the participants strongly agreed on the usefulness of documenting the SSC executions.

Figure 2. Checklist items displayed in the smart glasses app versus wall poster.

Table 2. Checklist completion rates by recall medium.

Checklist item	Recall medium			
	Memory (n=20)	Poster (n=12)	Memory + poster (n=8)	Smart glasses (phase 1; n=15)
Correct patient	20 (100)	11 (83)	8 (100)	15 (100)
Correct procedure	17 (85)	11 (92)	8 (100)	15 (100)
Correct position	17 (85)	12 (100)	4 (50)	15 (100)
Correct operative site/side	1 (5)	6 (50)	0 (0)	15 (100)
Consent completed, accurate, and signed	16 (80)	12 (100)	8 (100)	15 (100)
Surgical site marked by surgeon and visible after preparation/after drape	3 (15)	9 (75)	5 (62)	15 (100)
Confirmation of allergies	19 (100)	12 (100)	8 (100)	15 (100)
Images/implants available	12 (60)	9 (75)	3 (38)	15 (100)
Prophylactic antibiotic given	20 (100)	12 (100)	8 (100)	15 (100)
DVT[a] prophylaxis	18 (90)	10 (83)	8 (100)	15 (100)
Procedure duration	16 (80)	11 (92)	6 (75)	15 (100)
Any patient-specific concerns, are we all in agreement?	15 (75)	12 (100)	6 (75)	15 (100)
Is EBL[a]>500 cc or is there possibility of major blood loss?	18 (80)	10 (83)	8 (100)	15 (100)
Introduction by roles	11 (55)	1 (8)	6 (75)	11 (73)
Average	14.4 (72)	10 (83)	6.2 (77)	14.7 (98)
Standard deviation	3.2 (16)	1 (8)	0.8 (10)	0.45 (3)

[a]DVT: deep vein thrombosis.

[a]EBL: estimated blood loss.

Implementation of Contextualized Checklists to Increase Checklist Benefits

One of the main issues when using static mediums such as posters and papers lies in the inability of the checklist content to be adapted. While each surgical procedure type requires the verification of specific items, such as surgical phases and equipment, our interviews revealed the paper-style checklists could only provide very generic items. To evaluate the benefits of customized checklists, we developed five checklists that represented the most frequent surgery types we investigated in this research: laparoscopic sleeve gastrectomy, laparoscopic gastric bypass, esophagogastroduodenoscopy, per-oral endoscopic myotomy, and laparoscopic cholecystectomy. In addition, we numbered each of the verifications to indicate the progression in the checklist and moved the “introduction by role” to the first item, as suggested by WHO.

Figure 3. Completion rates across all mediums.

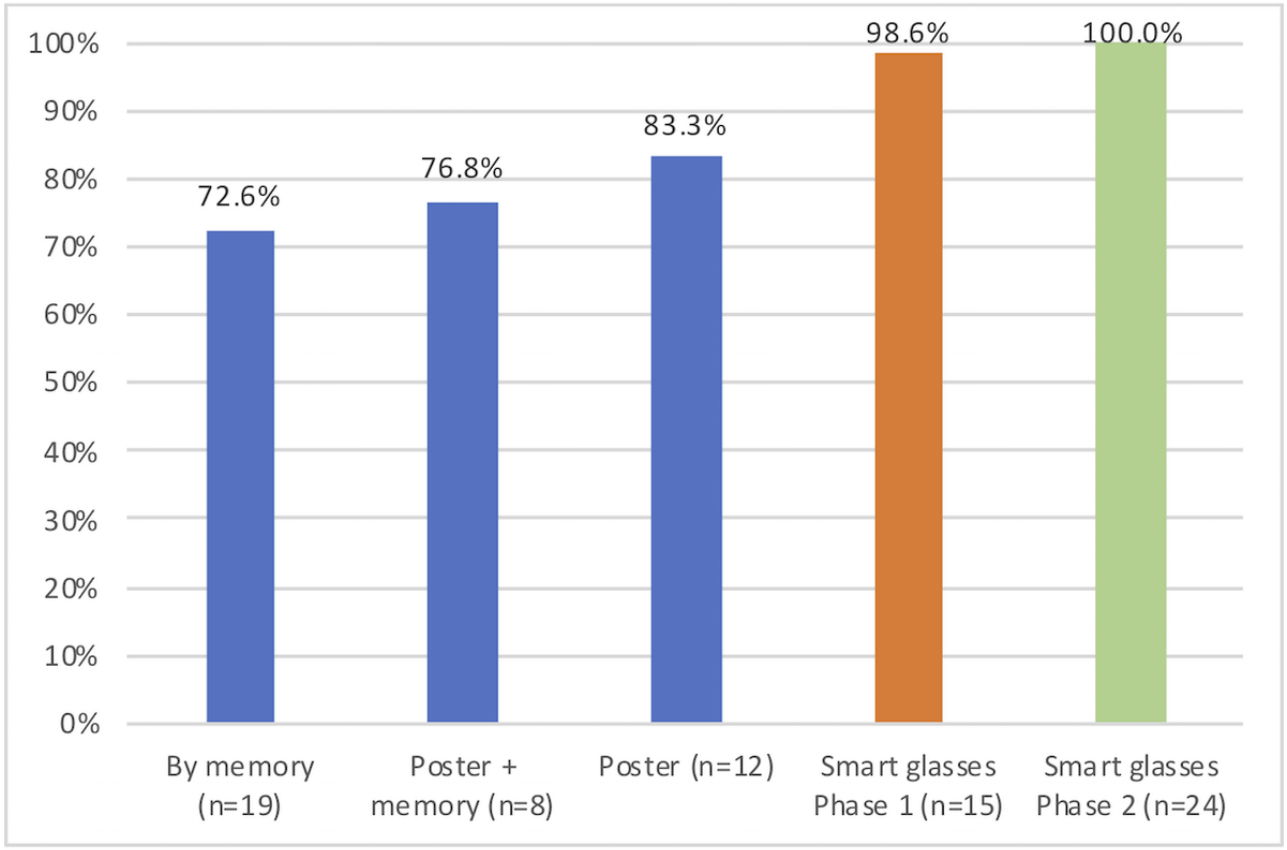


Table 3.

	Memory	Memory + poster	Poster	Smart glasses (phase 1)	Smart glasses (phase 2)
Average duration (seconds)	54.8	58.5	60.8	85.7	47.7
Standard deviation (seconds)	10.9	14.6	15.6	28.5	13.2

After 24 checklist executions with the smart glasses in phase 2, the completion rate was 100% as shown in Figure 3. Contextualized checklists also had an impact on the checklist duration given that only relevant checklist items were verified. Our evaluation revealed that contextualized checklists via smart glasses required less time compared with using the generic checklist, as shown in Table 3. It is not surprising to observe that executing the checklist for the first time with Google Glass required some adaptation time. However, making lists more relevant as they were contextualized resulted in decreased checklist duration and preserved high completion rates of all relevant items.

From the interviews, all participants strongly agreed that the use of contextualized checklists brings more efficiency to checklist executions.

Discussion

Principal Findings

In this research, we evaluated the use of smart glasses to improve the completion rate and, hence, the benefits of time-out checklist executions. In interviews, surgeons stressed the importance of differentiating between voluntary versus involuntary checklist noncompletion. The first is legitimate and happens when the checklist and the context are not aligned (eg, patient-specific items to verify). The latter occurs when surgeons unintentionally exclude items. Our evaluation confirms that smart glasses have the capability to address both types of noncompletion. They not only improve the completion rate but also increase the checklist's relevance and benefits. We observed that the checklist content differed by up to 40% across the selected checklists of this study. The implementation of smart glasses to contextualize SSCs allows deliberate list customization for specific surgeons and procedures, preventing the need for voluntary checklist noncompletion. At the same time, surgeons do not experience the need to improvise any additional checklist steps on the go as required when reading from a standardized list that does not adapt to real procedural circumstances. Furthermore, our results also show that the use of smart glasses was particularly relevant when no poster was available. Certainly, smart glasses had some hindrances; for example, we observed that they tended to isolate surgeons from the rest of the OR staff and that surgeons would move and interact less when using smart glasses. It is unclear whether this was a result of ergonomics, visual constraints, or just the awareness of wearing a novel head-mounted display.

Questions pertaining to implementation costs, maintenance, and use of such devices remain open. To the best of our knowledge, no hospitals have been using smart glasses at a large scale on a regular basis. Potential problems related to cleaning the device, overall reliability, and fragility are responsibilities that hospitals will soon have to accept. Beyond the US $1500 cost of each Google Glass headset, simple implementation decisions can also influence costs (eg, whether to adopt wired or wireless update and information transfer). While most ORs are equipped with wireless connectivity, for safety reasons, data must be encrypted, which increases the complexity. Shall the hospital equip each OR or each surgeon? If the former, the OR staff will be responsible for cleaning, updating, and ensuring that the device functions when the surgeon requires it. If the latter, the surgeon would be responsible for the device.

Limitations

This research had some limitations. There was no randomization in this study. While randomization could have been achieved fairly simply, there was limited availability of smart glasses and of some key members of the research team. Perhaps in the near future, these described limitations could be overcome with universal availability of this technological platform.

Conclusion

Our results show that smart glasses have the capacity to yield checklist completion rates of 100% by providing better usability than traditional mediums. Another benefit of using smart glasses lies in the use of contextualized checklist items. These allow surgeons to focus on context-relevant verifications while irrelevant checklist items are removed. In addition, smart glasses can be used to automatically document and transfer checklist executions to the patient's medical record.

Beyond smart glasses, this research also demonstrates the inefficiency of merging multiple verifications into one checklist item. Our observations show that most of the forgotten verifications are those that are merged with others.

This research is a first step toward clinically evaluating the efficiency of smart glasses in a long-term study. We encourage researchers and clinicians to further evaluate the use of smart glasses to execute checklists in surgery and other interventional procedures.

Acknowledgments

We would like to thank Sharon Wulfovich and Karolina Twardowska as well as Drs Nuria Martinez, Ioana Baiu, Dan Elison Azagury, and Matt Cooper for helping us evaluate the different prototypes. This research was partially funded by the Swiss National Science Foundation (P1LAP2-168407) and the Foundation of the University of Lausanne (Switzerland).

Conflicts of Interest

None declared.

References

1. Vincent C, Moorthy K, Sarker SK, Chang A, Darzi AW. Systems approaches to surgical quality and safety: from concept to measurement. Ann Surg 2004 Apr;239(4):475-482. [Medline: 15024308]
2. ElBardissi AW, Sundt TM. Human factors and operating room safety. Surg Clin North Am 2012 Feb;92(1):21-35. [doi: 10.1016/j.suc.2011.11.007] [Medline: 22269258]
3. Carayon P, Schoofs HA, Karsh B, Gurses AP, Alvarado CJ, Smith M, et al. Work system design for patient safety: the SEIPS model. Qual Saf Health Care 2006 Dec;15 Suppl 1:i50-i58 [FREE Full text] [doi: 10.1136/qshc.2005.015842] [Medline: 17142610]
4. Wickens C, Hollands J, Banbury S, Parasuraman R. Engineering Psychology and Human Performance. 4th Edition. London: Psychology Press; 2012.
5. WHO guidelines for safe surgery 2009: safe surgery saves lives.: World Health Organization URL: https://apps.who.int/iris/bitstream/handle/10665/44185/9789241598552_eng.pdf?sequence=1&isAllowed=y [accessed 2019-03-19]
6. WHO surgical safety checklist. 2008. URL: https://www.who.int/patientsafety/safesurgery/tools_resources/SSSL_Checklist_finalJun08.pdf [accessed 2019-03-19] [WebCite Cache ID 76zok8OYW]
7. Haugen AS, Søfteland E, Almeland SK, Sevdalis N, Vonen B, Eide GE, et al. Effect of the World Health Organization checklist on patient outcomes: a stepped wedge cluster randomized controlled trial. Ann Surg 2015 May;261(5):821-828. [doi: 10.1097/SLA.0000000000000716] [Medline: 24824415]
8. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AS, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med 2009 Jan 29;360(5):491-499. [doi: 10.1056/NEJMsa0810119] [Medline: 19144931]
9. Fourcade A, Blache J, Grenier C, Bourgain J, Minvielle E. Barriers to staff adoption of a surgical safety checklist. BMJ Qual Saf 2011 Nov 07;21(3):191-197. [doi: 10.1136/bmjqs-2011-000094] [Medline: 22069112]
10. Warnock G. Of surgeons and safety checklists. Can J Surg 2010;53(6):364-365. [Medline: 21092427]
11. AORN. Recommended practices for maintaining a sterile field. AORN J 2006;83(2):402-416. [Medline: 16544860]

12. Russ S, Sevdalis N, Moorthy K, Mayer E, Rout S, Caris J, et al. A qualitative evaluation of the barriers and facilitators toward implementation of the WHO Surgical Safety Checklist across hospitals in England: lessons from the Surgical Checklist Implementation Project. Ann Surg 2015;261(1):81-91. [Medline: 25072435]
13. Mayer E, Sevdalis N, Rout S, Caris J, Russ S, Mansell J, et al. Surgical checklist implementation project: the impact of variable WHO checklist compliance on risk-adjusted clinical outcomes after national implementation: a longitudinal study. Ann Surg 2015;263(1):58-63. [Medline: 25775063]
14. Clay-Williams R, Colligan L. Back to basics: checklists in aviation and healthcare. BMJ Qual Saf 2015;24(7):428-431. [Medline: 25969512]
15. Hales B, Pronovost P. The checklist—a tool for error management and performance improvement. J Crit Care 2006 Sep;21(3):231-235. [Medline: 16990087]
16. Davis CR, Rosenfield LK. Looking at plastic surgery through Google Glass: part 1. systematic review of Google Glass evidence and the first plastic surgical procedures. Plast Reconstr Surg 2015 Mar;135(3):918-928. [doi: 10.1097/PRS.0000000000001056] [Medline: 25719707]
17. Kolodzey L, Grantcharov PD, Rivas H, Schijven MP, Grantcharov TP. Wearable technology in the operating room: a systematic review. BMJ Innov 2016 Dec 21;3(1):55-63. [doi: 10.1136/bmjinnov-2016-000133]
18. Wei NJ, Dougherty B, Myers A, Badawy SM. Using Google Glass in surgical settings: systematic review. JMIR Mhealth Uhealth 2018 Mar 06;6(3):e54 [FREE Full text] [doi: 10.2196/mhealth.9409] [Medline: 29510969]
19. Starner T. Wearable computing: meeting the challenge. In: Barfield W, editor. Fundamentals of Wearable Computers and Augmented Reality. 2nd Edition. Boca Raton: CRC Press; 2015.
20. Vallurupalli S, Paydak H, Agarwal SK, Agrawal M, Assad-Kottner C. Wearable technology to improve education and patient outcomes in a cardiology fellowship program: a feasibility study. Health Technol 2013 Oct 11;3(4):267-270. [doi: 10.1007/s12553-013-0065-4]
21. Block JF, Yablok D, McDonald J. Clinical evaluation of the head-up display of anesthesia data. Int J Clin Monit Comput 1995;12(1):21-24. [Medline: 7782663]
22. Collman A. First ever surgery conducted by doctor wearing Google Glass.: Daily Mail; 2013. URL: https://www.dailymail.co.uk/news/article-2346442/First-surgery-conducted-doctor-wearing-Google-glass.html [accessed 2019-03-18] [WebCite Cache ID 76y8x7iVY]
23. Guillén P. Spanish doctor performs first surgery transmitted live via Google Glass. 2013 Sep 02. URL: http://www.kurzweilai.net/spanish-doctor-performs-first-surgery-transmitted-live-via-google-glass [accessed 2019-03-20] [WebCite Cache ID 770twzhE0]
24. Moshtaghi O, Kelley KS, Armstrong WB, Ghavami Y, Gu J, Djalilian HR. Using Google Glass to solve communication and surgical education challenges in the operating room. Laryngoscope 2015 Oct;125(10):2295-2297. [doi: 10.1002/lary.25249] [Medline: 25775933]
25. Dougherty B, Badawy SM. Using Google Glass in nonsurgical medical settings: systematic review. JMIR Mhealth Uhealth 2017 Oct 19;5(10):e159 [FREE Full text] [doi: 10.2196/mhealth.8671] [Medline: 29051136]
26. Armstrong DG, Rankin TM, Giovinco NA, Mills JL, Matsuoka Y. A heads-up display for diabetic limb salvage surgery: a view through the google looking glass. J Diabetes Sci Technol 2014 Sep;8(5):951-956 [FREE Full text] [doi: 10.1177/1932296814535561] [Medline: 24876445]
27. Hamann D, Mortensen WS, Hamann CR, Smith A, Martino B, Dameff C, et al. Experiences in adoption of teledermatology in Mohs micrographic surgery: using smartglasses for intraoperative consultation and defect triage. Surg Innov 2014 Dec;21(6):653-654. [doi: 10.1177/1553350614552735] [Medline: 25389145]
28. Liebert CA, Zayed MA, Aalami O, Tran J, Lau JN. Novel use of Google Glass for procedural wireless vital sign monitoring. Surg Innov 2016 Aug;23(4):366-373. [doi: 10.1177/1553350616630142] [Medline: 26848138]
29. Muensterer OJ, Lacher M, Zoeller C, Bronstein M, Kübler J. Google Glass in pediatric surgery: an exploratory study. Int J Surg 2014;12(4):281-289. [doi: 10.1016/j.ijsu.2014.02.003] [Medline: 24534776]
30. Widmer A, Schaer R, Markonis D, Muller H. Facilitating medical information search using Google Glass connected to a content-based medical image retrieval system. Conf Proc IEEE Eng Med Biol Soc 2014:4507-4510. [doi: 10.1109/EMBC.2014.6944625] [Medline: 25570993]
31. Borgmann H, Rodríguez SM, Salem J, Tsaur I, Gomez RJ, Barret E, et al. Feasibility and safety of augmented reality-assisted urological surgery using smartglass. World J Urol 2017 Jun;35(6):967-972. [doi: 10.1007/s00345-016-1956-6] [Medline: 27761715]
32. Ponce BA, Menendez ME, Oladeji LO, Fryberger CT, Dantuluri PK. Emerging technology in surgical education: combining real-time augmented reality and wearable computing devices. Orthopedics 2014 Nov;37(11):751-757. [doi: 10.3928/01477447-20141023-05] [Medline: 25361359]
33. Hashimoto DA, Phitayakorn R, Fernandez-del Castillo C, Meireles O. A blinded assessment of video quality in wearable technology for telementoring in open surgery: the Google Glass experience. Surg Endosc 2016 Jan;30(1):372-378. [doi: 10.1007/s00464-015-4178-x] [Medline: 25829065]
34. Datta N, MacQueen IT, Schroeder AD, Wilson JJ, Espinoza JC, Wagner JP, et al. Wearable technology for global surgical teleproctoring. J Surg Educ 2015;72(6):1290-1295. [doi: 10.1016/j.jsurg.2015.07.004] [Medline: 26276303]

35. Yoon JW, Chen RE, Han PK, Si P, Freeman WD, Pirris SM. Technical feasibility and safety of an intraoperative head-up display device during spine instrumentation. Int J Med Robot 2017 Sep;13(3). [doi: 10.1002/rcs.1770] [Medline: 27569781]
36. Spencer RJ, Chang PH, Guimaraes AR, Firth PG. The use of Google Glass for airway assessment and management. Paediatr Anaesth 2014 Sep;24(9):1009-1011. [doi: 10.1111/pan.12496] [Medline: 25039494]
37. Wu TS, Dameff CJ, Tully JL. Ultrasound-guided central venous access using Google Glass. J Emerg Med 2014 Dec;47(6):668-675. [doi: 10.1016/j.jemermed.2014.07.045] [Medline: 25281180]
38. Kolodzey L, Grantcharov P, Rivas H, Schijven M, Grantcharov T. Wearable technology in the operating room: a systematic review. BMJ Innov 2016 Dec 21;3(1):55-63 [FREE Full text] [doi: 10.1136/bmjinnov-2016-000133]
39. Mitrasinovic S, Camacho E, Trivedi N, Logan J, Campbell C, Zilinyi R, et al. Clinical and surgical applications of smart glasses. Technol Health Care 2015;23(4):381-401. [doi: 10.3233/THC-150910] [Medline: 26409906]
40. Sein M, Henfridsson O, Purao S, Rossi M, Lindgren R. Action design research. J MIS Quarterly 2011;35(1):37-56.
41. Weiser TG, Haynes AB, Lashoher A, Dziekan G, Boorman DJ, Berry WR, et al. Perspectives in quality: designing the WHO Surgical Safety Checklist. Int J Qual Health Care 2010 Oct;22(5):365-370. [doi: 10.1093/intqhc/mzq039] [Medline: 20702569]

Abbreviations

OR: operating room
SSC: surgical safety checklist
WHO: World Health Organization

Edited by G Eysenbach; submitted 20.01.19; peer-reviewed by K Helling, E Varela, J Nosta; comments to author 27.02.19; revised version received 07.03.19; accepted 07.03.19; published 29.04.19.

Please cite as:
Boillat T, Grantcharov P, Rivas H
Increasing Completion Rate and Benefits of Checklists: Prospective Evaluation of Surgical Safety Checklists With Smart Glasses
JMIR Mhealth Uhealth 2019;7(4):e13447
URL: http://mhealth.jmir.org/2019/4/e13447/
doi:10.2196/13447
PMID:31033451

Corrigenda and Addenda

Author Correction: The Impact of Text Messaging on Medication Adherence and Exercise Among Postmyocardial Infarction Patients: Randomized Controlled Pilot Trial

Avinash Pandey[1*]; Alexis A Krumme[2,3*], MS; Tejal Patel[4*], PharmD; Niteesh K Choudhry[2,3*], MD, PhD

[1]University of Western Ontario, London, ON, Canada

[2]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States

[3]Center for Healthcare Delivery Sciences, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States

[4]School of Pharmacy, University of Waterloo, Waterloo, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Niteesh K Choudhry, MD, PhD
Brigham and Women's Hospital and Harvard Medical School
1620 Tremont St Suite 3030
Boston, MA, 02120
United States
Phone: 1 617 278 0930
Fax: 1 617 232 8602
Email: nkchoudhry@bwh.harvard.edu

Related Article:

Correction of: http://mhealth.jmir.org/2017/8/e110

(JMIR Mhealth Uhealth 2019;7(4):e9254) doi:10.2196/mhealth.9254

The corresponding author for "The Impact of Text Messaging on Medication Adherence and Exercise Among Postmyocardial Infarction Patients: Randomized Controlled Pilot Trial" (JMIR MHealth UHealth 2017 Aug 3;5(8):e110) has been changed from Alexis Krumme to Niteesh Choudhry, whose contact information is as follows:

Niteesh K Choudhry, MD, PhD

Brigham and Women's Hospital, Harvard Medical School

1620 Tremont Street, Suite 3030

Boston, MA, 02120

Phone: 1 617 278 0930

Fax: 1 617 232 8602

Email: nkchoudhry@bwh.harvard.edu

Additionally, middle initials have been added to the following authors' names:

Alexis A Krumme

Niteesh K Choudhry

The correction will appear in the online version of the paper on the JMIR website on April 3, 2019, 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 also has been resubmitted to those repositories.

Edited by G Eysenbach; submitted 08.12.17; this is a non–peer-reviewed article;accepted 01.02.19; published 03.04.19.

Please cite as:
Pandey A, Krumme AA, Patel T, Choudhry NK
Author Correction: The Impact of Text Messaging on Medication Adherence and Exercise Among Postmyocardial Infarction Patients: Randomized Controlled Pilot Trial
JMIR Mhealth Uhealth 2019;7(4):e9254
URL: https://mhealth.jmir.org/2019/4/e9254/
doi:10.2196/mhealth.9254
PMID:30943170

Corrigenda and Addenda

Authorship Correction: Perspectives of Nonphysician Clinical Students and Medical Lecturers on Tablet-Based Health Care Practice Support for Medical Education in Zambia, Africa: Qualitative Study

Sandra Barteit[1], MA; Florian Neuhann[1], MD; Till Bärnighausen[1,2,3], MSc, MD, ScD, PhD; Annel Bowa[4], MSc; Sigrid Lüders[5], MD; Gregory Malunga[4], MSc; Geoffrey Chileshe[4], BSc; Clemence Marimo[6], MD; Albrecht Jahn[1], MSc, MD, PhD

[1]Heidelberg Institute of Global Health, Heidelberg, Germany

[2]Harvard TH Chan School of Public Health, Department of Global Health and Population, Boston, MA, United States

[3]Africa Health Research Institute (AHRI), KwaZulu-Natal, South Africa

[4]Chainama College of Health Sciences, Lusaka, Zambia

[5]SolidarMed, Lusaka, Zambia

[6]School of Medicine, University of Zambia, Lusaka, Zambia

Corresponding Author:
Sandra Barteit, MA
Heidelberg Institute of Global Health
Im Neuenheimer Feld 130.3
Heidelberg, 69120
Germany
Phone: 49 06221 ext 5634030
Email: barteit@uni-heidelberg.de

Related Article:

Correction of: http://mhealth.jmir.org/2019/1/e12637/

(JMIR Mhealth Uhealth 2019;7(4):e13431) doi:10.2196/13431

The authors of "Perspectives of Nonphysician Clinical Students and Medical Lecturers on Tablet-Based Health Care Practice Support for Medical Education in Zambia, Africa: Qualitative Study" (JMIR Mhealth Uhealth 2019;7(1):e12637) inadvertently omitted Annel Bowa, MSc (Chainama College of Health Sciences, Lusaka, Zambia) from the list of authors, although initially he had been included as a co-author and had signed the License to Publish form. He has now been added to the list of authors directly after Till Bärnighausen.

The correction will appear in the online version of the paper on the JMIR website on April 3, 2019, 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 also has been resubmitted to those repositories.

Multimedia Appendix 1

Letter from authors requesting correction.

[PDF File (Adobe PDF File), 465KB - mhealth_v7i4e13431_app1.pdf]

Edited by G Eysenbach; submitted 17.01.19; this is a non–peer-reviewed article;accepted 18.01.19; published 03.04.19.

Please cite as:
Barteit S, Neuhann F, Bärnighausen T, Bowa A, Lüders S, Malunga G, Chileshe G, Marimo C, Jahn A
Authorship Correction: Perspectives of Nonphysician Clinical Students and Medical Lecturers on Tablet-Based Health Care Practice Support for Medical Education in Zambia, Africa: Qualitative Study
JMIR Mhealth Uhealth 2019;7(4):e13431
URL: https://mhealth.jmir.org/2019/4/e13431/
doi:10.2196/13431
PMID:30943169

Review

Methodological Strategies for Ecological Momentary Assessment to Evaluate Mood and Stress in Adult Patients Using Mobile Phones: Systematic Review

Yong Sook Yang[1], MSN; Gi Wook Ryu[1], MPH; Mona Choi[1], PhD

Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Mona Choi, PhD
Mo-Im Kim Nursing Research Institute
College of Nursing
Yonsei University
50 Yonsei-ro, Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 3341
Fax: 82 2 2227 8303
Email: monachoi@yuhs.ac

Abstract

Background: Ecological momentary assessment (EMA) has utility for measuring psychological properties in daily life. EMA has also allowed researchers to collect data on diverse experiences and symptoms from various subjects.

Objective: The aim of this study was to review methodological strategies and useful related information for EMA using mobile phones to capture changes of mood and stress in adult patients seeking health care.

Methods: We searched PubMed, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Embase, the Cochrane Library, PsycINFO, and Web of Science. This review included studies published in peer-reviewed journals in English between January 2008 and November 2017 that used basic- or advanced-feature mobile phones to measure momentary mood or stress in adult patients seeking health care in outpatient departments. We excluded studies of smoking and substance addictions and studies of mental disorder patients who had been diagnosed by physicians.

Results: We reviewed 12 selected articles that used EMA via mobile phones to measure momentary mood and stress and other related variables from various patients with chronic fatigue syndrome, breast cancer, migraine, HIV, tinnitus, temporomandibular disorder, end-stage kidney disease, and traumatic brain injury. Most of the selected studies (11/12, 92%) used signal contingency and in 8 of the 12 studies (67%) alarms were sent at random or semirandom intervals to prompt the momentary measurement. Out of 12 studies, 7 (58%) used specific apps directly installed on mobile phones, 3 (25%) used mobile phones to link to Web-based survey programs, and 2 (17%) used an interactive voice-response system.

Conclusions: This study provides researchers with useful information regarding methodological details for utilizing EMA to measure mood and stress in adult patients. This review shows that EMA methods could be effective and reasonable for measuring momentary mood and stress, given that basic- and advanced-feature mobile phones are ubiquitous, familiar, and easy to approach. Therefore, researchers could adopt and utilize EMA methods using mobile phones to measure psychological health outcomes, such as mood and stress, in adult patients.

(JMIR Mhealth Uhealth 2019;7(4):e11215) doi:10.2196/11215

KEYWORDS

review; experience sampling method; ecological momentary assessment; mobile apps; mood; stress

Introduction

Momentary assessment techniques, such as ecological momentary assessment (EMA), have a long tradition as a prospective and repeated-measures longitudinal research methodology [1]. Originally, paper diaries were used in combination with pagers or electronic wristwatches. As technology became more advanced, data collection logistics

and reliability were improved by the use of personal digital assistants and mobile phone apps [2]. The method focuses on symptoms and adaptive function, such as well-being, and aims to map daily psychological function [3]. This method captures fluctuations by taking measurements multiple times day-to-day, unlike retrospective reporting, and has produced many findings with respect to psychological properties in the daily life of subjects [4-7].

EMA methods have ecological validity because assessments are made in natural and real-life environments, which reduces recall bias and avoids aggregation since it assesses the actual moment of interest repeatedly at multiple time points [3,8]. These repeated measures over time can reduce assessment error and improve the validity, reliability, and transparency of individual pattern assessments [3]. These aspects of increasing accuracy [8] and sensitivity to changes [9] in various properties have made EMA advantageous to study psychological state, quality of life, mobility, social networks, and more [3]. This method is considered suitable for understanding daily changes in psychological features such as mood and stress [10-12]. Traditionally, mood and stress have been assessed using retrospective measures [13]. EMA methods might provide health care providers with more accurate data than retrospective and global self-reporting methods. This may increase access to effective treatments by enabling enhanced understanding of the daily mood and stress of subjects, which are closely related to environmental factors.

The prevalence of mobile phones is increasing. In addition, advanced mobile technology has rendered mobile phones a novel, plausible way to implement EMA methods utilizing mobile technology, which is already available and familiar to many populations [14-16]. In an EMA study of police officers using a mobile phone app, participants indicated that the EMA correctly measured their mood and stress; they also felt comfortable using the app installed on their own mobile phones [12].

There have been systematic reviews of EMA methods monitoring adult patients with psychiatric disorders. A review study of depressive symptoms or affective disorders showed that the monitoring system using a mobile phone-based EMA method was feasible and accurate in predicting mood, but this study did not include postpartum, postnatal, or pregnant women with depressive symptoms [17]. Another review of studies on anxiety disorders, such as panic disorder, generalized anxiety disorder, social phobia, posttraumatic stress disorder, and obsessive-compulsive disorder [18], found that EMA methods have the potential to illuminate patients' anxiety in their everyday lives.

However, there is no extant review of the feasibility and use of EMA methodology using basic- or advanced-feature mobile phones to capture changes of mood and stress in adult patients without diagnoses of psychiatric disorders such as affective, anxiety, or mood disorders. Therefore, this review provides methodological details for the use of EMA technology to assess mood and stress in adult patients.

Methods

Information Source and Search Strategy

The search included studies that used mobile apps to measure momentary mood or stress in adults; the studies were published in peer-reviewed journals in English between January 2008 and November 2017. We performed database searches on six online biomedical databases—PubMed, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Embase, the Cochrane Library, PsycINFO, and Web of Science. We also performed hand-searches of the Journal of Medical Internet Research (JMIR) and the website of the Society for Ambulatory Assessment. We used the following search terms: ("ecological momentary assessment" [MeSH] OR "experience sampling" OR "ecological momentary" OR "event sampling" OR "ambulatory assessment" OR "structured diary method" OR "real-time data capture studies" OR "real-time data capture study" OR "beeper studies" OR "beeper study" OR "intensive longitudinal assessment") AND ("stress, psychological" [MeSH] OR "affect" [MeSH] OR "mood" OR "emotion" OR "affection" OR "stress") AND ("mobile applications" [MeSH] OR "smartphone" [MeSH] OR "cell phones" [MeSH] OR "smartphone*" OR "cell phone" OR "cellular phone" OR "mobile app*"). The articles identified were inspected, including their reference lists and in-text citations of relevant articles (see Multimedia Appendix 1).

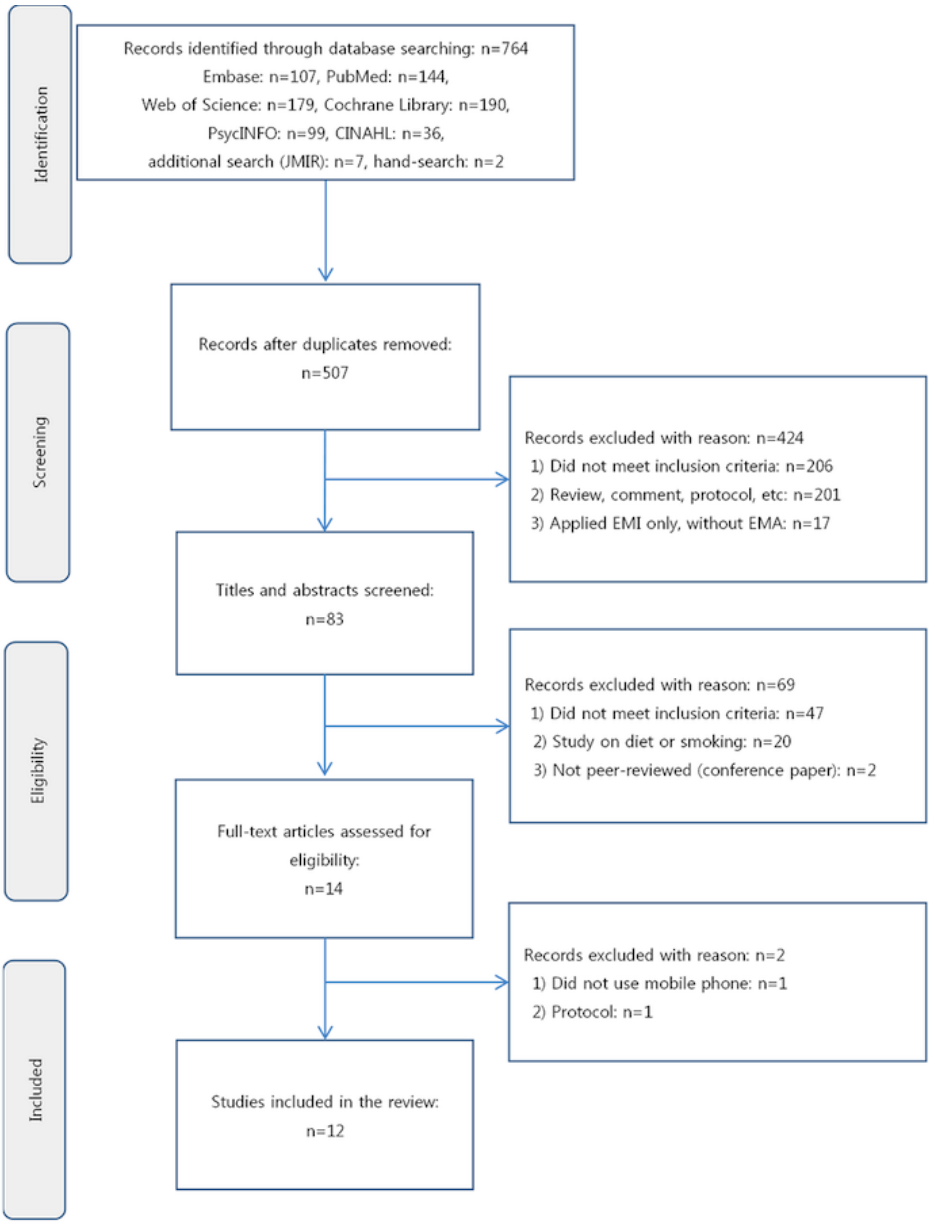
Study Selection

Studies were included that used basic- and advanced-feature mobile phones to measure momentary mood or stress in adult patients. We included those studies that were published in peer-reviewed journals in English. Specifically, included studies used basic- or advanced-feature mobile phones to deliver EMAs. Included studies also involved adult patients in community settings who were diagnosed with a certain disease by their physicians and cared for in outpatient settings. We also included studies that involved people who had mood or stress problems without diagnosis by their physicians of psychiatric disorders, such as affective, anxiety, and mood disorders or of substance addictions. The year 2008 was chosen as the earliest year of publication because the first app downloaded on a mobile device was in 2008 [19]. Studies were excluded if they were studies of smoking, diet, addictions, major psychological problems, or child populations.

Screening Procedure

A total of 764 articles were retrieved from the six databases, in which 257 records were duplicated. For 507 articles, two reviewers (YSY and GWR) independently screened titles and abstracts. After that, the same two reviewers independently reviewed full-text articles to decide whether each article was relevant to the review. In case of disagreement, a third person (MC) was consulted to reach consensus. Ultimately, 12 full-text articles were selected according to the criteria and relevant data were extracted. Figure 1 shows the process of study selection based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [20].

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart providing an overview of the study selection process. CINAHL: Cumulative Index to Nursing and Allied Health Literature; EMA: ecological momentary assessment; EMI: ecological momentary intervention; JMIR: Journal of Medical Internet Research.



Data Extraction

The following information was extracted: study purpose, sample characteristics, main momentary measurement, data analysis method, and methodological details of EMA, such as operating system, mode of response, contingency, duration of data collection, frequency, and alarm interval for each study.

Results

In total, 12 studies met the selection criteria. The following sections summarize how EMA approaches were applied to the study populations along with methodological details.

Subject Characteristics and Main Momentary Measurements

Various clinical populations were included in this review: patients with chronic fatigue syndrome [21,22], HIV [5,23-25], tinnitus [6], migraine [26], minor traumatic brain injury [27], breast cancer [7], end-stage kidney disease [28], and temporomandibular disorder [29]. Out of 12 studies, 11 (92%) [5,7,21-29] measured mood or stress as major variables by the EMA method; 1 study (8%) assessed stress or stressors [6]; and 5 studies (42%) [5,6,24-26] measured mood or affect along with stress or stressors. A total of 2 studies out of 12 (17%) measured pain along with mood or stress [26,29]. The majority of articles (11/12, 92%) captured changes in mood or stress over time without applying interventions [5-7,21-28]; 1 article out of 12 (8%) reported changes of mood from a pre-post approach [29] (see Table 1).

Table 1. Study details including study purpose, sample characteristics, and main momentary measurements.

Author (year), country	Study purpose	Sample characteristics	Sample size, n	Age in years	Main momentary measurements
Band et al (2016) [21], United Kingdom	To examine relationship between significant others' responses and patient outcomes	Pairs of CFS[a] patients and significant others	23	35.5 (14.0)[b]	Affects, significant others' responses, symptom severity, disability, and activity management strategies
Band et al (2017) [22], United Kingdom	To investigate whether activity patterns occurred according to patient symptom experience and affect	CFS patients	23	35.5 (14.0)[b]	Patient activity management strategies, patient affects, and symptoms
Farmer et al (2017) [5], United States	To assess stress, frequency of stressors, stressful life events, and behaviors	Patients with HIV	32	46.0 (23-64)[c]	Stressors, stress level, emotional and physical states, medication adherence, and sexual activity
Moore et al (2017) [23], United States	To examine feasibility, acceptability, and initial validity of using mobile phone-based EMA[d]	Older adults with HIV	20	58.8 (4.3)[b]	Mood and cognitive symptoms
Cook et al (2017) [24], United States	To test whether momentary motivation was a mechanism by which everyday experiences affect medication adherence	Patients with HIV	87	40.0 (8.8)[b]	Control beliefs, mood, stress, coping, and social support
Cook et al (2017) [25], United States	To test predictors of electronically monitored adherence at both the state and trait levels and to compare relative effects	Patients with HIV	87	40.0 (8.8)[b]	Thoughts, mood, stress, coping, social support, and treatment motivation
Wilson et al (2015) [6], United States	To explore feasibility of EMA as a tool to more accurately assess the level of bother from tinnitus	Tinnitus patients	20	55 (38-65)[c]	Bother, loudness, and stress
Houtveen et al (2013) [26], the Netherlands	To test prodromal functioning relative to the interictal state	Migraine patients	87	44.5 (25-68)[c]	Migraine attacks and prodromal features: fatigue, cognitive functioning, affects, and stressors
Juengst et al (2015) [27], United States	To assess pilot feasibility and validity of a mobile health system for tracking mood-related symptoms after traumatic brain injury	Traumatic brain injury patients	20	36.7 (12.4)[b]	Depressive and anxious mood, impact of fatigue, and affects
Kim et al (2016) [7], South Korea	To evaluate the potential of a mobile mental health tracker, the impact of adherence on reporting, and its accuracy	Breast cancer patients	78	44.4 (7.0)[b]	Sleep satisfaction, mood, and anxiety
Abdel-Kader et al (2014) [28], United States	To evaluate day-to-day and diurnal variability of fatigue, sleepiness, exhaustion, and related symptoms	End-stage kidney disease patients	55	56.7 (17.3)[b]	Mood, cognition, sleepiness, and exhaustion
Litt et al (2009) [29], United States	To determine whether cognitive-behavioral therapy treatment operates by effecting changes in cognitions, affects, and coping behaviors	Temporomandibular disorder patients	54	41.0 (11.9)[b]	Pain, coping, and affects

[a]CFS: chronic fatigue syndrome.

[b]Mean (SD).

[c]Median (range).

[d]EMA: ecological momentary assessment.

Out of 12 studies, 2 studies of patients with chronic fatigue syndrome (CFS) (17%) assessed patients' affect [21,22]. Out of these 2 studies, 1 (50%) focused on activity management strategies, affect, and symptoms to investigate whether activity

patterns occurred according to patients' symptom experience and affect [22]. The other study (1/2, 50%) examined the relationship between significant others' responses and patient outcomes such as affect, symptom severity, disability, and activity management strategies [21].

Out of 12 studies, 4 EMA studies of HIV patients (33%) measured several variables. Out of these 4 studies, 1 (25%) evaluated momentary mood and cognitive symptoms of HIV patients [23], and another (1/4, 25%) assessed control beliefs, mood, stress, coping, and social support to examine whether momentary motivation is a mechanism by which everyday experiences affect adherence to medication therapy [24]. Cook et al's study (1/4, 25%) measured thoughts, mood, stress, coping, social support, and treatment motivation to test predictors of electronically monitored adherence at both state and trait levels [25]. Out 4 studies, 1 EMA study (1/4, 25%) also investigated stress, frequency of stressors, stressful life events, and behaviors of HIV patients [5].

Out of 12 studies, 1 (8%) assessed experience of migraine attacks and prodromal features, such as fatigue, cognitive functioning, affect, effort spent (eg, working hard and feeling strained), and stressors, to test and identify individual prodromal features related to the interictal state in moderate-to-severe migraine patients [26].

In a study of minor traumatic brain injury patients (1/12, 8%), mood and affect were assessed to evaluate feasibility and validity of a mobile health system app [27]. A study of patients with breast cancer (1/12, 8%) measured sleep satisfaction, mood, and anxiety to evaluate the potential of a mobile, mental health tracker app using daily mental health ratings as indicators of depression [7].

Out of 12 studies, 1 (8%) evaluated day-to-day and diurnal variability of fatigue, sleepiness, exhaustion, and related symptoms in end-stage kidney disease patients [28]. Out of 12 studies, a pre-post EMA design in 1 study (8%) was applied to measure pain, coping, and affect in order to evaluate the effect of cognitive-behavioral treatment for temporomandibular disorder patients in the context of painful episodes [29].

In terms of main momentary measurement, half of the included studies (6/12, 50%) measured momentary mood, affect, or stress with standardized scales for validation [7,23-25,27,29], while others (6/12, 50%) did not administer or specify them [5,6,21,22,26,28]. Band et al's study (1/12, 8%) captured mood changes by using two subscales of positive affect and negative affect of CFS patients [22]. In this study, positive affect was assessed using five items: excited, happy, satisfied, relaxed, and cheerful (Cronbach alpha=.87); negative affect was assessed using five items: sad, annoyed, irritated, anxious, lonely, and guilty (Cronbach alpha=.87). In the other study of CFS patients (1/12, 8%), affect was measured by a single item, *feeling distressed*, which was included with standard items examining patients' affect at a momentary level [21].

The standardized measures of the Beck Depression Inventory-II and the Profile of Mood States were administered to measure state mood and stress in comparison to the momentary item for assessing mood of older adults with HIV; correlates with state mood (ie, sadness, happiness, and tiredness) and stress were evaluated by item questions developed in the study [23].

In a study of HIV patients (1/12, 8%) [24], three items for mood (Cronbach alpha=.93) and six items for stress (Cronbach alpha=.67) from the Diary of Ambulatory Behavioral States were used after piloting [30]. Another study of HIV patients (1/12, 8%) used the mood scale from the Diary of Ambulatory Behavioral States and the stress scale from the Daily Hassles Scale; they were validated by the trait measurement tools from the Center for Epidemiological Studies-Depression scale and the HIV/AIDS-Targeted Quality of Life instrument [25]. Both trait-level mood and stress predicted their respective state-level measures.

In 1 study out of 12 (8%), the Daily Mood and Affect scale for momentary assessment was developed; the Positive and Negative Affect Schedule and the 9-item Patient Health Questionnaire as standardized measures were applied [27]. In a study of breast cancer patients (1/12, 8%), the author used 3-item short scales for anxiety, mood, and sleep satisfaction, rated by facial emoticon scales, and evaluated the concurrent validity with the standardized mood scale of the 9-item Patient Health Questionnaire [7].

In 1 study out of 12 (8%), evaluating the effects of cognitive-behavioral therapy of patients with temporomandibular pain, a standardized tool—the Center for Epidemiological Studies-Depression scale—was used to compare pre- to posttreatment change of affect using a mood item borrowed from the Coping Strategies Questionnaire [29].

Out of 12 studies, 4 (33%) reported on feasibility or validity of an EMA app [6,23,26,27]. A study with EMA design for patients suffering from tinnitus (1/4, 25%) indicated that they would suggest an EMA method to a friend [6]. Participants expressed their experience with the EMA method positively [23,27]; they reported that they accepted it as usable and were satisfied with the EMA method [26,27].

An evaluation of the usefulness or perceptions by participants of the EMA methods was conducted in another study (1/12, 8%); the results indicated that the EMA using mobile phones was useful and reliable for self-monitoring of functioning ability in daily routines [5]. EMA showed promising results in the field of screening depressive moods in a clinical population by evaluating accuracy of depression screening via the EMA method (1/12, 8%) [7].

Methodological Details of Ecological Momentary Assessment

Table 2 shows information on methodological details of EMA used in the studies, such as the operating system of mobile phones, mode of response, contingency, duration of data collection, frequency per day, and alarm interval. Different operating systems were used to install the mobile apps, but more than half of the studies (7/12, 58%) used Android operating systems [5,7,21,22,24,25,27].

Table 2. Completion rate and momentary data analysis method.

Author (year)	Operating system	Mode	Contingency	Duration in days, n	Frequency per day, n	Total frequency, n	Alarm interval
Band et al (2016) [21]	Android	App	Signal	6	10	60	Semirandom
Band et al (2017) [22]	Android	App	Signal	6	10	60	Semirandom
Farmer et al (2017) [5]	Android	App	Signal and event	42	1 (medication adherence); 4 (emotional and physical states); 7 (stressor)	42 (medication adherence); 168 (emotional and physical states); 294 (stressor)	Fixed; fixed; self-initiated time (event-based)
Moore et al (2017) [23]	Android	App	Signal	7	5	35	Fixed (adjusted for participant)
Cook et al (2017) [24]	Android	Link to online survey	Signal	70	1	70	Random
Cook et al (2017) [25]	Android	Link to online survey	Signal	70	1	70	Random
Wilson et al (2015) [6]	Not specified	Link to online survey	Signal	14	4	56	Random (09:00-20:00)
Houtveen et al (2013) [26]	Nokia	App	Signal	28	4	112	Random (09:30-16:00); semirandom at get-up time and bedtime
Juengst et al (2015) [27]	Not specified	App	Signal	56[a]	1	56	Fixed by preference
Kim et al (2016) [7]	Not specified	App	Not specified	336	1	336	Not specified
Abdel-Kader et al (2014) [28]	Not specified	IVR[b]	Signal (call)	7	4	28	Fixed
Litt et al (2009) [29]	Not specified	IVR	Signal (call)	7 (pre); 14 (post)	4	28; 56	Random (08:00-22:00)

[a]Repeated four times over 8 weeks.

[b]IVR: interactive voice response.

Out of the 12 studies, 7 (58%) used specific apps directly installed onto mobile phones [5,7,21-23,26,27]; 3 studies (25%) used a Web-based online survey program hyperlinked from the mobile phones [6,24,25]; and the remaining 2 studies (17%) applied an EMA method using an interactive voice-response system [28,29]. A daily repeated voice-recorded EMA design could be a good system for patients with motor dysfunction, instead of a mobile phone app or online survey in which patients have to operate the phones to respond.

Out of 12 studies, 11 (92%) [5,6,21-29] included in the review used signal contingency to prompt momentary measurement; there was 1 study (8%) where the contingency method was not specified [7]. In a study of patients with HIV (1/12, 8%), both signal-based contingency and event-based self-report were applied [5]. Frequency of the contingency varied from once per day [7] to 10 times per day [21,22], and the study durations ranged from a minimum of 6 days [21,22] to a maximum of 48 weeks, which equals 336 days [7].

The studies with the shortest period (2/12, 17%) had the highest frequency per day of assessment [21,22], and studies with lengthy periods of more than 6 weeks (4/12, 33%) had the lowest frequency [7,24,25,27]. Out of 12 studies, 1 (8%) tried various frequencies of momentary assessment by constructs: once a day for measuring medication adherence, four times a day for emotional and physical states, and seven times per day for stressors [5].

The interval of the reminder signal varied according to the study design from random, stratified semirandom, and semirandom to fixed time per participant. A total of 4 studies out of 12 (33%) had set the alarm time as fixed according to the preference or convenience of each participant to improve compliance to the EMA [5,23,27,28].

Study completion rates ranged from 64.6% [22] to 89.5% [26], excluding studies with no reported completion rates (see Table 3). A study of temporomandibular disorder patients (1/12, 8%) paid participants US $5 for every day that they completed at least 50% of scheduled daily assessments [29], while 2 studies of HIV patients (17%) provided incentives of US $25 and the mobile phone used in the study when they finished the EMA measurement [24]. In another study of patients with HIV (1/12,

8%), in which both signal-based and event-based EMA methods were applied, event-based self-reports were encouraged by applying incentives up to US $70 to reach the survey goal of seven times per day [5]. However, the study did not calculate completion rate, since the measure was reported in an event-based way [5]. Other studies included in this review (6/12, 50%) did not mention incentives [21-23,26-28]. No articles evaluated related factors affecting the completion rate.

Because EMA datasets include diverse sources of variance, various analysis methods have been employed to address the complexity and hierarchy of the data. Out of the 12 studies, 7 (58%) reviewed here undertook multilevel or mixed-modeling analysis [21,22,24-26,28,29]. A total of 2 studies out of 12 (17%) used the MTMIXED command in Stata (StataCorp LLC) for continuous outcome variables in multilevel modeling [21,22]. Linear mixed-model multilevel analysis with maximum likelihood estimation was employed [24-26] (see Table 3).

Of the 12 studies, 2 (17%) used descriptive analysis and correlation analysis [23,27], and 2 others (17%) applied the receiver operating characteristic and ordinary least squares according to the characteristics of the variables analyzed [6,7]. Kim et al's study [7] (1/12, 8%) estimated random-effects logistic regression parameters and thereafter used receiver operating characteristic plots to evaluate the screening accuracy of the model.

Of the 12 studies, 1 study of HIV patients (8%) applied EMA using both quantitative and qualitative measurement with various frequencies according to the target variables. The data analysis method for quantitative data was not specified, while a grounded thematic coding method in Dedoose (SocioCultural Research Consultants LLC), a Web-based mixed-method data analysis program, was applied for qualitative data of the user experience of the usefulness or perceptions regarding the EMA app [5].

While there is no standard for appropriate response rate to assess validity, 1 study out of 12 (8%) clarified that they used all available daily observations [24], and another (1/12, 8%) excluded participants who completed fewer than 20 assessments out of the total of 60 for preliminary analysis but retained all participants in the final analysis [22]; other studies did not specify inclusion criteria for response rate or number of observations for statistical analysis. A study of temporomandibular pain patients (1/12, 8%) used the observations selectively, in accordance with the study purpose, in which pain was nonzero and coping was recorded at the same time [29].

Of the 12 studies, 7 (58%) had a briefing or intake session to ensure that participants understood the EMA app before starting the survey. Participants could practice and ask questions regarding the app during the session. Informed consent and non-EMA measures, such as baseline or laboratory measurement, were also obtained during the session. After finishing the EMA phase, patients were debriefed to evaluate their experiences during the study.

Table 3. Completion rate and method used to analyze momentary data.

Author (year)	Completion rate of EMA[a], n/N (%) or % (where n/N was not available)	Analysis method
Band et al (2016) [21]	38.74/60 (65)	Multilevel models
Band et al (2017) [22]	893/1380 (64.71)	Multilevel models
Farmer et al (2017) [5]	Not reported	Ground thematic coding method (not specified for quantitative data analysis)
Moore et al (2017) [23]	30/35 (86)	Descriptive and correlation analysis
Cook et al (2017) [24]	73.0	Multilevel modeling analysis
Cook et al (2017) [25]	65.0	Multilevel modeling analysis
Wilson et al (2015) [6]	889/1120 (79.38)	Ordinary least squares robust regression analysis
Houtveen et al (2013) [26]	89.5	Linear mixed-model multilevel analysis
Juengst et al (2015) [27]	73.4	Descriptive and correlation analysis
Kim et al (2016) [7]	Not reported	Random-effect model of logistic regression and receiver operating characteristic
Abdel-Kader et al (2014) [28]	1252/1540 (81.30)	Linear mixed model
Litt et al (2009) [29]	72.0 (pre); 71.0 (post)	Mixed model

[a]EMA: ecological momentary assessment.

Discussion

Principal Findings

This review identified mobile phone-based systems for monitoring mood or stress of patients seeking health care in outpatient departments. Studies focused on EMA methods using mobile phones, which are feasible for measuring stress and mood in adult patients and elucidating relevant methodological details. The EMA methods used in the included studies were evaluated as feasible for recognizing changes with significant variation in assessment variables [27-29] and for measuring mood and stress of patients [6,7,23,27]. This review presented

strategic information on EMA methods, such as mode of response, ways of sending alarm contingencies, time intervals, frequencies, and study durations, along with information about the participants in the survey and the momentary measurements.

The studies in our review used three different modes of EMA response on mobile phones: via mobile app [5,7,21-23,26,27]; via hyperlink to online survey [6,24,25]; and via interactive voice response system [28,29]. Mode of response can be selected in accordance with participants' clinical conditions.

EMA methods are time-consuming and demanding [4]. Not all patients are willing to participate or comply strictly with the protocol. The studies included in this review showed completion rates that ranged from 64.6% to 89.5%, which was contingent on the nature of the participants. Although there is no agreed-upon gold standard for an acceptable compliance rate in EMA studies, Stone and Shiffman [31] noted that EMA data would not be representative of participants' daily lives if compliance was lower than 80%, while another study considered that analysis using observations of participants who responded over 75% of the time would be reasonable [6].

One challenge is the complexity of EMA data [32]. An EMA protocol usually must consider item selection, period, intensity, signaling algorithm, event recording, application type, and data storage. Our review showed that the frequency of data collection varied from 1 to 10 times per day over a time period of 6 days to 48 weeks. Repeatedly answering the same questions in an EMA method requires substantial involvement, which increases the respondent's burden, and this aspect can be frustrating for participants [33]. Related to this complexity of data collection, missing data also presents a limitation [31].

Regarding data analysis, EMA studies tend to produce multilevel datasets from multiple participants who answer a set of questions at multiple times. Therefore, standard linear and logistic regression analysis techniques are insufficient for analysis of EMA datasets. The complexity of EMA data analysis could hinder researchers and clinicians in using this method [5]. This should be taken into account when considering this technology-driven approach.

A limitation of this review is that we did not include studies that utilized other mobile devices, such as wearable sensors or personal digital assistants, since the purpose of this review was to provide insight into methodological strategies for EMA using mobile phones to assess mood or stress.

Future studies would include objective measures of related variables, such as heart rate, physical activity, and walking, which may be affected by mood and stress, to confirm dynamic relationships between symptoms and mood and stress. Additionally, multidisciplinary research involving areas such as medical diagnosis, consultation, nursing care, and ecological momentary interventions (EMIs) with EMA data collection could be an interesting focus. Through these multiple approaches, we expect to perform more accurate and valid mental and physical health monitoring and to provide optimized medical service for patients by applying patient-specific health care interventions.

Conclusions

Prevalence of basic- and advanced-feature mobile phones is high, and mobile technology is readily used as a ubiquitous resource. Mobile phones can be utilized easily in health research to assess patients' experiences in their daily lives, as they are convenient for patients to carry and are user friendly. In addition, patients may feel comfortable using their own familiar mobile phones with EMA methods installed.

This review provides researchers with information regarding methodological details, such as length of administration period, mode of response, contingency of sending alarms, frequencies and durations, incentives for improving compliance, and statistical methods for data analysis when utilizing EMA to measure mood and stress in adult patients.

Despite the limitations of this study, we believe this review shows that EMA is an effective and reasonable way of measuring momentary mood and stress in an era in which mobile phones are ubiquitous in the general population, including patients. In particular, individuals who have experienced mood changes or stress can benefit from EMA methods by using mobile phones to monitor or track their mood and stress vulnerabilities. This review supports the use of EMA methods to evaluate mood and stress and recommends that researchers utilize EMA methods to measure psychological health outcomes of mood and stress in various patient populations.

Acknowledgments

This study was supported by the Basic Science Research Program through the National Research Foundation of Korea, funded by the Korean Ministry of Education (2017R1D1A1B03030706).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search terms and results.

[PDF File (Adobe PDF File), 66KB - mhealth_v7i4e11215_app1.pdf]

References

1. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Ann Behav Med 1994 Jan 01;16(3):199-202 [FREE Full text] [doi: 10.1093/abm/16.3.199]
2. Doherty ST, Lemieux CJ, Canally C. Tracking human activity and well-being in natural environments using wearable sensors and experience sampling. Soc Sci Med 2014 Apr;106:83-92. [doi: 10.1016/j.socscimed.2014.01.048] [Medline: 24549253]
3. Augustine AA, Larsen RJ. Emotion research. In: Mehl MR, Conner TS, editors. Handbook of Research Methods for Studying Daily Life. New York, NY: Guilford Press; 2012:497-510.
4. Myin-Germeys I, Oorschot M, Collip D, Lataster J, Delespaul P, van Os J. Experience sampling research in psychopathology: Opening the black box of daily life. Psychol Med 2009 Sep;39(9):1533-1547. [doi: 10.1017/S0033291708004947] [Medline: 19215626]
5. Farmer S, Mindry D, Comulada WS, Swendeman D. Mobile phone ecological momentary assessment of daily stressors among people living with HIV: Elucidating factors underlying health-related challenges in daily routines. J Assoc Nurses AIDS Care 2017;28(5):737-751 [FREE Full text] [doi: 10.1016/j.jana.2017.04.001] [Medline: 28549526]
6. Wilson MB, Kallogjeri D, Joplin CN, Gorman MD, Krings JG, Lenze EJ, et al. Ecological momentary assessment of tinnitus using smartphone technology: A pilot study. Otolaryngol Head Neck Surg 2015 May;152(5):897-903 [FREE Full text] [doi: 10.1177/0194599815569692] [Medline: 25676150]
7. Kim J, Lim S, Min YH, Shin Y, Lee B, Sohn G, et al. Depression screening using daily mental-health ratings from a smartphone application for breast cancer patients. J Med Internet Res 2016 Dec 04;18(8):e216 [FREE Full text] [doi: 10.2196/jmir.5598] [Medline: 27492880]
8. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [Medline: 18509902]
9. Bolger N, Laurenceau JP. Introduction to intensive longitudinal methods. In: Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research. New York, NY: Guilford Press; 2013:1-10.
10. Cristóbal-Narváez P, Sheinbaum T, Ballespí S, Mitjavila M, Myin-Germeys I, Kwapil TR, et al. Impact of adverse childhood experiences on psychotic-like symptoms and stress reactivity in daily life in nonclinical young adults. PLoS One 2016;11(4):e0153557 [FREE Full text] [doi: 10.1371/journal.pone.0153557] [Medline: 27082442]
11. Wichers M, Simons CJP, Kramer IMA, Hartmann JA, Lothmann C, Myin-Germeys I, et al. Momentary assessment technology as a tool to help patients with depression help themselves. Acta Psychiatr Scand 2011 Oct;124(4):262-272. [doi: 10.1111/j.1600-0447.2011.01749.x] [Medline: 21838742]
12. Yang YS, Ryu GW, Han I, Oh S, Choi M. Ecological momentary assessment using smartphone-based mobile application for affect and stress assessment. Healthc Inform Res 2018 Oct;24(4):381-386 [FREE Full text] [doi: 10.4258/hir.2018.24.4.381] [Medline: 30443427]
13. Ebner-Priemer U, Trull TJ. Ecological momentary assessment of mood disorders and mood dysregulation. Psychol Assess 2009 Dec;21(4):463-475. [doi: 10.1037/a0017075] [Medline: 19947781]
14. Heron KE, Smyth JM. Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
15. Freedman MJ, Lester KM, McNamara C, Milby JB, Schumacher JE. Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. J Subst Abuse Treat 2006 Mar;30(2):105-111. [doi: 10.1016/j.jsat.2005.10.005] [Medline: 16490673]
16. Swendeman D, Ramanathan N, Baetscher L, Medich M, Scheffler A, Comulada WS, et al. Smartphone self-monitoring to support self-management among people living with HIV: Perceived benefits and theory of change from a mixed-methods randomized pilot study. J Acquir Immune Defic Syndr 2015 May 01;69 Suppl 1:S80-S91 [FREE Full text] [doi: 10.1097/QAI.0000000000000570] [Medline: 25867783]
17. Dogan E, Sander C, Wagner X, Hegerl U, Kohls E. Smartphone-based monitoring of objective and subjective data in affective disorders: Where are we and where are we going? Systematic review. J Med Internet Res 2017 Dec 24;19(7):e262 [FREE Full text] [doi: 10.2196/jmir.7006] [Medline: 28739561]
18. Walz LC, Nauta MH, Aan Het Rot M. Experience sampling and ecological momentary assessment for studying the daily lives of patients with anxiety disorders: A systematic review. J Anxiety Disord 2014 Dec;28(8):925-937. [doi: 10.1016/j.janxdis.2014.09.022] [Medline: 25445083]
19. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: A systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012. [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
21. Band R, Barrowclough C, Emsley R, Machin M, Wearden AJ. Significant other behavioural responses and patient chronic fatigue syndrome symptom fluctuations in the context of daily life: An experience sampling study. Br J Health Psychol 2016 Dec;21(3):499-514 [FREE Full text] [doi: 10.1111/bjhp.12179] [Medline: 26700742]

22. Band R, Barrowclough C, Caldwell K, Emsley R, Wearden A. Activity patterns in response to symptoms in patients being treated for chronic fatigue syndrome: An experience sampling methodology study. Health Psychol 2017 Dec;36(3):264-269 [FREE Full text] [doi: 10.1037/hea0000422] [Medline: 27819461]
23. Moore RC, Kaufmann CN, Rooney AS, Moore DJ, Eyler LT, Granholm E, et al. Feasibility and acceptability of ecological momentary assessment of daily functioning among older adults with HIV. Am J Geriatr Psychiatry 2017 Aug;25(8):829-840 [FREE Full text] [doi: 10.1016/j.jagp.2016.11.019] [Medline: 28065496]
24. Cook PF, Schmiege SJ, Bradley-Springer L, Starr W, Carrington JM. Motivation as a mechanism for daily experiences' effects on HIV medication adherence. J Assoc Nurses AIDS Care 2018;29(3):383-393. [doi: 10.1016/j.jana.2017.09.003] [Medline: 28967490]
25. Cook PF, Schmiege SJ, Starr W, Carrington JM, Bradley-Springer L. Prospective state and trait predictors of daily medication adherence behavior in HIV. Nurs Res 2017;66(4):275-285 [FREE Full text] [doi: 10.1097/NNR.0000000000000216] [Medline: 28654566]
26. Houtveen JH, Sorbi MJ. Prodromal functioning of migraine patients relative to their interictal state: An ecological momentary assessment study. PLoS One 2013;8(8):e72827 [FREE Full text] [doi: 10.1371/journal.pone.0072827] [Medline: 23977358]
27. Juengst SB, Graham KM, Pulantara IW, McCue M, Whyte EM, Dicianno BE, et al. Pilot feasibility of an mHealth system for conducting ecological momentary assessment of mood-related symptoms following traumatic brain injury. Brain Inj 2015;29(11):1351-1361. [doi: 10.3109/02699052.2015.1045031] [Medline: 26287756]
28. Abdel-Kader K, Jhamb M, Mandich LA, Yabes J, Keene RM, Beach S, et al. Ecological momentary assessment of fatigue, sleepiness, and exhaustion in ESKD. BMC Nephrol 2014 Feb 06;15:29 [FREE Full text] [doi: 10.1186/1471-2369-15-29] [Medline: 24502751]
29. Litt MD, Shafer DM, Ibanez CR, Kreutzer DL, Tawfik-Yonkers Z. Momentary pain and coping in temporomandibular disorder pain: Exploring mechanisms of cognitive behavioral treatment for chronic pain. Pain 2009 Sep;145(1-2):160-168 [FREE Full text] [doi: 10.1016/j.pain.2009.06.003] [Medline: 19553018]
30. Cook PF, McElwain CJ, Bradley-Springer LA. Feasibility of a daily electronic survey to study prevention behavior with HIV-infected individuals. Res Nurs Health 2010 Jun;33(3):221-234. [doi: 10.1002/nur.20381] [Medline: 20499392]
31. Stone AA, Shiffman S. Capturing momentary, self-report data: A proposal for reporting guidelines. Ann Behav Med 2002;24(3):236-243. [doi: 10.1207/S15324796ABM2403_09] [Medline: 12173681]
32. Schwartz JE, Stone AA. Strategies for analyzing ecological momentary assessment data. Health Psychol 1998 Jan;17(1):6-16. [Medline: 9459065]
33. Verhagen SJ, Hasmi L, Drukker M, van Os J, Delespaul PA. Use of the experience sampling method in the context of clinical trials. Evid Based Ment Health 2016 Aug;19(3):86-89 [FREE Full text] [doi: 10.1136/ebmental-2016-102418] [Medline: 27443678]

Abbreviations

CFS: chronic fatigue syndrome
CINAHL: Cumulative Index to Nursing and Allied Health Literature
EMA: ecological momentary assessment
EMI: ecological momentary intervention
IVR: interactive voice response
JMIR: Journal of Medical Internet Research
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach; submitted 26.06.18; peer-reviewed by T Donker, P Chow, D Hidalgo-Mazzei; comments to author 06.10.18; revised version received 31.12.18; accepted 27.01.19; published 01.04.19.

Please cite as:
Yang YS, Ryu GW, Choi M
Methodological Strategies for Ecological Momentary Assessment to Evaluate Mood and Stress in Adult Patients Using Mobile Phones: Systematic Review
JMIR Mhealth Uhealth 2019;7(4):e11215
URL: https://mhealth.jmir.org/2019/4/e11215/
doi:10.2196/11215
PMID:30932866

Original Paper

Context-Sensitive Ecological Momentary Assessment: Application of User-Centered Design for Improving User Satisfaction and Engagement During Self-Report

Preethi Srinivas[1,2], PhD; Kunal Bodke[1,2], MS; Susan Ofner[3], MS; NiCole R Keith[1,2,4], PhD; Wanzhu Tu[1,3], PhD; Daniel O Clark[1,2,5], PhD

[1]Indiana University, Center for Aging Research, Indianapolis, IN, United States

[2]Regenstrief Institute, Indianapolis, IN, United States

[3]Indiana University, School of Medicine, Department of Biostatistics, Indianapolis, IN, United States

[4]Indiana University, School of Health and Human Sciences, Department of Kinesiology, Indianapolis, IN, United States

[5]Indiana University, School of Medicine, Department of Medicine, Indianapolis, IN, United States

Corresponding Author:
Daniel O Clark, PhD
Indiana University
Center for Aging Research
1101 West 10th Street
Indianapolis, IN, 46202
United States
Phone: 1 3172749000
Fax: 1 3172749307
Email: daniclar@iu.edu

Abstract

Background: Ecological momentary assessment (EMA) can be a useful tool for collecting real-time behavioral data in studies of health and health behavior. However, EMA administered through mobile technology can be burdensome, and it tends to suffer from suboptimal user engagement, particularly in low health-literacy populations.

Objective: This study aimed to report a case study involving the design and evaluation of a mobile EMA tool that supports context-sensitive EMA-reporting of location and social situations accompanying eating and sedentary behavior.

Methods: An iterative, user-centered design process with obese, middle-aged women seeking care in a safety-net health system was used to identify the preferred format of self-report measures and the look, feel, and interaction of the mobile EMA tool. A single-arm feasibility field trial with 21 participants receiving 12 prompts each day for momentary self-reports over a 4-week period (336 total prompts per participant) was used to determine user satisfaction with interface quality and user engagement, operationalized as response rate. A second trial among 38 different participants randomized to receive or not to receive a feature designed to improve engagement was conducted.

Results: The feasibility trial results showed high interface satisfaction and engagement, with an average response rate of 50% over 4 weeks. Qualitative feedback pointed to the need for auditory alerts. We settled on 3 alerts at 10-min intervals to accompany each EMA-reporting prompt. The second trial testing this feature showed a statistically significant increase in the response rate between participants randomized to receive repeat auditory alerts versus those who were not (60% vs 40%).

Conclusions: This paper reviews the design research and a set of design constraints that may be considered in the creation of mobile EMA interfaces personalized to users' preferences. Novel aspects of the study include the involvement of low health-literacy adults in design research, the capture of data on time, place, and social context of eating and sedentary behavior, and reporting prompts tailored to an individual's location and schedule.

Trial Registration: ClinicalTrials.gov NCT03083964; https://clinicaltrials.gov/ct2/show/NCT03083964

(JMIR Mhealth Uhealth 2019;7(4):e10894) doi:10.2196/10894

KEYWORDS

mhealth; health status; obesity; ecological momentary assessment

Introduction

Background

Precision medicine is an approach to care that involves classifying individuals into subpopulations that differ in their susceptibility to a particular disease or in their response to a specific treatment [1]. Subpopulations can be defined by genetics, but they can also be defined by behavioral and environmental exposures that lead to differential responses to biomedical, behavioral, and environmental interventions. The latter targets might be referred to as precision health interventions and may be advanced by better measures of behavioral and environmental exposures [2]. Environmental exposures can be relatively constant or vary over relatively short intervals of time, which together make up what we will refer to in this study as situations.

Large-scale precision health research efforts are in final planning [3] or just underway [4] and include the measurement of real-world and real-time physiological data from sensors such as accelerometers, heart rate, and glucose monitors, among others. Although detailed physiological measures will help with early detection of changes in physiological states and thus improve prevention or early treatment, precision health will also require measuring the situations and behaviors that directly or indirectly affect physiology [5]. Social, behavioral, and environmental factors contribute as much or more to health and longevity as other major domains including medical care and genetics [6]. Thus, better measures of behavior and situations are the keys to not only understanding behavior but also precision health interventions and ultimately better health and longevity.

Furthermore, 1 of the techniques widely used to obtain a situational or contextual understanding of daily life includes experience-sampling methodology [7]. This method includes self-report measurement but in a form where a person responds to subjective questions multiple times a day. This technique has often been attributed to overcoming methodological problems owing to memory and recall [8,9]. Furthermore, this method has high ecological validity and supports within-subject investigations [10,11]. Previous health research has attempted to execute experience sampling on technology devices such as personal digital assistants and pagers. With advancing technology, experience sampling has also been executed on mobile devices such as mobile phones. Often referred to as ecological momentary assessment (EMA) in health research, EMA is typically completed as persons experience something in their natural environment.

EMA is considered the gold standard of experiential sampling in health research [8]. However, self-report through EMA can be a burden, given the need to administer instruments multiple times in a day. Furthermore, the collected data can suffer from poor adherence and misreporting, especially if the instrument is cumbersome to use or does not suit individually-variable reporting needs and preferences (eg, sleep and work schedules and location triggers) [12]. This motivates a need for sampling tools that not only support situation-dependent, real-time self-report multiple times in a day but also are (1) low burden, (2) supportive of recurrent use, and (3) tailored to users' needs.

Objectives

This paper addresses these needs with a case study involving self-report measurement of location and social situations accompanying eating or sedentary behavior. This work was carried out in the context of a randomized trial among middle-aged, obese women cared for in a safety-net health system (NCT03083964) [13]. We report here the design and implementation of self-report measures of eating and movement behavior specific to users' location and social contexts. Each measure was developed through an iterative, user-centered design process involving obese, middle-aged women and deployed in a field trial to establish usability. The specific contributions of this paper are (1) a series of design constraints identified as important to consider and satisfy when designing mobile EMA interfaces, which are personalized to users' preferences, (2) 6 refined measures for self-report of eating and sedentary behavior, specific to location and social context using a mobile device, (3) a characterization of the ways in which individuals prefer to self-report eating and movement, along with perceived benefits and challenges of this self-reporting, and (4) field trials of feasibility with attention to response rates.

Methods

Overview

This study used an iterative user-centered research and design approach that comprised 4 phases (Table 1) to support the design and development of a mobile app. All sessions were audio and video recorded. Recordings were reviewed by stakeholders and designers before making design modifications to the EMA system. Qualitative thematic analysis was performed to identify and iteratively refine themes. This included 1 researcher coding and analyzing data using ATLAS.ti 8 Mac (ATLAS.ti). The codes and themes were reviewed, iteratively rectified, and agreed upon in consultation with other researchers on the team. The findings of Phases 1 and 2 were paired with design literature to guide the development of prototypes for evaluation in Phases 3 and 4 field trials. Overall, iterative participatory design and review sessions helped progress the identified measures from low-fidelity paper sketches to high-fidelity prototypes.

Table 1. Iterative and participatory user research and design of ecological momentary assessment system.

Phase	Research method	Duration	Stakeholders/users, n
Phase 1: Exploratory ideation	Focus group with stakeholders; 1-1 design session with users	60 min; 4560 min	6 stakeholders; 5 users
Phase 2: In-lab evaluation	Scenario-based think-aloud usability evaluation	4560 min	6 users
Phase 3: Field Trial 1	User evaluation of ecological momentary assessment (EMA) system in the field (feasibility test)	4 weeks	21 users
Phase 4: Field Trial 2	Response rate comparison of 2 versions of EMA system	4 weeks	38 users

Setting and Users

The target users of the EMA system live within a single city-county area of the Midwest. This study recruited patients aged 35 to 64 years and currently receiving care in 1 of Eskenazi Health's federally qualified health centers (FQHCs). Eskenazi Health is 1 of the 5 largest safety-net health systems in the nation. FQHCs gave us access to obese, middle-aged women who had had a provider-referral to the Healthy Me program. A primary care provider may refer a patient with a body mass index (BMI) of 30 or higher to meet with a Healthy Me coach. In this clinic-based program, health coaches counsel adult, obese patients and create an action plan for increased physical activity and encourage patients to make healthy food choices with an emphasis on controlling portions of food they consume. Healthy Me coaches are certified in behavior change counseling and fitness instructions, and they are present 2 or more days per week in each of the FQHCs [14].

Phase 1: Exploratory Design Ideation

Phase 1 involved ideation and designing of self-report measures on the basis of existing literature and user and stakeholder requirements. It included (1) a focus group session with the research team's primary stakeholders (social scientist, exercise physiologist, visual communication experts, and personal trainers) and (2) a 1-on-1 design session with 5 Healthy Me patients (P1-P5).

Methods

The focus group session lasted for an hour and identified the core self-report measures of behavior and situation specific to the project—food and drink consumption, physical movement, users' location, and social copresence. The focus group session also resulted in the development of 4 questions, each representing a single measure (version 1, V1). Questions in V1 were straightforward and required selection from a list of prepopulated response options. For instance, questions on eating/drinking and social copresence included options *yes* and *no*, whereas questions on location and physical activity included options created by stakeholders from expert knowledge and previous literature.

Following the initial focus group session, we conducted 1-on-1 design sessions with 5 Healthy Me patient volunteers (Participant 1 through Participant 5 or P1 through P5). Participants for this session were identified through announcements by coaches at the end of Healthy Me classes. The 45 to 60 min design session had 2 parts. During the first part, participants were presented with the problem domain and V1 questions created during the focus group session. This was followed by a brainstorming phase, where participants helped identify components for version 2 (V2). Each participant was compensated with a US $25 gift card at the end of a session.

Results

The fundamental principles of human-computer interaction informed us in designing an interactive prototype to support a user in responding to V2 EMA questions. As such, differences between V1 and V2 included refined and context-based response choices with images, icons, and a time component to every question to aid users' recall from within a referenced time frame. Specifically, V2 questions included (1) context-based suggestions for location in addition to search option (eg, prepopulated options to choose from, on the basis of the device's current location), (2) simplified response choices for physical activity, (3) detailed response choices for eating behavior to help users provide more detailed data (eg, meal or snack as opposed to a *yes*), and (4) detailed response choices for social copresence.

Phase 2: In-Lab Evaluation Studies

This phase involved evaluation of V2 measures identified in Phase 1.

Methods

In this phase, 6 female Healthy Me patients aged 35 to 64 years (50% African American, 50% nonHispanic white), recruited through snowball sampling, evaluated V2 measures. We refer to these participants as P6-P11. Each session lasted 45 to 60 min, where participants qualitatively evaluated and provided feedback on V2 questions delivered through an interactive, mobile prototype. Findings from this session resulted in the development of version 3 (V3) questions.

Results

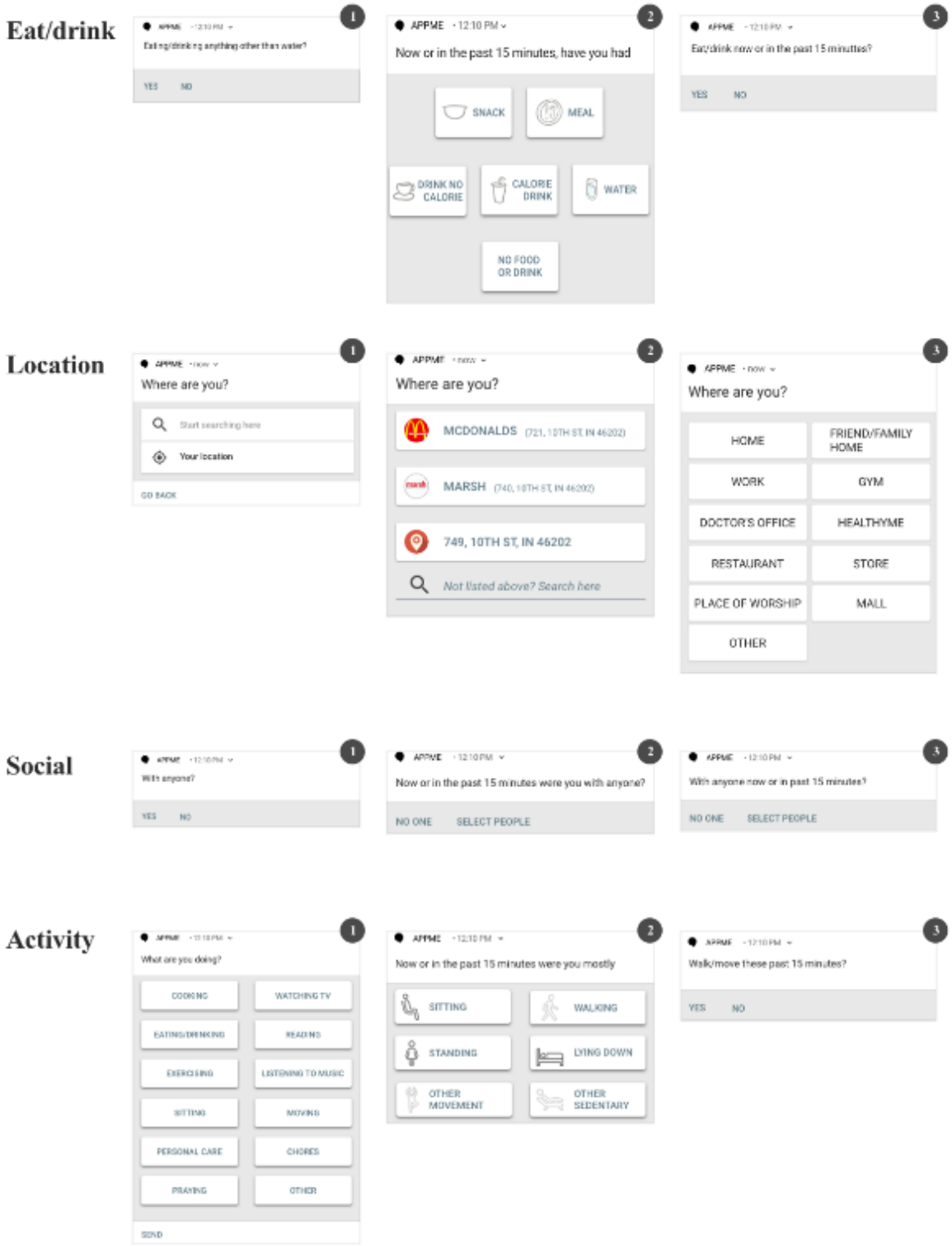
Some of the findings from the in-lab evaluation session included mixed support for image icons placed next to each response option and the need for simplified question and response options. For instance, although some participants found the prototype intuitive, others were confused with associating an image icon with the response option. This led to an executive decision on rolling back or excluding image icons as the immediate project scope was to design a mobile EMA tool supporting the core qualitative findings from Phases 1 and 2. Overall, the differences between V2 and V3 questions included excluding image icons, refined questions with fewer words and a time frame reference, and refined response choices that allowed the user to maintain privacy. Specifically, V3 questions included (1) response choices that helped users select the type of location as opposed to providing the actual physical address, (2) simplified response

choices for physical activity that focused on users reporting whether they had walked or not, (3) simplified response choices for eating behavior that direct users to consider everything other than drinking water as an eating event, and (4) means to prepopulate social connections with names or relationships.

Summary of Findings From Phases 1 and 2

Phase 1 (ideation) and Phase 2 (in-lab evaluation) work produced the 6 qualitative themes below that guided our design decisions on questions and response choices among versions V1, V2, and V3 (Figure 1).

Figure 1. Three versions of ecological momentary assessment questions resulting from iterative design process.



First: Privacy

Most participants disliked the idea of sharing their location with a technology solution:

> *My husband is very paranoid when it comes to technology and often says to me that someone is watching us through it. I would rather not share my exact address with the app.* [P3]

Participants attributed this feeling toward a lack of trust as to where their data were stored and who had access to them. The following is an example:

> *I don't know who is going to look at this data. It is like I am being policed* [P1]

Such participants suggested reporting an approximate location by selecting names associated with the location's address:

> *I don't mind saying I am at a restaurant as opposed to reporting I am at the McD on the corner of 16thstreet, if you know what I mean.* [P11]

Second: Active or Sedentary

Participants reported difficulties in identifying if they had or had not been moving enough to self-report physical activity while responding to an EMA question:

> *If I were sitting down but moving my arms up and down, wouldn't that still count as movement?* [P4]

A few participants valued and thought an interface would be more intuitive and natural if the question included images depicting the action that was of interest:

> *It would be easier for me to remember what I was doing when I see a picture of someone performing the activity, like a walking figure.* [P2]

On the other hand, some participants reported they would be confused if they saw images with actions that represented an unfamiliar action:

> *I see a picture of someone standing on one leg and it says here other movement. Is this yoga or dancing?Perhaps something else?* [P10]

Participants felt strongly about the need for a simplified and direct list of options from which to choose when reporting activity:

> *If all the app wants to know is if I moved or not, why not just ask that as opposed to asking me more details?* [P2]

Third: Reporting Food/Drink Consumption

Participants were very diverse in the ways they thought about and wanted to report food or drink consumption. Although some thought water would fall under food/drink self-report, others considered water being independent of any other food or drink they consumed:

> *I wouldn't want to report when I had water to drink. I am drinking water all day. The app would think I am eating or drinking something all day then.* [P5]

In such cases, participants suggested improving clarity on the question to help remind them not to think about water while self-reporting food/drink consumption:

> *Maybe the question can be–did you eat or drink something other than water?* [P8]

Several participants suggested improving their recall ability by including more response options:

> *It would be more easier if the question listed out options like snack or meal for food and maybe if it was diet or normal drink to help me remember. Yes or no is fine, but it would help if I saw better options.* [P6]

However, a few other participants felt they would face a challenge identifying and associating a category to whatever they had consumed. Several participants doubted the accuracy of this technique:

> *I think half a sandwich is a snack, that is just me. But I am not sure if that is a snack for others.* [P9]

Fourth: Who Is Around Me

Participants discussed the need for maintaining the privacy of their social connections while using the app. Some felt strongly about not sharing the name of the person they were spending time with when responding to the EMA questions as they thought they were violating others' privacy without their consent:

> *I don't want to tell some app Sue is with me without her consent.* [P4]

In such cases, participants suggested means for the app to allow selecting a relationship type as opposed to a person's name:

> *Why don't I say my sister is with me instead?* [P4]

Several participants reported increased desire and flexibility in prepopulating the names or relationships of most people with whom they interacted. This prepopulated list was used to create response options for the social copresence question. Participants appreciated reporting *other* in instances the response options did not hold a specific name of a person or relationship:

> *...it is easy to choose other when I don't see the name of the person who is with me.* [P7]

Fifth: Quick Interaction

All participants strongly preferred selecting a value through a single tap:

> *I like how I tap and can move onto next question. This is really quick.* [P8]

That being said, some participants discussed instances when they thought a single tap to move onto next question could be error prone:

> *This is really quick. But, what would happen if I tapped yes to the eating question by mistake? Can I go back and re-do the selection?* [P11]

Sixth: Time-Based Reporting

Several participants expressed concern toward disruptions at times they prefer not being disturbed (eg, while asleep):

> *Will I be receiving these messages during the night? I don't think I can respond. Can I get these questions when I am awake?* [P2]

Some participants felt strongly about the need to restrict the time frame on the basis of which they recalled before self-reporting:

> *I think it will make more sense if the app asked me what I did in the past 10 or 15 minutes* [P6]

Some participants rejected the time-oriented interface and strongly preferred a more relaxed qualitative instrument or something they can tune to their own subjective experience:

> *It is hard to know what I was doing exactly 15 minutes ago.* [P8]

These participants explained that the more relaxed the time frame such as since the last time they responded would enable them to be more accurate in remembering from the last time they responded:

> *I will remember what I said last time to the app.* [P9]

Discussion

The 6 qualitative findings from the exploration (described above) and in-lab evaluation phases were coupled with self-report and mobile usability literatures [15-24] to create 5 core design constraints (described below) that guided our design decisions for phases 3 and 4 (Table 2).

Design to Support Rapid Recall

A significant portion of Phase 1 and Phase 2 findings included discussing the need to reduce effort during recall. First, participants discussed the need for aiding in rapidly recalling their behavior on the basis of simplified response choices. Second, although the stakeholders discussed the need for capturing responses that represent behavior in the time between 2 EMA prompts, the Healthy Me patients expressed confusion in determining a time window they had to use as reference while performing a recall. Discussions included confusion when users skip/miss EMA prompts and the need for a time window as reference to help recall information. For instance, P1 stated the following:

> *What would happen if I didn't respond all day and see something at the end of the day?*

P3 stated the following:

> *Do I respond to the question considering what I was doing at the moment I saw it or am I responding based-off of something in the recent past?*

Design for Low Effort From User

Self-report can be a burden given the need to respond multiple times a day and while in the midst of daily activities. The need to respond to a notification can force users to interrupt an ongoing activity, consequently leading to increased burden/frustrations. This is true especially when the task of responding is effortful. Given this, it was a critical need for the system to require minimal effort while responding to a question post recall.

Table 2. Design decisions on the basis of constraints.

Constraint	Design decision
Design to support rapid recall	Rapid recall can be supported through the provision of a reference time frame where users can perform recall by focusing on the time window.
Design for low effort from user	A system that is capable of sending self-report questions as a group of message notifications can allow users to respond on the go, with a single tap on a mobile device. The burden can be further reduced if the group of messages is dependent on the users' context. For instance, the EMA[a] system can skip asking about food/drink consumption if the user is physically located in a restaurant when a self-report message group is sent, or the EMA system can provide suggestions for location on the basis of the device's location to help users avoid searching for an actual address.
Design to capture situations that accompany a behavior	To ensure a response is captured close to a situation, the EMA system should not allow users to respond after a set number of minutes have passed as it is likely the context of the user changes over time. To support this need, participants suggested a 30-min window for capturing responses and context, that is, a notification with a question disappears from the user's screen if the user does not respond within 30 min of receiving the notification.
Design to capture better quantity and quality of data	To maximize data capture, the EMA system can prompt a response at times when the user is awake and does not want to be disturbed. Furthermore, 1 way to personalize this experience for users is to have an onboarding process where the users can set preferences for when they will typically be available to receive EMA questions. Moreover, 1 way to capture data with improved quality, specifically for eating behavior, is to include response options where users can choose the type of food or drink. Similarly, social copresence can also include detailed response choices, which users can prepopulate with their social connections before EMA usage and select those when prompted for social copresence.
Design for user's safety	A notification can be held back if the EMA system identifies the user is moving in a vehicle. This avoids putting the user in danger.

[a]EMA: ecological momentary assessment.

Design to Capture Situations That Accompany a Behavior

It became clear from the ideation sessions that users are present in dynamically changing contexts of time, space, activity, and social connection. Hence, the EMA system should not only capture self-report measures but also the context or situations that entail particular user responses.

Design to Capture Better Quantity and Quality of Data

Both Phase 1 and Phase 2 participants pointed to the need for capturing better quality, in addition to quantity of data. Stakeholders discussed the need for capturing detailed information on social copresence to help identify patterns in eating or movement behaviors specific to social connections. Similarly, several participants from the 1-on-1 design sessions expressed concern about the lack of detail in reporting eating behavior. As such, all participants had differences in perception in determining when they have had something to eat or drink. For instance, P2 stated the following:

> *If I had a piece of candy do I still report yes to this question? I don't think that is eating really.*

P4 stated the following:

> *I usually know when I have eaten a meal, which can be a bigger portion size, as opposed to something like a small snack. It would be nice if I can report clearly what I had to eat because otherwise the app is going to think I am eating something all day.*

Design for Users' Safety

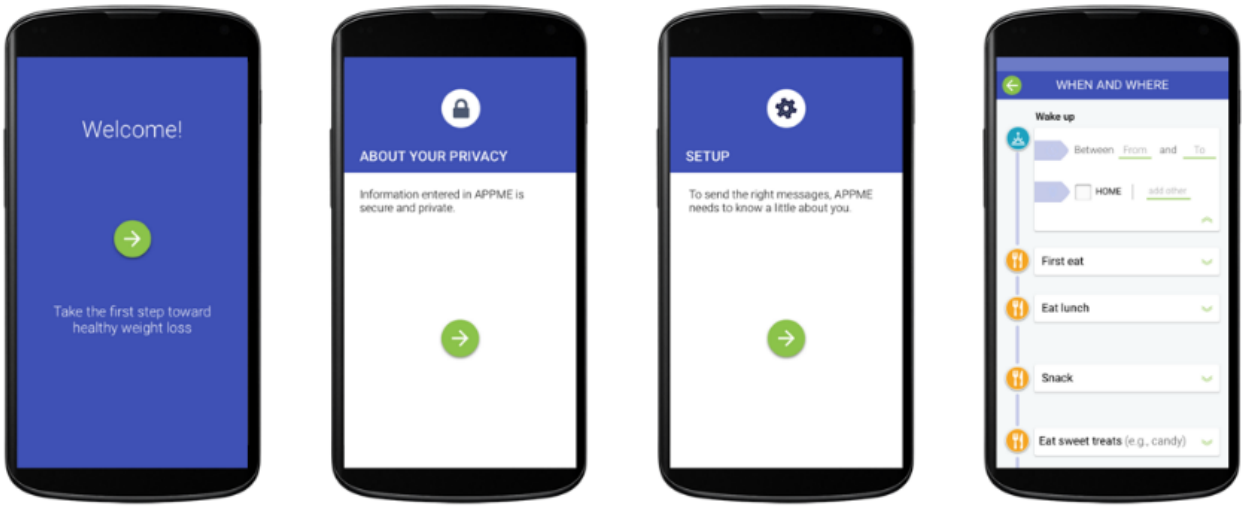
Participants from both phases raised the issue of impacting user safety in instances the user is required to, but is unable to, respond to message notifications. Hence, there is a need to design for safety and avoid penalizing a user for a nonresponse at unsafe moments (eg, while driving).

Ecological Momentary Assessment System: V3-Simple

From the 6 qualitative findings and 5 design constraints, an interactive and functioning EMA app was designed and developed to run on an Android smartphone. We chose to develop an Android app owing to (1) the availability of reusable Google libraries and services that can help identify the device's location or movement, (2) access to open-source code in the Android ecosystem that can be reused to save engineering effort, and (3) engineers' existing knowledge of Android app development practices. Although we expect modifications to the user interaction design, we do not anticipate modifications to the design constraints if this tool were to be deployed in a different smartphone such as an iPhone to suit operating system and hardware needs. To achieve maximum timely awareness and context learning, this system also included an onboarding process (Figure 2).

The onboarding process included users performing a 1-time setup by selecting approximate times when they woke up, ate, slept, and wished to not be disturbed. We envisioned this information to guide the timing of EMA questions for each participant. The EMA system was programmed to begin sending EMA questions the day following the onboarding setup to ensure the complete capture of data in a day. Any EMA question was designed to mimic a notification as identified by Google Material Design and comprised components such as header, content, and action buttons. Furthermore, EMA questions were developed to always appear first in the device's notification drawer (Figure 2) and included playing a sound clip (device's default for receiving a notification) every time a question was received on the device. The system was programmed to send EMA questions in groups. Overall, 12-question groups were prescheduled at the beginning of the day for every participant. Each response notification was timed to occur within the waking time and not during sleep or *do not disturb* times as selected by a participant during the onboarding process. By default, each message group comprised the 4 EMA questions regarding location, social presence, eating, and movement behaviors. However, the system removes EMA questions from a group on the basis of the participant's context, such as device's movement and location. The logic rules that guide the system and ultimately the user experience are shown in Figure 3.

Figure 2. One-time onboarding screens (top) and example ecological momentary assessment question in device's notification drawer (bottom left) and structure of an example ecological momentary assessment question (bottom right).



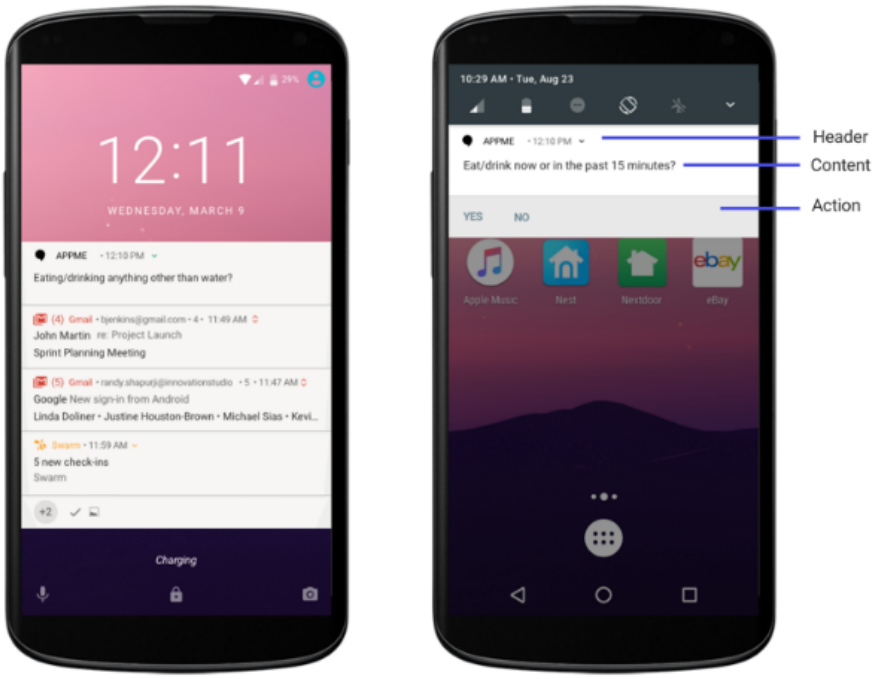
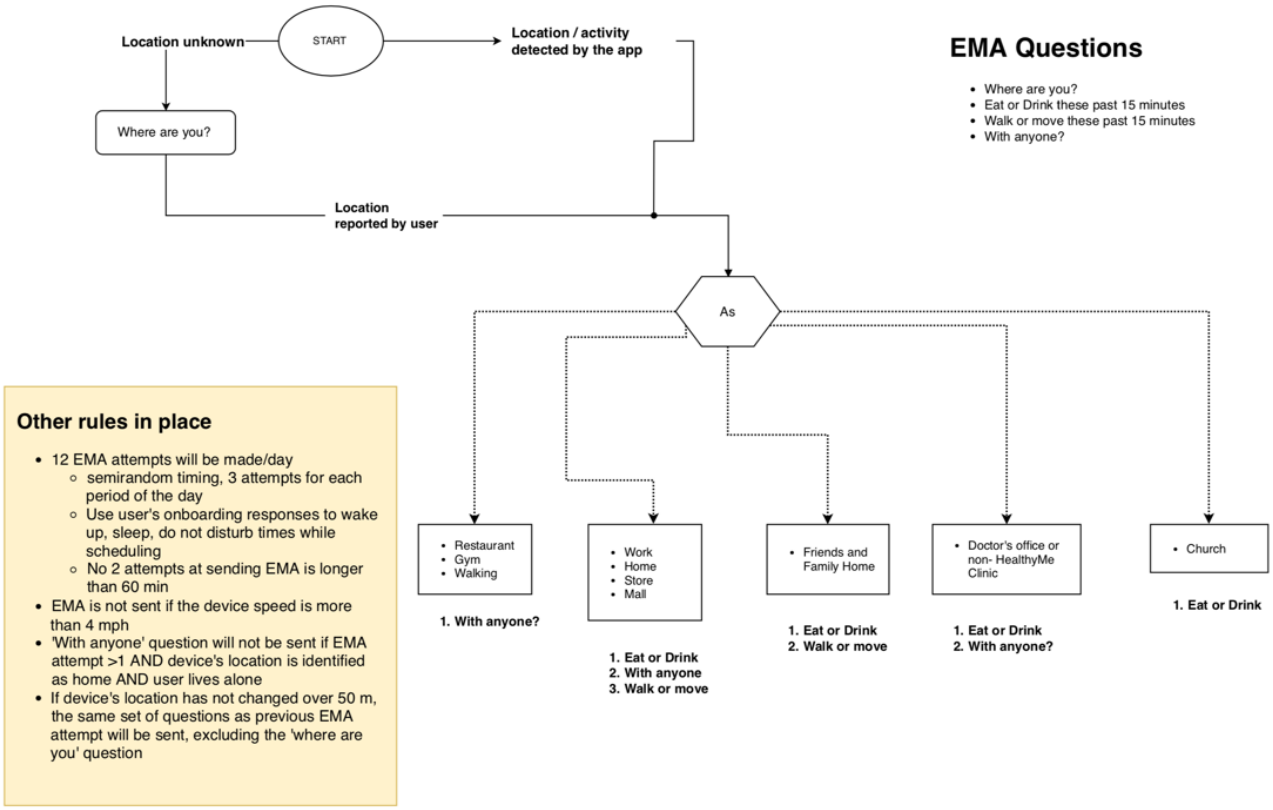


Figure 3. Flowchart depicting the logic used to identify the ecological momentary assessment questions in a group. EMA: ecological momentary assessment .



Identifying the Device's Location and Movement

In general, the system uses a phone's sensors and Google's movement and location identification services to identify the device's movement and location, respectively. For locations not in Google's service and particularly the participant's home, the system uses the following approach. The location *Home* is self-labeled by the participant (with help from study staff as needed) during onboarding. This system uses the recorded location coordinates within a 50-meter radius to identify whether or not a device is at the participant's home. For instances when the system is unable to locate the device, the participant's response to the location question is used to present appropriate follow-up EMA questions specific to that context. Any location coordinate recorded by the system is anonymized and stored in an encrypted, highly secure study server. This functionality of the system was communicated to all the participants (during the consent process), with participants having the additional option to remove their home location coordinates from the system if they changed their mind during the study period.

Phase 3: Field Trial 1

This field trial included testing the feasibility of the EMA system (V3-simple), with a focus on participants' understanding of EMA questions and response choices.

Methods

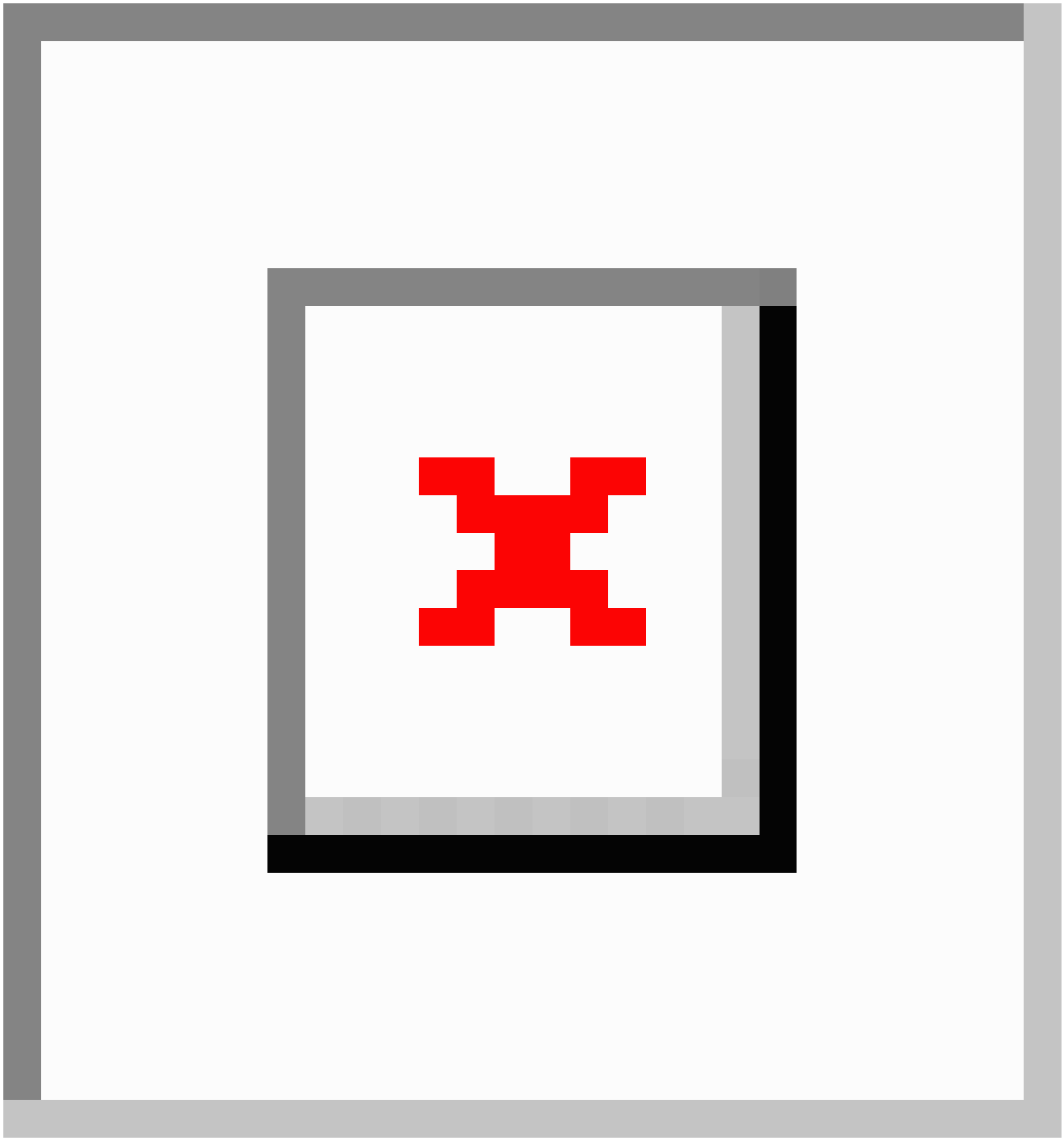
Participants for Phases 3 and 4 were recruited and enrolled following the human subjects' protection approval and protocols of the larger trial under which this work occurred [13]. This included an incentive of 10 cents for each completed EMA question, up to a maximum value of $10 over 4 weeks. Before the EMA trials, research assistants completed an in-home baseline assessment that captured sociodemographic and BMI data, and the Newest Vital Sign score for health literacy [25] data (see Table 3). During the 4-week trial, 21 Healthy Me patients (P12-P33) received 12 EMA message groups per day, every day on the basis of their onboarding times. The actual EMA questions for each group were determined on the basis of the phone's context and participant responses to the location question. Each EMA question behaved similar to a short message service notification in a smartphone, with a sound alert every time a question was received on the phone. This field trial included gathering qualitative feedback on the question and response choices, overall satisfaction [26], and perceived interface quality [27]. Table 3 shows the characteristics of participants in Field Trial 1 as well as Field Trial 2.

Table 3. Descriptive statistics for participants in field trials 1 and 2.

Participant characteristics	Field Trial 1 (n=21)	Field Trial 2 (n=38)
Age (years), mean (SD)	52.2 (7.6)	52.4 (8.5)
Race, n (%)		
Black or African American	16 (76.2)	33 (89.2)
White	5 (23.8)	4 (10.8)
Number of households, mean (SD)	2.2 (1.5)	2.8 (1.9)
Household income, mean (SD)	20,480 (23,022)	19,088 (12,771)
Education level, n (%)		
High school	7 (33.3)	1 (2.7)
College/university	14 (66.7)	22 (59.5)
Work status, n (%)		
Not working	17 (81.0)	14 (37.8)
Working	4 (19.0)	23 (63.9)
Hours worked, n (%)		
1-10	1 (25.0)	13 (36.1)
11-20	1 (25.0)	2 (15.4)
21-30	1 (25.0)	11 (84.6)
31-40	1 (25.0)	7 (53.8)
Shift work, n (%)		
First shift (6-8 am)	2 (50.0)	7 (53.8)
Second shift (2-5 pm)	__[a]	3 (23.1)
Varies	2 (50.0)	3 (23.1)
Weight (lbs), mean (SD)	257.0 (57.6)	250.9 (56.8)
Body mass index, mean (SD)	45.5 (8.4)	45.2 (10.7)
Low health literacy, n (%)		
No	11 (52.4)	12 (32.4)
Yes	10 (47.6)	25 (67.6)

[a]Data not available.

Figure 4. Mean weekly response rates for participants using version 3-simple in Field Trial 1.



Results

This trial was also used to determine participants' response rate to the EMA system. A common finding on reviewing the quantitative data suggested that all participants who began responding to the first question in a message group continued to complete all the questions in that group. Hence, the presence of a response to the first question within an EMA message group was used to calculate a participant's response rate. The overall response formula was: total number of first question responses over the total number of EMA question groups.

On the basis of this formula, the overall average response rate for participants in Field Trial 1 was identified as 50.3%. The mean weekly response rates were consistently close to 50% (Figure 4). At the end of 4 weeks, participants reported on Likert scales (1=strongly agree to 7=strongly disagree) delivered by research assistants, which captured overall satisfaction, perceived likability, and pleasantness of the smartphone app. Data captured on paper-based survey forms were transcribed and stored in a secure database. As such, participants reported being satisfied ($M_{Satisfaction}$=1.3) and expressed pleasantness and likability ($M_{Interface\ quality}$=1.4).

A common qualitative finding from this trial included participants' misunderstanding of the EMA questions, especially when they perceived a group of questions to be related to 1 another:

> *When I saw the "with anyone" question after the eating questions, I thought the app was asking me if I was eating with someone.* [P30]

Feedback from Field Trial 1 also included the need for balancing repetitive reminders versus interruptions that may be caused by an ongoing activity. Although several participants suggested resending an unanswered EMA question every few minutes as a reminder, a few discussed their inability to respond, irrespective of the reminder as they could not interrupt the ongoing activity:

> *It would help if the phone dinged after a couple of minutes to remind me in case I missed hearing the first time* [P18]

> *It won't matter sending reminders really. My phone is not with me and I work in a warehouse with no network.* [P20]

As such, we identified an additional design constraint for the EMA system, namely *designing for increased engagement*. We hypothesized that engagement would increase from 2 additional auditory alerts on the mobile phone on receipt of an EMA question. That is, the phone will play an audio sound indicating that the first EMA question in a group is awaiting a response when a phone has received the question but when the user has not responded. In particular, the EMA system was programmed to sound an alert every 10 min if the mobile phone had an unanswered first EMA question in its notification drawer. As such, in Field Trial 2, we tested the hypothesis that sounding an alert every 10 min in a 30-min window of sending the first EMA question would increase the response rate as opposed to only sounding an alert once when the first EMA question is sent to the mobile device.

Phase 4: Field Trial 2

This field trial included testing the hypothesis from Field Trial 1 for its effect on the response rate.

Methods

We used 2 versions, V3-simple and V3-ding, for use by 2 groups of participants. In all, 38 Healthy Me patients (P34-P72, Table 3) were randomly assigned to use either V3-simple or V3-ding for 4 weeks. During the 4-week field trial, both groups of participants received 12 EMA prompts daily, on the basis of their onboarding times. Similar to Field Trial 1, the actual EMA questions received by participants were dependent on their phone's context or their response to the location question. All participants were instructed at study setup to consider each question independent of the other while responding, to overcome the confusion experienced by participants in Field Trial 1. Although participants using V3-simple received a single auditory alert, participants testing V3-ding received 2 additional auditory alerts when the first EMA question in a group was received by the mobile phone. The 2 sound alerts were spaced to occur at 10-min intervals when a participant did not respond. Both versions included the same group of questions seen by participants in Field Trial 1. This trial also included gathering qualitative feedback (semistructured interview or occasional phone conversations where the research assistants assisted participants during the trial when needed) on repeated notifications in addition to perceptions on question and response choices either at the end of the trial or during the trial at times when the participants needed assistance.

Results

We calculated the response rate using the formula identified in Field Trial 1. Response rates of the participants using V3-simple were compared between trials 1 and 2 by means of a linear model with the independent variable representing mean response rate of Trial 1 versus Trial 2. In this model, mean response rates of participants using V3-simple between Trials 1 and 2 were not significantly different (P=.16).

V3-simple is compared with V3-ding in Figure 5. We found a statistically and significantly higher response rate for V3-ding respondents. However, a greater percentage of participants worked in the V3-ding group as opposed to the percentage of participants in the V3-simple group (P=.03). Hence, the statistical test comparing the response rates between V3-simple and V3-ding groups in Field Trial 2 was estimated again with an adjustment for work status. This adjustment did not change the findings. There was a significantly higher response rate for participants in the V3-ding group than the V3-simple group (Table 4).

Figure 5. Chart depicting mean weekly response rate comparison between groups version 3 (V3)-simple and V3-ding for participants in Field Trial 2. V: version.

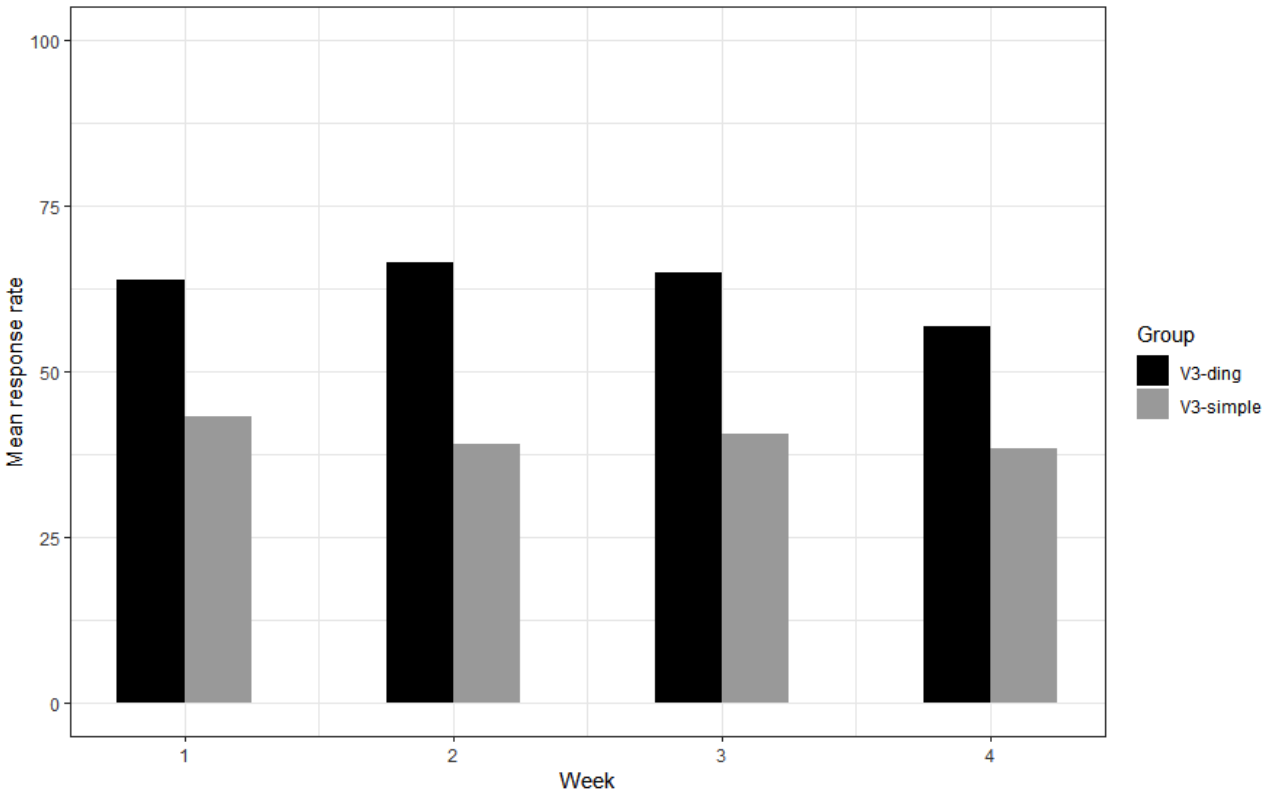


Table 4. Estimated means for response rate adjusted for group and work status.

Treatment	Estimate	Standard error	*P* value	Difference	95% CI
V3-simple	40.35	4.66	<.01	–19.96	–34.76 to –5.16
V3-ding	60.32	5.50	__[a]	—	—

[a]Not applicable.

The qualitative and quantitative findings pointed to a decline in the response rate with weekly progression. For instance, the qualitative reports on V3-ding indicated that some participants considered repeat sound alerts as nondisruptive whereas others perceived this feature as a drawback:

> *It is not a big deal.* [P45]
>
> *It felt like the system was eating up my phone's battery.* [P60]

This feedback concurs with the occurrence of a statistical difference in the response rates between weeks 1 and 4 (*P*=.03).

Another common finding included participants being confused between whether they had to consider drinking water as a drinking activity:

> *Do I select yes when I have had water to drink?* [P38]
>
> *Water is a drink too. I reported yes every time I had just had water when I received eat question. I hope the system doesn't think I am always eating or drinking something.* [P42]

Discussion

Our aim in this research was to identify the constraints and design parameters for a mobile EMA system capable of capturing self-report measures of eating and movement behavior coupled with participants' location and social contexts. We viewed the capture of these and similar health-related behaviors as they occur in everyday contexts as critical to progress in precision health research and interventions. Personalization of EMA processes may improve engagement in the data capture process and may also advance precision health efforts by facilitating just-in-time and just-in-place interventions [5]. By working with a diverse set of individuals, we were able to design an effective reporting interface. Following an iterative exploration and ideation process, we identified the constraints that users and stakeholders reported to be particularly important. We evaluated the interface and the self-report measures through 2 field trials. Although the first trial helped identify feasibility, the second trial tested a design feature that improved user engagement. Our interface was well received with high satisfaction and interface quality ratings.

Our target users were female patients with obesity seeking care in a safety-net health system; most had very limited household

resources or technology experience and low health literacy. Achieving a satisfactory interface for real-world, real-time self-reports of weight-related behaviors and context required significant time with users and many design iterations and prototypes over a 6-month period. From this extensive effort, we suggest that future work attempting to conduct momentary assessments should achieve the following minimum specifications to support increased user engagement: (1) an onboarding process to personalize the times when an assessment is delivered to a participant, (2) a question pruning through passive sensing of device location to present only questions most relevant to a participant's context, (3) increase participant's attention with a notification that is prioritized over other app notifications in the receiving device, (4) limit the time window within which a participant can respond to a question to capture situations accompanying a behavior, and (5) repeat auditory alerts to remind participants to respond. Specific to reducing burden while capturing a user response, we suggest designing a system that (1) supports tap interaction to record a response, (2) uses simple-worded, direct questions with fewer words that are easier to read and quicker for the participant to understand, (3) has simple response options that are easier to read, quicker for the participant to understand and select from, and (4) includes instruction as to whether a participant has to consider each question in the assessment independent of one another while selecting a response.

Our Contribution in Comparison With Previous Work

Although participants in our study used descriptions of the behavior similar to the findings reported here [28], we found that using icons to visually represent target behavior can be difficult to understand. Overall, we highlight that personalizing or tailoring the EMA experience to an individual can have several positive outcomes. First, customizing the times for EMA prompts to suit an individual's daily routine can increase adherence. For instance, participants are more likely to not respond to an EMA question at times they do not want to be disturbed. Second, personalization can help to only present EMA questions that are most relevant to an individual's situation or context. For instance, our system can automatically identify an individual's location from the individual's smartphone and present only the most relevant EMA questions. This can reduce the response burden and better match a respondent's cognitive process for recall and self-assessment. Finally, our solution involving tap-to-report interaction highlights a design solution for reducing the response burden.

Limitations

A potential side effect of EMA indicated in existing literature [29] is that EMA systems may serve as an intervention and not just a measurement tool. We heard some feedback supporting/suggesting that this may be the case. For instance, several participants expressed becoming more aware of their eating and movement behavior after using our EMA system. In other words, although participants were self-reporting eating or movement behavior, they were also being nudged to pay close attention to these self-behaviors. To determine to what extent this affected behavior or weight over time would require work outside of the scope of this report. Similarly, it is also outside the scope of this report to carry out validation of the EMA data. Future work could investigate the validity of EMA reports of movement, as this can be compared with accelerometry data. Social copresence could theoretically be validated by corroboration of the copresent individual. In the case of eating, validation would be more difficult as there is no objective measure of eating beyond direct observation; however, we do have work in the field using 24-hour dietary recall and EMA eating questions in the same time window. Future work can investigate a hybrid human-reported and automated data collection system such as the one described here [30] to support users to provide useful insights into their behavior at times when the automated system is inactive.

Conclusions

With full appreciation of the potential limitations, we are intrigued by the possibilities of this EMA platform in 3 broad areas of future work. First, if asking an individual about a behavior frequently and in varied context has implications for the practice of that behavior, broad use of this EMA platform could result in an extraordinarily low-burden, low-cost, and highly scalable intervention tool. Thus, we believe the impact of the EMA reporting process on behavior, particularly weight-related behavior in our case, deserves careful investigation. Second, expanding this EMA platform to provide contextually appropriate feedback, whether human-reported or automated, is an interest of ours. Keeping to the design principles learned here while developing easy, simple, positively reinforcing feedback could enhance any effect EMA reporting may have on behavior. Third, the use of this tool to measure behavior and context for other projects and programs is significant as precision health advancements may well depend on EMA [31]. Examples in the literature point to exciting possibilities for work in smoking cessation [32], drug abuse and recovery [33-35], and mental health [36] to name a few. We are hopeful that ongoing work in these and other areas could lead to a new generation of behavioral data and interventions.

Acknowledgments

The authors wish to acknowledge the contributions from persons who provided key expertise and resources to this project. Cori McCorkle and Anietra Butler recruited study participants and completed assessments with those participants. Tony French and Andrew Martin provided computer programming expertise to the project. This study is funded by the National Institute of Health, Lung, and Blood Institute (R01 HL128494) and is approved by the Indiana University Institutional Review Board.

Conflicts of Interest

None declared.

References

1. National Research Council. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. Washington, DC: The National Academies Press; 2011.
2. Riley K, Kalichman S. Mindfulness-based stress reduction for people living with HIV/AIDS: preliminary review of intervention trial methodologies and findings. Health Psychol Rev 2015;9(2):224-243. [doi: 10.1080/17437199.2014.895928] [Medline: 26209210]
3. National Institutes of Health. The future of health begins with you URL: https://allofus.nih.gov/ [accessed 2019-02-18] [WebCite Cache ID 76GoGDQss]
4. Manos D. Rocking the baseline: Verily, Duke, and Stanford aim to make medicine more predictive with a new baseline study. Clinical OMICs 2017;4(3). [doi: 10.1089/clinomi.04.03.01]
5. Gambhir SS, Ge TJ, Vermesh O, Spitler R. Toward achieving precision health. Sci Transl Med 2018 Feb 28;10(430) [FREE Full text] [doi: 10.1126/scitranslmed.aao3612] [Medline: 29491186]
6. McGovern L. Robert Wood Johnson Foundation. A brief measure for assessing generalized anxiety disorder: the GAD-7 URL: https://www.rwjf.org/en/library/research/2014/08/the-relative-contribution-of-multiple-determinants-to-health-out.html [accessed 2019-02-18] [WebCite Cache ID 76GsZhGvi]
7. Hektner JM, Hektner JA, Csikszentmihalyi M. Experience Sampling Method: Measuring the Quality of Everyday Life. Thousand Oaks, California: Sage Publications; 2007.
8. Kahneman D, Krueger AB, Schkade DA, Schwarz N, Stone AA. A survey method for characterizing daily life experience: the day reconstruction method. Science 2004 Dec 3;306(5702):1776-1780 [FREE Full text] [doi: 10.1126/science.1103572] [Medline: 15576620]
9. Kahneman D. Thinking, Fast And Slow. New York: Farrar, Straus And Giroux; 2019.
10. Csikszentmihalyi M, Larson R. Validity and reliability of the experience-sampling method. In: Csikszentmihalyi M, editor. Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi. Netherlands: Springer; 2014:35-54.
11. Dockray S, Grant N, Stone AA, Kahneman D, Wardle J, Steptoe A. A comparison of affect ratings obtained with ecological momentary assessment and the day reconstruction method. Soc Indic Res 2010 Nov;99(2):269-283 [FREE Full text] [doi: 10.1007/s11205-010-9578-7] [Medline: 21113328]
12. Trull TJ, Ebner-Priemer UW. Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. Psychol Assess 2009 Dec;21(4):457-462 [FREE Full text] [doi: 10.1037/a0017653] [Medline: 19947780]
13. Clark D, Srinivas P, Bodke K, Keith N, Hood S, Tu W. Addressing people and place microenvironments in weight loss disparities (APP-Me): design of a randomized controlled trial testing timely messages for weight loss behavior in low income Black and White Women. Contemp Clin Trials 2018 Apr;67:74-80 [FREE Full text] [doi: 10.1016/j.cct.2018.01.006] [Medline: 29357313]
14. Clark D, Keith N, Chrysler L, Perkins AJ, Willis D. Design and reach of a primary care weight management program. J Health Care Poor Underserved 2008 Feb;19(1):171-179. [doi: 10.1353/hpu.2008.0014] [Medline: 18263993]
15. Conner T, Tennen H, Fleeson W, Barrett LF. Experience sampling methods: a modern idiographic approach to personality research. Soc Personal Psychol Compass 2009 May 1;3(3):292-313 [FREE Full text] [doi: 10.1111/j.1751-9004.2009.00170.x] [Medline: 19898679]
16. Gong J, Tarasewich P. Guidelines for handheld mobile device interface design. 2004 Presented at: DSI 2004 Annual Meeting; November 2004; Boston Massachusetts p. 3751-3756 URL: https://pdfs.semanticscholar.org/4350/ed959d6a638168ba24058892338e8ee37338.pdf
17. Hansson R, Ljungstrand P, Redström J. Subtle and Public Notification Cues for Mobile Devices. In: Proceedings of the 3rd international conference on Ubiquitous Computing. 2001 Dec 1 Presented at: UbiComp'01; September 30-October 2, 2001; Atlanta, Georgia, USA p. 240-246.
18. Hofmann W, Baumeister RF, Förster G, Vohs KD. Everyday temptations: an experience sampling study of desire, conflict, and self-control. J Pers Soc Psychol 2012 Jun;102(6):1318-1335. [doi: 10.1037/a0026545] [Medline: 22149456]
19. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
20. Macht M, Haupt C, Salewsky A. Emotions and eating in everyday life: application of the experience-sampling method. Ecol Food Nutr 2004 Jul;43(4):11-21. [doi: 10.1080/03670240490454723]
21. Saffer D. Microinteractions: Designing with Details. Sebastopol, California: O'Reilly Media; 2013.
22. Stone A, Kessler RC, Haythornthwaite JA. Measuring daily events and experiences: decisions for the researcher. J Pers 1991 Sep;59(3):575-607. [doi: 10.1111/j.1467-6494.1991.tb00260.x] [Medline: 1960643]

23. Tourangeau R, Rips L, Rasinski K. The Psychology of Survey Response. Cambridge, UK: Cambridge University Press; 2000.
24. Vannette DL, Krosnick JA. Answering questions: a comparison of survey satisficing and mindlessness. In: Langer EJ, Ngnoumen CT, Ie A, editors. The Wiley Blackwell Handbook of Mindfulness. New Delhi, India: John Wiley & Sons; 2014:312-327.
25. Weiss B, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3(6):514-522 [FREE Full text] [doi: 10.1370/afm.405] [Medline: 16338915]
26. Lewis JR. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. ACM SIGCHI Bulletin 1991;23(1):78-81. [doi: 10.1145/122672.122692]
27. Lewis J. Psychometric evaluation of the post-study system usability questionnaire. 1992 Presented at: The Human Factors and Ergonomics Society Annual Meeting; October 12-16, 1992; Atlanta, Georgia p. 1259-1260 URL: https://www.researchgate.net/publication/230786756_Psychometric_evaluation_of_the_post-study_system_usability_questionnaire_The_PSSUQ
28. Connelly K, Stein KF, Chaudry B, Trabold N. Development of an ecological momentary assessment mobile app for a low-literacy, Mexican-American population to collect disordered eating behaviors. JMIR Public Health Surveill 2016 Jul 14;2(2):e31 [FREE Full text] [doi: 10.2196/publichealth.5511] [Medline: 27418020]
29. 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]
30. Zhang P, Schmidt D, White J, Mulvaney S. Towards precision behavioral medicine with IoT: iterative design and optimization of a self-management tool for type 1 diabetes. 2015 Dec 1 Presented at: 2018 IEEE International Conference on Healthcare Informatics; June 4-7, 2018; New York, NY, USA. [doi: 10.1109/ICHI.2018.00015]
31. Juarascio A, Manasse SM, Goldstein SP, Forman EM, Butryn ML. Review of smartphone applications for the treatment of eating disorders. Eur Eat Disord Rev 2015 Jan;23(1):1-11 [FREE Full text] [doi: 10.1002/erv.2327] [Medline: 25303148]
32. Mitchell J, Dennis MF, English JS, Dennis PA, Brightwood A, Beckham JC, et al. Ecological momentary assessment of antecedents and consequences of smoking in adults with attention-deficit/hyperactivity disorder. Subst Use Misuse 2014 Sep;49(11):1446-1456 [FREE Full text] [doi: 10.3109/10826084.2014.912229] [Medline: 24827866]
33. Epstein D, Preston KL. Daily life hour by hour, with and without cocaine: an ecological momentary assessment study. Psychopharmacology (Berl) 2010 Aug;211(2):223-232 [FREE Full text] [doi: 10.1007/s00213-010-1884-x] [Medline: 20532873]
34. Epstein D, Tyburski M, Craig IM, Phillips KA, Jobes ML, Vahabzadeh M, et al. Real-time tracking of neighborhood surroundings and mood in urban drug misusers: application of a new method to study behavior in its geographical context. Drug Alcohol Depend 2014 Jan 1;134:22-29 [FREE Full text] [doi: 10.1016/j.drugalcdep.2013.09.007] [Medline: 24332365]
35. Preston K, Epstein DH. Stress in the daily lives of cocaine and heroin users: relationship to mood, craving, relapse triggers, and cocaine use. Psychopharmacology (Berl) 2011 Nov;218(1):29-37 [FREE Full text] [doi: 10.1007/s00213-011-2183-x] [Medline: 21336579]
36. Moore R, Depp CA, Wetherell JL, Lenze EL. 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]

Abbreviations

BMI: body mass index
EMA: ecological momentary assessment
FQHC: federally qualified health center
V1: version 1
V2: version 2
V3: version 3

Edited by G Eysenbach; submitted 01.05.18; peer-reviewed by B Chaudhry, E Hebert, B Smaradottir, S Chokshi; comments to author 25.08.18; revised version received 30.11.18; accepted 31.12.18; published 03.04.19.

Please cite as:
Srinivas P, Bodke K, Ofner S, Keith NR, Tu W, Clark DO
Context-Sensitive Ecological Momentary Assessment: Application of User-Centered Design for Improving User Satisfaction and Engagement During Self-Report
JMIR Mhealth Uhealth 2019;7(4):e10894
URL: http://mhealth.jmir.org/2019/4/e10894/
doi:10.2196/10894
PMID:30942698

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org