JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 5 (2017), Issue 11 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Original Papers

Original Paper

Detecting Acute Otitis Media Symptom Episodes Using a Mobile App: Cohort Study

Annemarijn C Prins-van Ginkel[1], MSc; Marieke LA de Hoog[1], PhD; C Uiterwaal[1], MD, PhD; Henriette A Smit[1], PhD; Patricia CJ Bruijning-Verhagen[1], MD, PhD

Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands

Corresponding Author:
Annemarijn C Prins-van Ginkel, MSc
Julius Center for Health Sciences and Primary Care
University Medical Center Utrecht
Huispost nr STR 6.131
PO Box 85500
Utrecht, 3508 GA
Netherlands
Phone: 31 887568044
Fax: 31 887568099
Email: A.C.Prins@umcutrecht.nl

Abstract

Background: Population cohort studies are useful to study infectious diseases episodes not attended by health care services, but conventional paper diaries and questionnaires to capture cases are prone to noncompliance and recall bias. Use of smart technology in this setting may improve case finding.

Objective: The objective of our study was to validate an interactive mobile app for monitoring occurrence of acute infectious diseases episodes in individuals, independent of health care seeking, using acute otitis media (AOM) symptom episodes in infants as a case study. We were interested in determining participant compliance and app performance in detecting and ascertaining (parent-reported) AOM symptom episodes with this novel tool compared with traditional methods used for monitoring study participants.

Methods: We tested the InfectieApp research app to detect AOM symptom episodes. In 2013, we followed 155 children aged 0 to 3 years for 4 months. Parents recorded the presence of AOM symptoms in a paper diary for 4 consecutive months and completed additional disease questionnaires when AOM symptoms were present. In 2015 in a similar cohort of 69 children, parents used an AOM diary and questionnaire app instead.

Results: During conventional and app-based recording, 93.13% (17,244/18,516) and 94.56% (7438/7866) of symptom diaries were returned, respectively, and at least one symptom was recorded for 32.50% (n=5606) and 43.99% (n=3272) of diary days (P<.01). The incidence of AOM symptom episodes was 605 and 835 per 1000 child-years, respectively. Disease questionnaires were completed for 59% (17/29) of episodes when participants were using conventional recording, compared with 100% (18/18) for app-based recording.

Conclusions: The use of the study's smart diary app improved AOM case finding and disease questionnaire completeness. For common infectious diseases that often remain undetected by health care services, use of this technology can substantially improve the accurateness of disease burden estimates.

(JMIR Mhealth Uhealth 2017;5(11):e181) doi:10.2196/mhealth.7505

KEYWORDS

smartphone; mobile app; infectious diseases, cohort studies, acute otitis media, underreporting, patient compliance; mobile applications; communicable diseases; otitis media

Introduction

A key issue in many prospective infectious disease epidemiological studies, both interventional and observational, is completeness of detecting disease events of interest among study participants. This is particularly true for events that cannot be comprehensively and reliably detected through health care-based research, such as self-limiting respiratory or gastrointestinal infections. For instance, in the Netherlands, it is estimated that health care encounters are not involved for up to 50% of acute otitis media (AOM) episodes in young children. AOM is therefore notoriously underdetected when health care contacts alone are relied on [1-3]. There is evidence that AOM poses a substantial burden, both on the child and the family, and results in economic losses due to workdays lost irrespective of health care use, thus requiring assessment of participant-reported disease events when population disease burden is estimated [1,4,5]. Moreover, health-seeking behavior is influenced by factors other than severity and could therefore introduce bias. In the field of infectious diseases epidemiology, this forms a major challenge in conducting research. For these reasons, there is a great interest in methods that can improve detection of infectious diseases events in epidemiological studies beyond the current health care-related scope.

Research on mobile apps for monitoring or promoting patient health is expanding rapidly, forming an entire new area within epidemiology. Most research focuses on evaluations of mHealth apps as an intervention in which apps are used to improve health [6,7]. However, well-designed mobile apps can also serve as valuable instruments to document health and disease among study participants. Up to now, apps have been little used as an alternative data-recording tool, while this technology offers several advantages over traditional methods of participant follow-up, such as paper, Internet, or telephone surveys. Interactive and dynamic features can be applied to improve participant compliance, detection, and disease burden estimation [7,8].

We performed a proof-of-concept validation study of an interactive mobile app for monitoring occurrence of acute infectious diseases episodes in individuals, independent of health care seeking, using AOM symptom episodes in infants as a case study. Our aim was to determine participant compliance and app performance in detecting and ascertaining (parent-reported) AOM symptom episodes with this novel tool compared with traditional methods used for monitoring study participants.

Methods

We compared use of traditional survey methods (paper diary sheets) with use of a smartphone diary app for the purpose of prospectively detecting AOM symptom episodes and measuring their disease burden in infants. We made comparisons by applying the different methods consecutively over 2 periods (in 2013 and 2015) of a nested AOM study within a larger ongoing birth cohort study.

WHISTLER Study

Both the first (2013) and second (2015) AOM study period were (partly) nested within the Wheezing and Illnesses Study Leidsche Rijn (WHISTLER) birth cohort that recruited healthy, term neonates between 2001 and 2012 to study perinatal and infant risk factors for wheezing illness. The study design and rationale of WHISTLER are described in detail elsewhere [9]. WHISTLER enrolled newborns before 2 months of age and living in the Leidsche Rijn district of Utrecht, the Netherlands. WHISTLER had traditionally been using paper diary sheets to monitor parent-reported respiratory symptoms during the first year of life. For this purpose, a diary sheet containing a list of respiratory symptoms printed for each day of the month was distributed monthly by mail. Symptoms included cough, wheeze, fever (>38°C), nasal cold, otorrhea, otalgia, and sore throat (Figure 1). Parents were asked to mark whenever a symptom was present on a particular day. Additionally, the diary sheet contained a monthly questionnaire about environmental risk factors of respiratory infections printed on the back. After the end of each month, parents mailed the filled-in paper sheet in return envelopes, and the sheets were scanned for digital processing [9].

Design of the 2013 Acute Otitis Media Study Period

In January 2013, we invited by mail 300 participating WHISTLER parents (randomly selected out of 594 parents) with children aged 0 to 3 years for additional participation in our nested AOM study. According to the WHISTLER traditional method, parents reported respiratory symptoms and answered the monthly questionnaire using the paper diary sheets. For children older than 1 year, we requested parents to restart recording during 4 consecutive months (February to May 2013), as symptoms were recorded only in the first year of life according to the WHISTLER protocol. In addition to the daily recording, parents were asked to contact the study team within 24 hours by email, text message, or telephone call when they recorded a combination of symptoms suggestive of an AOM symptom episode. For this, parents received detailed instructions upon enrollment explaining which (combination of) symptoms was suggestive of an AOM symptom episode and should prompt notification. Researchers verified these symptoms over the telephone. Subsequently, we asked parents to complete paper versions of a validated AOM severity score (AOM-SOS) during 7 consecutive days and an additional disease questionnaire on day 7 [10]. Questionnaires and scoring lists were returned by mail. We used these to ascertain the AOM symptom episode by assessing AOM severity, health care use, and impact on family life.

Design of the 2015 Acute Otitis Media Study Period

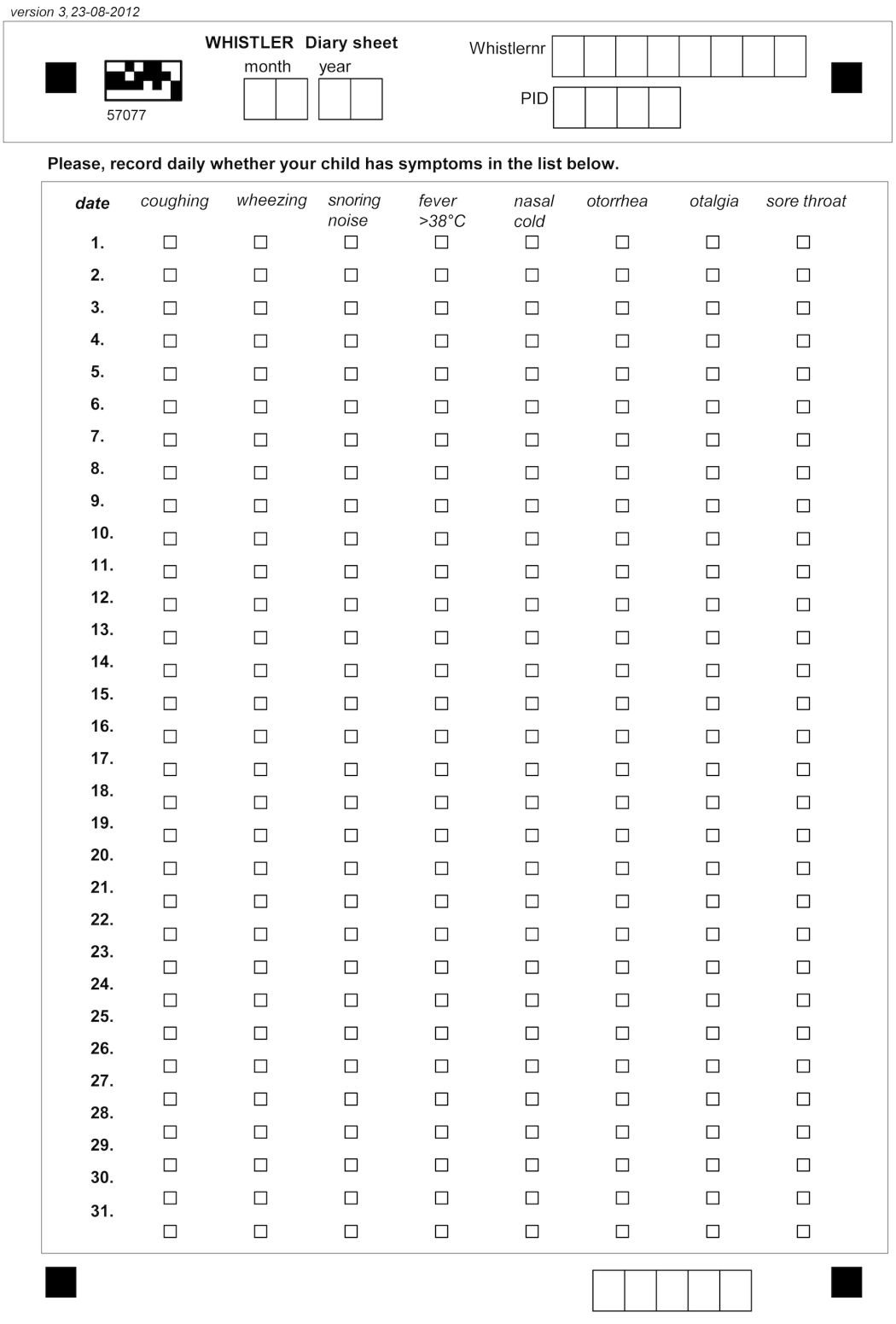
We invited 404 parents of 0- to 3-year-old children by mail to participate in the (nested) AOM study. As WHISTLER completed recruitment in January 2013, no infants under 1 year of age were participating in WHISTLER in January 2015. Thus, we invited 91 WHISTLER age-eligible (ie, <3 years) participants and an additional 313 non-WHISTLER participants aged between 3 months and 2 years who also lived in the Leidsche Rijn district. Instead of using the paper diary sheet, all participating parents were now instructed to use a mobile

device diary app, the InfectieApp, which we developed for this study, during 4 consecutive months (February to May 2015).

During each study period, we used identical criteria for occurrence of an AOM symptom episode: a combination of fever and either otalgia or otorrhea on the same day. This combination of symptoms had to be actively reported by the parents in the 2013 pilot study, while this combination was automatically recognized and reported using the InfectieApp in the 2015 pilot study.

Each period of the AOM study received separate approval by the medical ethics committee of the University Medical Centre Utrecht. Written informed consent was given by all participating parents.

Figure 1. Paper diary sheet used in the 2013 study period. WHISTLER: Wheezing and Illnesses Study Leidsche Rijn.



InfectieApp

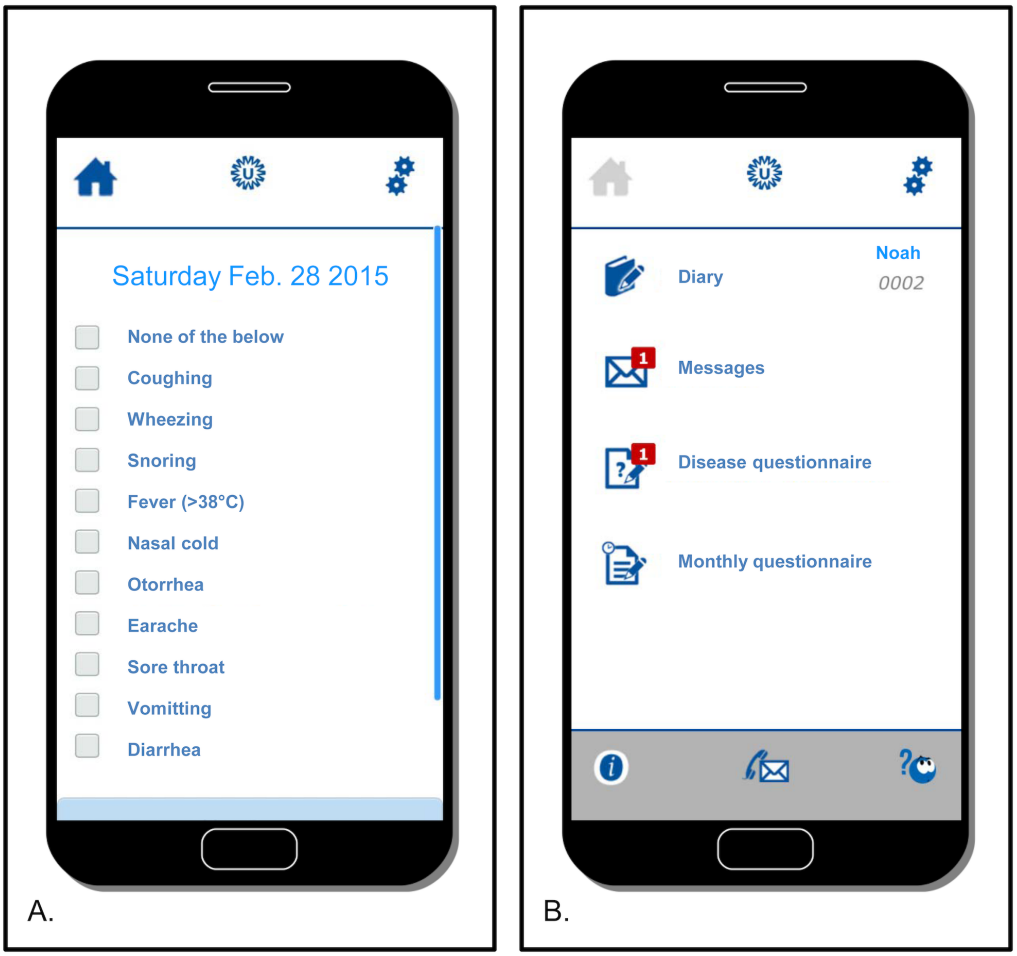
The InfectieApp software app was custom made in 2014 by University Medical Centre Utrecht in collaboration with VitalHealth Solutions, a company specializing in eHealth, based in Uddel, the Netherlands, and was compatible with iOS, Android, and Windows Mobile operating systems. The InfectieApp was developed for use by study participants to self-report 3 types of data: (1) symptom diary data (Figure 2, panel A), (2) monthly questionnaire responses on risk factors, and (3) disease questionnaire responses and scoring lists in the event of an AOM symptom episode. The symptoms listed in the app diary, disease questionnaire, and scoring list were identical to those in the questionnaires used in the 2013 study period. Participants could access additional study-related information in the app organized in various submenus, including general patient information related to the study, frequently asked questions, and contact details of the research team. All app content was organized in menus, accessible through the home screen (Figure 2, panel B).

The onset and ending of an AOM symptom episode was detected based on diary entries, using built-in algorithms: fever together with either earache or otorrhea occurring on the same day marked the onset of an episode. An episode ended when fever was not recorded for 7 consecutive days. Detection of a new-onset AOM episode triggered additional app content: participants received an app message explaining that AOM symptoms were detected and that additional questionnaires would follow in the coming days. For ascertainment of the AOM symptom episodes, the AOM-SOS scoring list and disease questionnaire appeared in the app questionnaire menu. The message also contained a link to a Dutch independent, professional, patient website on AOM where parents could read general AOM medical information [11-13].

To encourage participant compliance with diary recording, daily reminders at 8 PM appeared as push notifications on the smartphone. A diary that had not been filled in remained accessible to the participant up to 7 days after the diary date. We could also decide to contact the household by telephone.

The InfectieApp was password protected at first log-in; thereafter, the participant could use a 4-digit code to enter the app. When the app was used offline, data were stored encrypted in the InfectieApp. The recorded data were sent to the server via a secure connection (hypertext transfer protocol secure, HTTPS), meaning that the data were sent encrypted to the server. When the app was online, the recorded data were sent immediately to the server, providing us with the opportunity to monitor the participants in real time both for the occurrence of AOM symptom episodes and for compliance with the questionnaires.

Figure 2. Screenshots of the diary app (InfectieApp) used in the 2015 study period. (A) symptom diary; (B) home screen.



A security access layer determined which actions a user could perform, meaning that, depending on the rights assigned to the user, the user could or could not perform certain actions. This server was hosted by VitalHealth Solutions. The research team could access the study data stored on the server using a Web interface. Each researcher had access to the decrypted data using their personal username and passwords.

Data Analysis

For the primary outcome, we compared the proportion of AOM symptom episodes ascertained by complete disease questionnaires between both study periods. Next, we compared the number of diaries completed and the number of monthly questionnaires filled in. For the 2015 study period, we compared the difference in number of diaries and monthly questionnaires completed between the WHISTLER and non-WHISTLER participants.

In the 2015 study period, we could assess the number of diaries completed more reliably, because a confirmation of absence of symptoms was required (symptom checkbox "None of the below"; Figure 2, panel A), while no such checkbox existed on the paper diary sheet. We therefore decided to assess and compare diary completeness using 2 different outcomes. First, we compared the proportion of diaries returned by mail (first period) with the proportion entered using the app (second period). For this, we assumed that, for a returned paper diary sheet, each day of the month was completed, although a day without a reported symptom could represent a missing day. Second, we compared the proportion of days with at least one reported symptom in both study periods, thereby automatically excluding potentially missing days on diary sheets in the first study period.

We estimated the incidence rate of AOM symptom episodes as the number of AOM symptom episodes per 1000 child-years and compared the 2 study periods by using OpenEpi (Open Source Epidemiologic Statistics for Public Health, version 3.01 [14]). For the 2013 study period, we based the incidence on the number of paper diaries returned.

The supplementary analysis included an assessment of characteristics of the ascertained AOM symptom episodes in both periods. We compared AOM-SOSs, as well as other characteristics, including health care use, medication, and parental work absenteeism derived from the disease questionnaire.

Comparisons were made using chi-square and independent-sample *t* tests, where appropriate. Calculations were conducted using IBM SPSS version 22.0 (IBM Corporation). A *P* value of <.05 was considered statistically significant.

Results

Study Population

Of the 300 invited WHISTLER participants, 155 (51.7%) participated during the first study period (2013) and returned at least one paper diary sheet. For the second study period (2015), 69 (17.1%) of 404 invited parents participated and completed at least one monthly app questionnaire. Of these, 36 (52%) were former WHISTLER participants (Figure 3) and of whom 11 (16%) had also participated in the 2013 study period. The baseline participant characteristics were comparable during both study periods (Table 1).

Figure 3. Flowcharts of the study population. WHISTLER: Wheezing and Illnesses Study Leidsche Rijn.

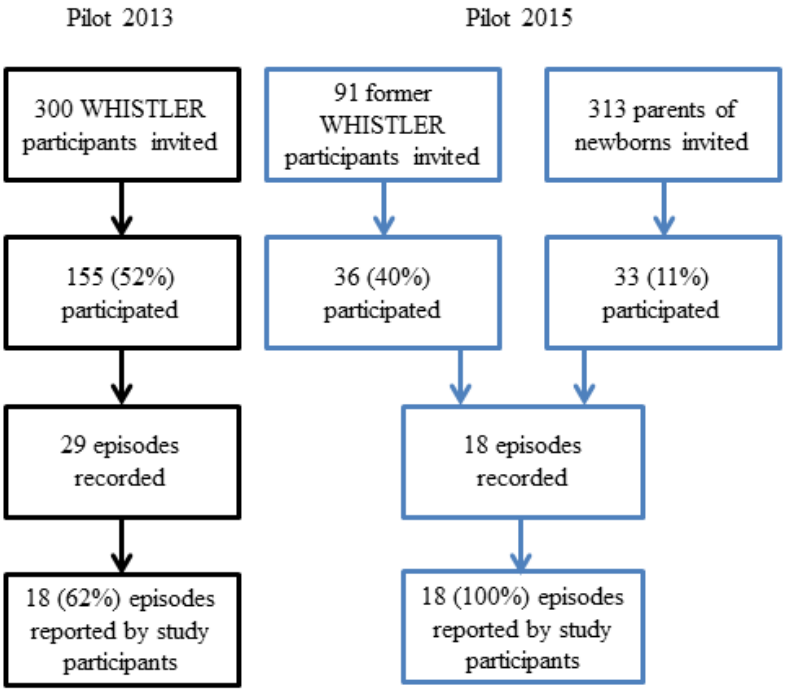


Table 1. Baseline characteristics of the 2013 and 2015 study populations.

Characteristics	Study		*P* value
	2013 paper diary (n=155)	2015 diary app (n=69)	
Infant characteristics			
Age (years), mean (SD)	1.50 (0.72)	1.76 (0.78)	.02
Male sex, n (%)	69 (45)	32 (46)	.88
Participant with at least 1 sibling, n (%)	93 (60)	29 (54)	.43
Daycare visit during study months[a], n (%)	138 (89)	55 (80)	.09
Former WHISTLER[b] study participant, n (%)	155 (100)	36 (52)	<.001
≥1 AOM[c] symptom episode, n (%)	24 (16)	13 (19)	.56
Parent characteristics			
Age (years), mean (SD)	36.2 (3.95)	35.7 (3.75)	.38
High level of education[d], n (%)	131 (89)	60 (92)	.47

[a]Minimum of 1 month of daycare during study period.

[b]WHISTLER: Wheezing and Illnesses Study Leidsche Rijn.

[c]AOM: acute otitis media.

[d]Defined as 1 or both parents having a high vocational or university degree.

Questionnaire Completeness

In 29 diary sheets of 24 different children, the criteria for an AOM symptom episode (fever in combination with otalgia or otorrhea) were met in the 2013 study period, and 18 (62%) of these episodes were actively reported to the study team by the parents. For 17 (59%) episodes, an AOM-SOS scale was completed for 7 days by the parents, and 15 (52%) disease questionnaires were completed. In the 2015 study period, the app automatically detected 18 AOM symptom episodes in 13 different children. For all (100%) of these episodes, the parents completed the questionnaire about health care use and family impact, and 7 days of the AOM-SOS scale. Ascertainment of AOM symptom episodes (*P*=.003) and completeness of AOM reporting (*P*=.003), including AOM-SOS and disease questionnaire (*P*=.001), were significantly higher in the 2015 study period (Table 2).

The 2013 study period contained a total of 18,516 observation days. Data were received for 17,244 days (93.13%). In 2015, data were recorded by the app for 7438 of the 7866 observation days (94.56%, *P*<.001 for difference in number of recorded observation days between the 2 study periods; Table 2). Of the 17,244 reported days in 2013, symptoms were recorded for 5605 (32.50%) as compared with 3272 symptom days out of 7438 days (43.99%) in the 2015 study period (*P*<.001).

During the 2013 study period, of the 617 monthly questionnaires that could have been completed, 575 (93.2%) were returned to the study team. In the 2015 study period, 299 of 329 (90.9%) monthly questionnaires were completed (*P*=.20).

During the 2015 study period, for the former WHISTLER participants, data were retrieved for 4090 of the 4271 observation days (95.8%), while for the non-WHISTLER participants, data were retrieved for 3348 of the 3595 observation days (93.1%) (*P*<.001). The percentage of completed monthly questionnaires was not significantly different between the former WHISTLER participants and the newly recruited parents of newborns in the 2015 study period (*P*=.13).

The incidence of AOM symptom episodes was 605 per 1000 child-years in 2013 and 835 per 1000 child-years in 2015 (Table 2). Table 3 shows characteristics and AOM-SOSs of AOM symptom episodes in both study periods. There were no significant differences in characteristics and AOM-SOSs of ascertained AOM symptom episodes between study periods.

Table 2. Acute otitis media (AOM) incidence and participant compliance with study procedures.

Questionnaire results	Study period		*P* value
	2013	2015	
AOM incidence			
AOM symptom episodes, n	29	18	.003
AOM incidence/1000 child-years	605	835	.29
AOM-SOS[a] questionnaires completed, n (%)	17 (59)	18 (100)	.003
Disease questionnaire completed, n (%)	15 (52)	18 (100)	.001
Participant compliance with recording[b]			
Total days for which data received, n/N (%)	17,244/18,516 (93)	7438/7866 (95)	<.001
Total days with ≥1 symptom reported in diary, n (%)	5605 (33)	3272 (44)	<.001
Monthly questionnaires completed, n/N (%)	575/617 (93)	299/329 (91)	.20

[a]AOM-SOS: acute otitis media severity score.

[b]The degree of compliance was compared between all participants of the 2013 and 2015 study periods (n=155 vs n=69).

Table 3. Characteristics of parent-reported acute otitis media symptom episodes.

Characteristics	Study period	
	2013	2015
Episodes with otalgia, n/N (%)	26/29 (90)	17/18 (94)
Episodes with otorrhea, n/N (%)	11/29 (38)	6/18 (33)
Number of days with fever, median (range)	3.0 (1-11)	2.0 (1-5)
Episodes for which parents stayed home, n/N (%)	6/15 (40)	9/18 (50)
Episodes when parents worried regularly to a lot, n/N (%)	7/15 (47)	9/18 (50)
Episodes for which antibiotics were prescribed, n/N (%)	5/15 (33)	6/18 (33)
General practitioner visits, n/N (%)	9/15 (60)	8/18 (44)
Highest AOM-SOS[a], mean (SD)	8.6 (3.0)	9.9 (3.4)

[a]AOM-SOS: acute otitis media severity score. Highest possible AOM-SOS is 14 for each day. This score consists of 7 discrete items: tugging of ears, crying, irritability, difficulty in sleeping, diminished activity, diminished appetite, and fever. Parents were asked to rate these symptoms daily during 7 days following symptom onset in comparison with the child's usual state, as "none," "a little," or "a lot," with corresponding scores of 0, 1, and 2. Higher scores indicated more severe symptoms. For this study the AOM-SOS scale was translated into Dutch [10].

Discussion

This study evaluated the use of a symptom diary app to detect the occurrence of parent-reported AOM in comparison with conventional (paper) survey methods accompanied by written instructions. The results of this study showed that improved case finding and completeness of disease burden information can be achieved by using an interactive app. Moreover, participants stayed well engaged with app procedures, resulting in 95% completeness of diary data.

Epidemiological research on common infectious diseases often struggles with underdetection of disease events [1-3]. This poses a threat to validity of incidence estimates and quantification of disease burden. For measurement of population incidence and disease burden of numerous infectious diseases, the use of interactive diary apps such as the prototype evaluated in this study could substantially improve accurateness of the estimates. Similarly, diary apps could be valuable in vaccine studies that typically require comprehensive detection and ascertainment of both disease events and potential vaccine side effects [15]. For such purposes, our prototype research app could be easily expanded to include tasks such as specimen collection. The study of Quee et al expanded the app with specimen collection and yielded similar high compliance rates for the questionnaires to those in this study, with 97% diary completeness and 93% to 98% completeness for the 3 different disease questionnaires used in Quee's study (FA Quee, email communication, November 2016) [16].

The field of mobile apps health research (mHealth) is rapidly expanding [7]. Studies reporting on mHealth apps, applied as interventions to improve patient self-management or enhance patient monitoring and disease management, are numerous [17-29]. However, very little has been published on the value of apps as methodological tools in health-related human participant research. To our knowledge, our study is the first to evaluate the use of an interactive symptom diary app as a research tool for detection of infectious disease events that are difficult to capture by health care-based research. We could

identify only 1 previous methodological study that compared the use of a diary app versus a paper diary for the reporting of pain-related symptoms. The use of the pain diary app increased completeness of records by approximately 34% [8]. While we evaluated the value of this diary app in epidemiological research, its application in clinical care to closely monitor a patient's health and aid disease management could be a valuable expansion of the app.

The increasing number of smartphone users, the fact that most smartphone owners have their smartphone on or near them most of the day, and the computer features of smartphones make them attractive tools for health care and research [6,30,31]. The use of apps compared with conventional survey methods, such as paper- or Web-based questionnaires, improves usability for the end user because of the opportunity to include dynamic content and decreased administrative handling [8,31-33]. Moreover, mobile technology offers instant data transfer, allowing for real-time monitoring of and interaction with study participants [15,31]. Several studies have confirmed that study participants prefer the use of apps over paper versions or Web-based questionnaires [8,31,33-38]. While we did not systematically assess parental preferences in our study, the 2015 parents who were former WHISTLER participants, and had thus used paper diary sheets in the past, mentioned a clear preference for the symptom diary app. Combined, these advantages create a potential for increased response and study compliance, improved case detection, and reduced costs when apps are used.

One possible threat when using apps for epidemiological research is the possibility of introducing selection bias because of the requirement of smartphone ownership. In the Netherlands the use of smartphones is widespread and steadily increasing, especially in the age group of young parents, where it now reaches almost 90% [39]. Therefore, introducing apps as research tool does not create selection bias in this particular setting [39,40]. However, when designing a study, this should always be evaluated upfront, especially in countries with lower smartphone penetration or for elderly populations. Concerns have also been raised about the quality of the response when using apps instead of conventional paper questionnaires. However, a recent systematic review by Marcano Belisario et al confirmed that there were no quality differences between data collected via tablet or smartphone apps and paper questionnaires [31]. Another possible limitation of our study is the difference in participants included in the study periods. In 2013, only WHISTLER participants were included, while 2015 included both WHISTLER participants and newly recruited participants. Both the 2013 and the 2015 WHISTLER participants consented to the AOM study as an add-on to the main WHISTLER study and were probably a relatively committed subset. This may explain the somewhat different percentages of completed questionnaires for WHISTLER and non-WHISTLER participants in the 2015 period, where the latter had a lower response rate. Inclusion of newly recruited parents in the 2015 study period might therefore have resulted in an underestimation of the compliance difference and thus the added value of app-based monitoring. Also, 11 parents participated in both study periods. The fact that these parents consented to participate in both study periods, in addition to the WHISTLER study, probably reflects their commitment to scientific research, which was also reflected in the high response rates for both study periods. By analyzing the 2 cohorts as being independent, we may have underestimated the effect of diary app use on the completion rate, since the 11 participants had little room for improvement due to app use.

In conclusion, our results indicate that intensive follow-up of study participants by means of an interactive app has the potential to improve the data quality of infectious diseases occurrence in populations, especially for health events that are difficult to capture by health care-based research. Our findings could have important implications for design and execution of research, both observational and interventional, involving population disease burden quantification, especially when conducted in populations with a high percentage of smartphone users, because digitization will continue and, over time, paper questionnaires may become less accepted.

Acknowledgments

The authors thank all parents and children who participated in the WHISTLER study and in our AOM studies, and all coworkers who took part in the measurements and data management.

The development of the mobile app was supported by an Innovation Grant provided by Pfizer.

Conflicts of Interest

None declared.

References

1. Fortanier AC, Venekamp RP, de Hoog ML, Uiterwaal CSPM, van der Gugten AC, van der Ent CK, et al. Parent-reported symptoms of acute otitis media during the first year of life: what is beneath the surface? PLoS One 2015;10(4):e0121572 [FREE Full text] [doi: 10.1371/journal.pone.0121572] [Medline: 25849847]
2. Sarna M, Ware RS, Sloots TP, Nissen MD, Grimwood K, Lambert SB. The burden of community-managed acute respiratory infections in the first 2-years of life. Pediatr Pulmonol 2016 Dec;51(12):1336-1346. [doi: 10.1002/ppul.23480] [Medline: 27228308]

3. Wolleswinkel-van den Bosch JH, Stolk EA, Francois M, Gasparini R, Brosa M. The health care burden and societal impact of acute otitis media in seven European countries: results of an Internet survey. Vaccine 2010 Nov 19;28 Suppl 6:G39-G52. [doi: 10.1016/j.vaccine.2010.06.014] [Medline: 21075269]
4. Barber C, Ille S, Vergison A, Coates H. Acute otitis media in young children - what do parents say? Int J Pediatr Otorhinolaryngol 2014 Feb;78(2):300-306. [doi: 10.1016/j.ijporl.2013.11.030] [Medline: 24374143]
5. Dubé E, De Wals WP, Gilca V, Boulianne N, Ouakki M, Lavoie F, et al. Burden of acute otitis media on Canadian families. Can Fam Physician 2011 Jan;57(1):60-65 [FREE Full text] [Medline: 21252135]
6. 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]
7. Hussain M, Al-Haiqi A, Zaidan AA, Zaidan BB, Kiah MLM, Anuar NB, et al. The landscape of research on smartphone medical apps: coherent taxonomy, motivations, open challenges and recommendations. Comput Methods Programs Biomed 2015 Dec;122(3):393-408. [doi: 10.1016/j.cmpb.2015.08.015] [Medline: 26412009]
8. Garcia-Palacios A, Herrero R, Belmonte MA, Castilla D, Guixeres J, Molinari G, et al. Ecological momentary assessment for chronic pain in fibromyalgia using a smartphone: a randomized crossover study. Eur J Pain 2014 Jul;18(6):862-872. [Medline: 24921074]
9. Katier N, Uiterwaal CSPM, de Jong BM, Kimpen JLL, Verheij TJ, Grobbee DE, Wheezing Illnesses Study Leidsche Rijn Study Group. The Wheezing Illnesses Study Leidsche Rijn (WHISTLER): rationale and design. Eur J Epidemiol 2004;19(9):895-903. [Medline: 15499901]
10. Shaikh N, Hoberman A, Paradise JL, Rockette HE, Kurs-Lasky M, Colborn DK, et al. Responsiveness and construct validity of a symptom scale for acute otitis media. Pediatr Infect Dis J 2009 Jan;28(1):9-12. [doi: 10.1097/INF.0b013e318185a3a0] [Medline: 19077916]
11. Thuisarts.nl. Mijn kind heeft koorts. Utrecht, Netherlands: Nederlands Huisartsen Genootschap; 2016 Nov 02. URL: https://www.thuisarts.nl/koorts-bij-kinderen/mijn-kind-heeft-koorts [accessed 2017-11-14] [WebCite Cache ID 6uyGbjht4]
12. Thuisarts.nl. Mijn kind heeft een middenoorontsteking. Utrecht, Netherlands: Nederlands Huisartsen Genootschap; 2014 Dec 02. URL: https://www.thuisarts.nl/oorontsteking/mijn-kind-heeft-middenoorontsteking [accessed 2017-11-14] [WebCite Cache ID 6uyGpwKt0]
13. Thuisarts.nl. Mijn kind heeft een loopoor. Utrecht, Netherlands: Nederlands Huisartsen Genootschap; 2015 Jun 09. URL: https://www.thuisarts.nl/loopoor/mijn-kind-heeft-loopoor [accessed 2017-11-14] [WebCite Cache ID 6uyHCLceC]
14. Dean A, Sullivan K, Soe M. OpenEpi: Open Source Epidemiologic Statistics for Public Health, version 3.01. 2013 Apr 06. URL: http://www.openepi.com/Menu/OE_Menu.htm [accessed 2017-11-14] [WebCite Cache ID 6uyJR4m6r]
15. Hughes A, Landers D, Arkenau H, Shah S, Stephens R, Mahal A, et al. Development and evaluation of a new technological way of engaging patients and enhancing understanding of drug tolerability in early clinical development: PROACT. Adv Ther 2016 Jun;33(6):1012-1024 [FREE Full text] [doi: 10.1007/s12325-016-0335-4] [Medline: 27167621]
16. Quee F, De Hoog M, Bruijning-Verhagen P. Intensive monitoring of young families by interactive smartphone application reveals high incidence and transmission of (viral) gastroenteritis. 2016 Nov 28 Presented at: European Scientific Conference on Applied Infectious Disease Epidemiology (ESCAIDE); Nov 28-30, 2016; Stockholm, Sweden.
17. Årsand E, Muzny M, Bradway M, Muzik J, Hartvigsen G. Performance of the first combined smartwatch and smartphone diabetes diary application study. J Diabetes Sci Technol 2015 May;9(3):556-563 [FREE Full text] [doi: 10.1177/1932296814567708] [Medline: 25591859]
18. Beltran A, O'Connor T, Hughes S, Baranowski J, Nicklas TA, Thompson D, et al. Alpha test of a videogame to increase children's vegetable consumption. Games Health J 2012 Jun;1(3):219-222 [FREE Full text] [doi: 10.1089/g4h.2011.0027] [Medline: 24761315]
19. Blackman KCA, Zoellner J, Kadir A, Dockery B, Johnson SB, Almeida FA, et al. Examining the feasibility of smartphone game applications for physical activity promotion in middle school students. Games Health J 2015 Oct;4(5):409-419. [doi: 10.1089/g4h.2014.0120] [Medline: 26287931]
20. Chai YJ, Song J, Kang J, Woo J, Song R, Kwon H, et al. A comparative study of postoperative pain for open thyroidectomy versus bilateral axillo-breast approach robotic thyroidectomy using a self-reporting application for iPad. Ann Surg Treat Res 2016 May;90(5):239-245 [FREE Full text] [doi: 10.4174/astr.2016.90.5.239] [Medline: 27186567]
21. 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;17(8):e210 [FREE Full text] [doi: 10.2196/jmir.4568] [Medline: 26316499]
22. Domhardt M, Tiefengrabner M, Dinic R, Fötschl U, Oostingh GJ, Stütz T, et al. Training of carbohydrate estimation for people with diabetes using mobile augmented reality. J Diabetes Sci Technol 2015 May;9(3):516-524 [FREE Full text] [doi: 10.1177/1932296815578880] [Medline: 25883165]
23. Frøisland DH, Årsand E. Integrating visual dietary documentation in mobile-phone-based self-management application for adolescents with type 1 diabetes. J Diabetes Sci Technol 2015 May;9(3):541-548 [FREE Full text] [doi: 10.1177/1932296815576956] [Medline: 25901020]
24. 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 04;15:531 [FREE Full text] [doi: 10.1186/s12889-015-1845-8] [Medline: 26041131]
25. Huang X, Matricardi PM. Allergy and asthma care in the mobile phone era. Clin Rev Allergy Immunol 2016 May 21. [doi: 10.1007/s12016-016-8542-y] [Medline: 27209270]
26. Moore RC, Fazeli PL, Patterson TL, Depp CA, Moore DJ, Granholm E, et al. UPSA-M: feasibility and initial validity of a mobile application of the UCSD Performance-Based Skills Assessment. Schizophr Res 2015 May;164(1-3):187-192 [FREE Full text] [doi: 10.1016/j.schres.2015.02.014] [Medline: 25790903]
27. Muench F. The promises and pitfalls of digital technology in its application to alcohol treatment. Alcohol Res 2014;36(1):131-142 [FREE Full text] [Medline: 26259008]
28. Stinson JN, Jibb LA, Nguyen C, Nathan PC, Maloney AM, Dupuis LL, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. J Med Internet Res 2013;15(3):e51 [FREE Full text] [doi: 10.2196/jmir.2350] [Medline: 23475457]
29. Zairina E, Abramson MJ, McDonald CF, Li J, Dharmasiri T, Stewart K, et al. Study protocol for a randomised controlled trial evaluating the efficacy of a telehealth program--management of asthma with supportive telehealth of respiratory function in pregnancy (MASTERY©). BMC Pulm Med 2015 Jul 31;15:84 [FREE Full text] [doi: 10.1186/s12890-015-0082-3] [Medline: 26228390]
30. eMarketer. 2 billion consumers worldwide to get smart(phones) by 2016: over half of mobile phone users globally will have smartphones in 2018. 2014 Dec 11. URL: https://www.emarketer.com/Article/2-Billion-Consumers-Worldwide-Smartphones-by-2016/1011694 [accessed 2017-02-14] [WebCite Cache ID 6oGon9lFm]
31. Marcano Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. Cochrane Database Syst Rev 2015 Jul 27(7):MR000042. [doi: 10.1002/14651858.MR000042.pub2] [Medline: 26212714]
32. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
33. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
34. Bush NE, Skopp N, Smolenski D, Crumpton R, Fairall J. Behavioral screening measures delivered with a smartphone app: psychometric properties and user preference. J Nerv Ment Dis 2013 Nov;201(11):991-995. [doi: 10.1097/NMD.0000000000000039] [Medline: 24177488]
35. Kim JH, Kwon S, Shim SR, Sun HY, Ko YM, Chun D, et al. Validation and reliability of a smartphone application for the International Prostate Symptom Score questionnaire: a randomized repeated measures crossover study. J Med Internet Res 2014 Feb 10;16(2):e38 [FREE Full text] [doi: 10.2196/jmir.3042] [Medline: 24513507]
36. Newell SM, Logan HL, Guo Y, Marks JG, Shepperd JA. Evaluating tablet computers as a survey tool in rural communities. J Rural Health 2015;31(1):108-117 [FREE Full text] [doi: 10.1111/jrh.12095] [Medline: 25243953]
37. Salaffi F, Gasparini S, Ciapetti A, Gutierrez M, Grassi W. Usability of an innovative and interactive electronic system for collection of patient-reported data in axial spondyloarthritis: comparison with the traditional paper-administered format. Rheumatology (Oxford) 2013 Nov;52(11):2062-2070 [FREE Full text] [doi: 10.1093/rheumatology/ket276] [Medline: 23955646]
38. Stomberg MW, Platon B, Widén A, Wallner I, Karlsson O. Health information: what can mobile phone assessments add? Perspect Health Inf Manag 2012;9:1-10 [FREE Full text] [Medline: 23209453]
39. Centraal Bureau voor de Statistiek. ICT, kennis en economie 2016 ed. Den Haag, Netherlands: Centraal Bureau voor de Statistiek; 2016. URL: https://www.cbs.nl/-/media/_pdf/2016/26/ike2016_web.pdf [accessed 2017-02-14] [WebCite Cache ID 6oGpovovg]
40. Statista. Global mobile OS market share in sales to end users from 1st quarter 2009 to 1st quarter 2016. 2016. URL: https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/ [accessed 2017-02-14] [WebCite Cache ID 6oGpMzH12]

Abbreviations

AOM: acute otitis media
AOM-SOS: severity score

Edited by G Eysenbach; submitted 17.02.17; peer-reviewed by C Nguyen, Y Guo, L Dang, J Kaewkungwal, P Schulz; comments to author 27.06.17; revised version received 14.08.17; accepted 23.09.17; published 28.11.17.

Please cite as:
Prins-van Ginkel AC, de Hoog MLA, Uiterwaal C, Smit HA, Bruijning-Verhagen PCJ
Detecting Acute Otitis Media Symptom Episodes Using a Mobile App: Cohort Study
JMIR Mhealth Uhealth 2017;5(11):e181
URL: http://mhealth.jmir.org/2017/11/e181/
doi:10.2196/mhealth.7505
PMID:29183869

Original Paper

Electronic Brief Intervention and Text Messaging for Marijuana Use During Pregnancy: Initial Acceptability of Patients and Providers

Justin Gray[1], MD; Jessica R Beatty[1], PhD; Dace S Svikis[2,3,4], PhD; Karoline S Puder[5], MD; Ken Resnicow[6], PhD; Janine Konkel[6], MPH; Shetoya Rice[1], BA; Lucy McGoron[1], PhD; Steven J Ondersma[1], PhD

[1]Merrill Palmer Skillman Institute, Department of Psychiatry and Behavioral Neurosciences, Wayne State University, Detroit, MI, United States
[2]Department of Psychology, Virginia Commonwealth University, Richmond, VA, United States
[3]Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, United States
[4]Department of Obstetrics and Gynecology, Virginia Commonwealth University, Richmond, VA, United States
[5]School of Medicine, Department of Obstetrics and Gynecology, Wayne State Unviersity, Detroit, MI, United States
[6]School of Public Health, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Jessica R Beatty, PhD
Merrill Palmer Skillman Institute
Department of Psychiatry and Behavioral Neurosciences
Wayne State University
71 East Ferry
Detroit, MI, 48202
United States
Phone: 1 3136642518
Fax: 1 3136642555
Email: jbeatty@wayne.edu

Abstract

Background: Marijuana is the most widely used illicit substance during pregnancy. Technology-delivered brief interventions and text messaging have shown promise in general and pregnant samples but have not yet been applied to marijuana use in pregnancy.

Objective: The objective of the study was to evaluate, among pregnant women and prenatal care providers, the acceptability of an electronic brief intervention and text messaging plan for marijuana use in pregnancy.

Methods: Participants included patients (n=10) and medical staff (n=12) from an urban prenatal clinic. Patient-participants were recruited directly during a prenatal care visit. Those who were eligible reviewed the interventions individually and provided quantitative and qualitative feedback regarding software acceptability and helpfulness during a one-on-one interview with research staff. Provider-participants took part in focus groups in which the intervention materials were reviewed and discussed. Qualitative and focus group feedback was transcribed, coded manually, and classified by category and theme.

Results: Patient-participants provided high ratings for satisfaction, with mean ratings for respectfulness, interest, ease of use, and helpfulness ranging between 4.4 and 4.7 on a 5-point Likert scale. Of the 10 participants, 5 reported that they preferred working with the program versus their doctor, and 9 of 10 said the intervention made them more likely to reduce their marijuana use. Provider-participants received the program favorably, stating the information presented was both relevant and important for their patient population.

Conclusions: The findings support the acceptability of electronic brief intervention and text messaging for marijuana use during pregnancy. This, combined with their ease of use and low barrier to initiation, suggests that further evaluation in a randomized trial is appropriate.

(JMIR Mhealth Uhealth 2017;5(11):e172) doi:10.2196/mhealth.7927

KEYWORDS

pregnancy; marijuana; intervention study; text messaging

Introduction

A growing literature suggests that marijuana use during pregnancy is associated with low birth weight, stillbirth, and neurologic or developmental impairment [1-4]. Subtle sequelae are most apparent later in development, with impaired cognition, decreased academic achievement, and increased risk of abusing drugs later in life being the most prevalent [5]. Despite these possible adverse outcomes, marijuana use during pregnancy often goes unaddressed. Physicians and other providers must act with caution when considering drug use, especially marijuana, in a pregnant population. This has become increasingly important because past month marijuana use in the general population (22.2 million Americans, aged 12 or older) has increased from 6.2% in 2002 to 8.3% in 2015, with 19.8% (6.8 million) of those aged between 18 and 25 reporting past month use [6]. Among pregnant women, 6.4% of those in the age group of 18 to 25 years report past month use, and 1.3% of those aged between 26 and 44 years report past month use [7]. An increasing number of patients are turning to this substance for its antiemetic properties; nausea is an approved indication in all states where medical use of this drug has been legalized [8]. This puts an increasing number of pregnancies at risk.

Women with substance use disorders are less likely to enter treatment than men [9]. Furthermore, most persons in need of treatment neither receive it nor feel that they need it [10], suggesting a need for approaches that are brief enough to be acceptable even to those who are unwilling to engage in formal treatment. As a consequence, screening, brief intervention, and referral for treatment (SBIRT) has become a recommended core element of prenatal care [11].

However, integration of SBIRT into the clinical setting carries with it several practical challenges. Evidence from attempts to implement SBIRT for alcohol and tobacco have demonstrated large variations in efficacy, significant SBIRT-related training costs, and a notable increase in overall visit length [12,13]. Furthermore, additional physician and staff training in SBIRT focusing on smoking cessation has been shown to have modest or transient effects on trainee behavior, and few prenatal care providers fully implement recommended brief intervention strategies [14-17]. In one national survey, only 27.49% of women (1323 out of 4812) reported even being asked about alcohol use by their primary care provider [18]. These challenges—time, burden, and training-related—are only magnified when seeking to add additional screening and brief intervention foci such as marijuana on top of existing calls to address alcohol and tobacco. Furthermore, disclosure of substance abuse can be both socially stigmatizing and subject to potential legal consequences (at least in most states; it remains to be seen whether use in pregnancy will be viewed differently in states where it is legalized, as women are most often pursued via child welfare statutes rather than on drug-related charges). Regardless, this stigmatization and social risk suppresses disclosure, which in turn further limits the proportion of at-risk women who receive even a brief intervention.

Electronic administration of SBIRT (eSBIRT) may help to address these challenges. Computer and electronic approaches could potentially reduce financial barriers, increase reproducibility, and facilitate consistent delivery across patient populations [19]. These approaches also offer the ability to tailor parts of the program based on patient responses, conferring a feeling of personal relevance to participants that enhances message impact [20]. In addition, computer-delivered screening is associated with greater disclosure of drug use, as is provision of anonymity [21], which is possible via technology-delivered brief interventions.

Delivering SBIRT via technology also allows incorporation of text messaging, which can readily use data collected during screening and brief intervention to send subsequent tailored communications. Text messages have a number of advantages, including the near ubiquity of mobile phone ownership: 100% of adults in the age group of 18 to 29 years own a cell phone, and, as of 2016, 100% of people aged between 18 and 29 years who own a cell phone use it for texting at least occasionally [22]. Text messaging interventions carry relatively low operational costs that do not increase as the reach of the program increases [23]. Furthermore, a recent systematic review found that text messages were particularly effective as a supplement to Internet-based interventions [24]. This result may reflect some of the relative advantages of text messaging such as (1) the ability of text-based messaging to maintain multiple communications with a participant without reliance on a subsequent meeting; (2) the ability of text messages to reach the participants in their natural environment; and (3) the fact that text messages—unlike other forms of communication, such as mailings—are nearly always opened (99% are opened, 90% within the first 3 min of receipt [25]). For example, with regard to smoking in pregnancy, Naughton et al (2011) found that participants receiving tailored SMS (short message service) text messages were more likely to set a quit date and reported higher levels of self-efficacy, harm beliefs, and determination to quit than their control counterparts [26]. Additionally, Quit4baby (Voxiva Inc), a text message program designed to help women reduce tobacco use during pregnancy, was rated as helpful in getting them to make a quit attempt and providing ideas on ways to quit [27].

Text messaging and eSBIRT may therefore have utility in addressing marijuana use in pregnancy. However, neither of these approaches has previously been tested in this context. This study is a preliminary, qualitative, and quantitative examination of the acceptability of eSBIRT and a tailored texting protocol among 2 key samples: (1) pregnant women reporting regular marijuana use during the month before becoming pregnant and (2) prenatal clinic staff. Qualitative and quantitative feedback described in this study informed changes before the start of a larger clinical trial.

Methods

Participants

Patient-Participants

Patient-participants were 10 women recruited from a prenatal clinic in Detroit, Michigan. Inclusion criteria were self-report of marijuana use at least twice weekly in the month before

pregnancy, aged between 18 and 40 years, less than 20 weeks pregnant, and owning a cellphone (As participants would be responsible for any charges resulting from receiving text messages on their personal phone, all participants were specifically asked their willingness to receive text messages during a feedback interview.). Exclusion criteria included inability to understand English, inability to provide consent, consideration of an elective abortion or adoption for the current pregnancy, or past participation in any other study by the authors.

Provider-Participants

Of the medical staff from the same prenatal clinic, 12 members volunteered to participate in focus groups regarding the intervention materials. We offered participation to all physicians in the department; 5 physicians were available to attend one focus group, and 7 medical staff (all of the nurses, medical assistants, and reception staff from the clinic where recruitment took place) participated in the second focus group.

Procedure

Recruitment for pregnant participants began on March 10, 2015, and ended on July 10, 2015. In total, 61 participants were screened over 41 days of recruitment to obtain the 13 (21%) eligible participants. Of the 13 eligible patients, 10 (77%) participated in the feedback session. Participants were recruited until saturation was reached (high quantitative ratings of the intervention's ease of use, usefulness, and attitude toward using the intervention, and no new information obtained from interviews) [28]. Focus group data were collected during March 2015.

Patient-participants were given a flyer by clinical staff describing the study. Medical staff explained that it was a voluntary research project involving 5 to 10 min to determine eligibility for a larger study and that they would receive a small gift for their baby if they choose to participate. Those showing interest were introduced to the research assistant, who prescreened for age, gestation, ability to understand spoken English, and ability to receive text messages. Women passing this prescreen process reviewed an informed consent information sheet on a study provided on a tablet personal computer (PC) with headphones (with disposable sanitary covers). Those who consented completed eligibility screening.

Interested and eligible patient-participants reviewed a second research information sheet approved by the institutional review board (IRB) describing the anonymous, single-session study. Only patient-participants who were eligible were told that the study was focused on reducing marijuana use during pregnancy. Those who provided consent to participate spent approximately 1 hour reviewing the intervention materials on the tablet PC and completing a semistructured feedback interview (approximately 30 min). All patient-participant activities were conducted via individual sessions with research staff in a private office within the clinic.

All patient-participants responded to qualitative acceptability items using the same tablet PC used for consent and intervention delivery. The research assistant conducted the semistructured interview after patient-participants completed the intervention and acceptability items. The brief interview was designed to elicit the participant's overall impression of the software, evaluations of its helpfulness, likes and dislikes, and suggested changes. Unclear or short responses were probed to elicit more information. Once questions about the intervention and videos (included as part of the intervention) were answered, participants were provided with a list of text messages to review. It was explained that these messages would be sent on a schedule selected by the patient (ie, once, twice, or 3 times a week) over the course of their pregnancy. Patient-participants were told that those participating in the later clinical trial would receive text messages that were tailored based on the responses provided during the initial screening and software intervention. For example, when asked whether they had ever been prescribed medication for depression, anxiety, or any other emotional difficulties, patient-participants would receive one of two messages. If they answered in the affirmative, the following message would be sent:

> *Smoking may seem to help when you are feeling down, but after, it may make you feel worse. Pay attention to your body and emotions. Talk to your doc for help facing your struggles.*

If patients denied ever seeking outside services, they would receive the following message:

> *Every day you don't smoke, be extra good to yourself. You deserve it. Tell yourself, "What I am doing is amazing."*

The research assistant pointed out this single example of how text messages could be tailored before allowing the participant to review the list. After the participants reviewed the text message list, they completed the software acceptability items and were subsequently interviewed.

The research assistant received training in conducting semistructured interviews from one of the coauthors (JRB) and was supervised for several practice interviews to determine competency. All responses were transcribed verbatim during the interview for later analysis. A total of 39 questions were asked of all participants. After completion of the interview, all participants were provided with a referral guide, encompassing multiple areas of risk (eg, substance abuse, emotional health, and education and job training), for relevant services in the area. Procedures were in place for active referral to clinic staff and local treatment centers, but none of the participants expressed interest in this level of service at the time of the interview. Participants received a US $30 Target gift card for participation in the study.

The provider-participant focus groups were conducted on two different dates: one with physicians and one with other medical staff. Procedures were the same for both focus groups. An informed consent information sheet was reviewed before each focus group began. Medical staff were then shown the intervention on a large projector screen. Initial reactions were obtained halfway through the intervention and again after the end of the intervention. After all interview questions were answered, staff were handed a copy of the potential text messages and asked for reactions related to that information. Before concluding the focus group, all staff were asked about

the best ways to implement the study without disrupting the flow of ongoing care. Each focus group was audio-recorded and then transcribed. All procedures were reviewed and approved by the university IRB before any recruitment.

Measures

Software Acceptability (Patient-Participants)

Acceptability of the computer-delivered intervention was assessed using participant responses to 12 self-report questions (6 additional questions focused on the videos shown as part of the brief intervention and 10 on text messages). These 28 items, relating to ease of use, respectfulness, helpfulness, and likability, were rated using a 5-point Likert scale (with 1=not at all and 5=very much) after completion of the intervention. The question, "How much did some parts of the computer bother you?," was reverse-coded, with 1=very much and 5=not at all. These acceptability questions were similar to those developed and used successfully in previous usability testing [29] and are based on the Technology Acceptance Model [30]. Additionally, 1 yes-no question asked participants about change likelihood as a result of the intervention:

> *Are you more likely to be successful with this goal [of quitting marijuana] because of your participation here?*

Feedback Interview (Patient-Participants)

The 39-item open-ended interview used in this study was based on expert opinion (the majority of the coauthors were included in this group) and consensus on what information would be helpful in making modifications to the intervention. See Textbox 1.

Focus Group Interview (Provider-Participants)

Focus group questions focused on how the providers felt their patients would react to the intervention materials, how helpful they felt the intervention materials were, and how the study procedures could be best integrated into the ongoing clinic procedures. The questions were similar in language and content to those asked of the patient-participants.

Interventions

The brief intervention developed for this study uses patient responses to provide an individualized, interactive experience. A three-dimensional animated narrator guides the participant through the intervention. The narrator is able to speak, move, provide empathic reflections, and display appropriate emotional responses. The program includes aural as well as visual presentation of all content, and all answers are recorded by simply tapping responses from a list or by touching a visual analogue scale. The narrator reads aloud any written material on the screen, including response options (participants just have to click on the word or phrase to hear it read aloud).

The intervention content was adapted from brief intervention [31-33] and motivational interviewing techniques [34]. It was adapted and modified to specifically address marijuana use from a previous brief intervention designed to reduce alcohol use during pregnancy [35]. The intervention begins with a brief introduction, followed by an embedded video of a physician discussing potential benefits of reducing or quitting marijuana use during pregnancy and of a mother describing her own decision to quit using marijuana while pregnant (all were actors). Next, participants are asked to report how they feel about their use of marijuana while pregnant. Those who reported already stopping all marijuana use receive normed feedback designed to reinforce the decision to stop using, were asked to provide the reasons and advantages for their decision to quit, and were helped to develop a personalized plan for preventing relapse. This branch of the intervention was designed to be potentially efficacious with both women who have quit and women who have not quit but choose to indicate abstinence to avoid negative reactions. Participants who endorsed active use received content that was consistent with traditional brief intervention approaches and that included normed feedback (normed for age, pregnancy status, and gender), decisional balance exercises, and an optional change plan with a menu of change options. See Figure 1 for a visual diagram of the intervention flow.

The video embedded in the intervention featured a physician providing gain-framed information about the benefits of reducing marijuana use during pregnancy and a mother providing a testimonial regarding her decision to avoid marijuana while pregnant. Multiple versions of each video were available and were tailored to participants based on self-reported self-efficacy, race, and motivation.

The text messages were designed to be tailored on a number of factors, including self-efficacy, gestational stage, social support, and processes of change based on the theory of planned behavior [36] and self-determination theory [37]. Approximately two-thirds of messages are related to the participant's specific goal regarding marijuana: the remaining one-third of the content was designed to provide either community resources relevant to the pregnancy or inspirational quotes. All intervention materials were developed with ongoing expert feedback and final review before being presented to participants.

Textbox 1. Interview questions.

Intervention

1. What did you think about the program?
2. What parts did you like the most?
3. What parts did you like the least?
4. What, if anything, do you think should be changed?
5. What do you think about the introductions?
6. How was [the narrator's] voice?
7. Did any of the questions of parts of the program bother you?
8. What was useful to you in the program?
9. How has using it changed your thoughts on your marijuana use, if at all?
10. Did you make a personal plan for how to change your marijuana use?

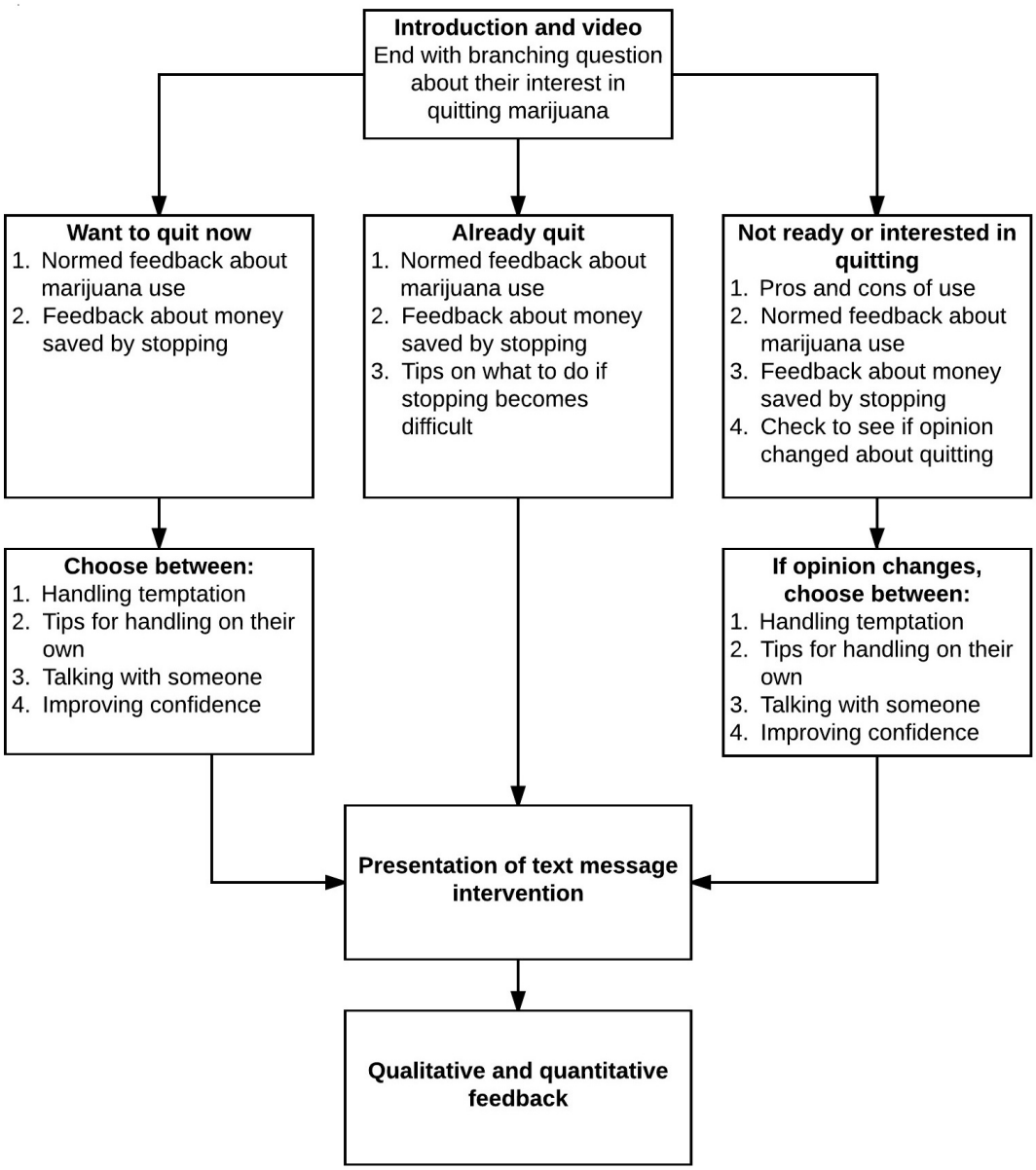
Video

1. What did you think about the videos?
2. What bothered you about the videos?
3. What did you like about the videos?
4. What could we do to improve the videos?
5. Did the videos give you the impression that we didn't understand marijuana or that we were acting like it's more dangerous than it really is?
6. Did the video feel preachy or judgmental?
7. What did you learn from the videos?
8. Was the woman someone you could relate to? Why or why not?
9. Did the videos change your opinion in any way?

Short message service

1. What did you think about the text messages?
2. Did any of the content on the messages bother you?
3. What words/language made you uncomfortable?
4. What was useful about the messages?
5. How much did it feel like the messages were intended for you? What could have made them feel more personalized?
6. How did you feel about the advice you were given?
7. How did you feel about the amount of information? Too much or too little?
8. What would you have liked to see added to text messages? Was there anything that didn't need to be there?
9. How have your feelings about your marijuana changed since you read the messages? What did it make you think about or want to do differently, if anything?
10. Would you be comfortable receiving these messages on your phone?
11. If you got these messages and someone saw them on your phone, how would you feel about that?
12. Do you have a phone with texting abilities? If yes, how often do you use texts?
13. If you were to receive text messages such as these, how often would you like to receive them?

Figure 1. Intervention flow.



Data Analysis

Means, standard deviations (SDs), and counts for demographic data and software acceptability ratings were analyzed via SPSS version 22 (IBM Corp). Qualitative data were subjected to manual, thematic coding of participants' oral responses of transcribed interviews (patient-participants) and focus groups (provider-participants). Responses were grouped by similar patterns or themes that occurred throughout the interviews [38]. Next, coders decided what construct label best described each of the groupings. The coding and classification of the responses were performed and validated independently by 2 members of the investigative team (JG and JRB). Both coders met to compare the independently generated groupings and themes. All disagreements were discussed until a consensus was reached for a total of 10 themes.

Results

Quantitative Feedback

Using a 5-point rating scale, average ratings for the intervention were all at or above 4.4, with the highest rating being for excitement about making a change in marijuana use (mean 4.9, SD 0.38; see Table 1). All 10 patient-participants reported feeling more likely to change their marijuana use because of their interaction with the computerized intervention. All ratings for the video were at or above 4.0 except for the item "The mom in the video looks like me" (mean 3.7, SD 1.49). Average ratings of the following text message items were below 4.0: likability (mean 3.7, SD 1.0), how interesting (mean 3.7, SD 1.23), helpfulness (mean 3.6, SD 1.06), got you thinking about your use (mean 3.3, SD 1.41), and interest in receiving text messages (mean 3.2, SD 1.09).

Table 1. Quantitative ratings of software, video, and short message service (SMS) acceptability (scored on a 5-point Likert scale, with 1=not at all and 5=very much; positive rating=score of 4 or 5; n=10 for intervention and video ratings, n=9 for text message ratings).

Question	Giving a positive rating, n (%)		
	Intervention	Video	Text messages
Likability	7 (70)	5 (50)	5 (56)
Ease of use	10 (100)	-	8 (89)
How interesting	8 (80)	-	5 (56)
Respectful	9 (90)	-	9 (100)
Bothered by parts of program[a]	8 (80)	-	7 (78)
Helpfulness	9 (90)	-	4 (44)
Feel better about yourself	8 (80)	-	7 (78)
Got you thinking about your use	9 (90)	-	5 (56)
Program seemed to understand you	9 (90)	-	-
Interest in working with program again/receiving messages	7 (70)	-	3 (33)
Excited about changing marijuana use during pregnancy	10 (100)	-	-
Think other moms would be helped	9 (90)	-	7 (78)
Learned new information	-	7 (70)	-
Learned useful information	-	7 (70)	-
More motivated to avoid marijuana	-	9 (90)	-
Already knew information presented	-	6 (60)	-
Mom in video was like me	-	5 (50)	-

[a]This item was reverse-scored.

Qualitative Feedback

Analysis of open-ended responses yielded 7 major themes among the 3 sections of the software arising from the patient-participant interviews, as well as 3 major themes arising from the medical staff focus feedback group; see Textbox 2.

Intervention

Liked Working With the Program

Of the patient-participants, 8 reported feeling some personal relevance of the information whereas the remaining 2 were ambiguous. When asked whether they would rather talk to the program or a person about their marijuana use, 5 patient-participants indicated that they preferred interacting with the program, whereas 4 said they would have preferred talking to their doctor; one participant reported that either option was fine. For example one patient-participant said:

> *It's hard to be truthful with people you don't know. It wouldn't feel so awkward as talking with real people you don't know, it's hard. It's easier to tell the whole truth with the computer.*

Supported or Changed Thoughts About Marijuana

After completing the computerized intervention portion of the study, 6 patient-participants reported making a plan to quit; 3 reported already ceasing use, and one felt she could stop on her own; 7 patient-participants reported being more likely to stop marijuana use after interacting with the program. When asked about the most useful part of the program, 4 women highlighted the tips on how to cut down or quit and 2 mentioned the information about how much money they could save by quitting. All but 2 patient-participants indicated that interacting with the program changed something about how they saw their use of marijuana during pregnancy. For example, another patient-participant said:

> *It was great because [the narrator] gave true facts about marijuana use. [The narrator] gave better information. It's better than listening to people on the street.*

Textbox 2. Themes and examples.

Intervention

- Liked working with the program
  - Real great, very educational too.
  - It was well balanced, covered both the good things and bad things and helped me not feel bad about my use.
  - [The program] makes it less awkward. I am okay talking with people I know. [The program] would be more helpful, because the whole truth will come out.
- Supported or changed thoughts about marijuana
  - It boosted my confidence up, like you can do it girl.
  - Standing firm against temptation: I like when [the narrator] said take it out of the house and stay away from people that do it.
  - I actually learned that marijuana may not affect your pregnancy at the moment with an infant, like it could be later on in life.
- Women in the videos were helpful
  - Relatability of actors/actresses used was important to most, if not all, of the participants.
  - The mother's testimony, I love that part. She told how she got through it by thinking about the baby.
  - More testimonials may hit people in a different way, it can relate to more people.
- More information about harmful effects
  - Dr said marijuana use can cause learning disability, I didn't know that.
  - The MD's video made me want to be as healthy as possible to have a healthy baby.
  - More in depth for ladies on the second or third pregnancies. Even though I'm on my fifth pregnancy, I'm pretty sure there is something I need to learn still.

Short message service feedback

- Text messages were helpful.
  - Really good topics, would pass it on to friends.
  - Great idea, a lot of people use cell phones and text.
  - I like them because it's a new idea and a way to motivate because they can be experiencing a moment of temptation.
  - Helpful...to deal with bad moods and things to do so you won't relapse.
- Learned about community resources
  - I like the dial 2-1-1 for help.
  - Text messages can help people quit and provide resources that we didn't know about.
- Information presented was overwhelming
  - Excessive, a lot of information.
  - Too much—make them briefer.
  - A little too much, people will get bored reading all that.

Medical staff focus group

- Concern about patient takeaway
  - Dose does matter. If someone does cut back, that's a victory too. I think this is good message if low motivation. A stronger message would be needed for high motivation women.
  - Sounds kind of soft in terms of message about marijuana—hedges too much.
  - Last line comes across as if you use less, than that's okay too. I think I would like an "abstinence" message, but not sure. We'll take what we can get, but thinks it leave the door open where it's okay to use a little.

- Integration into patient visit
  - The time after they are "roomed" is wasted. It would give them the opportunity to do something.
  - All the time moms sit [at the office] for glucose testing—would be a good opportunity to catch them or give a second dose of the intervention.
- Quality of the presentation
  - I liked it. It would connect well. A lot of my patients are surrounded by marijuana. I like how the woman [in the video] talks about how she made up her own mind—drug use is kind of rebellious and this allows her to keep this.
  - [Specific text message] should be at the beginning. That's when everything is nasty. They feel they need weed to eat. A lot of patients feel like that. They have to be told to try other foods.
  - A couple of the [screens] are too busy. Follow the rule of 6—no more than 6 words per line and 6 lines of text.
  - [The doctor in the video] really looks like an OB/GYN instead of an actress; authentic.

Video

Women in the Videos Were Helpful

Relatability of the testimonials was brought up by several patient-participants. Although many were critical of one of the actresses (saying that she wasn't believable and "seemed directed as if reading off a script"), most reported liking another actress who was African American and who spoke in a casual, unscripted manner. Overall, patient-participants found the testimonials to be helpful, and some found it to be the most helpful element. For example, one patient-participant said:

> *The mother's testimony, I love that part. She told how she got through it by thinking about the baby. I tell my friends to think about the baby too. The video boosted my confidence.*

More Information About Harmful Effects

In addition to asking for more testimonials, patient-participants also asked to see more information about the effects of marijuana on the mother and fetus. Several women stated they would like to hear more of this information from the physician in the video in addition to the patient testimonials. For example, one patient-participant said:

> *The doctor's video made me want to be as healthy as possible for my baby.*

Information about the real effects of marijuana and about the controversy in this area was also well received. When asked what she liked about the videos, a different patient-participant said:

> *The picture of the pregnant woman smoking and people on either side saying good and bad things. It didn't make me feel bad and it's good to show that.*

Text/SMS

Text Messages Were Helpful

Overall, patient-participants found the text messages easy to understand, helpful, and encouraging. One patient-participant said that the messages provided information she would not be able to get from her friends. When asked what they thought about the text messages, one patient-participant said:

> *A lot of them were helpful actually. The eating, changing the diet, getting a car seat immediately, the exercising and taking showers when you feel like you want to smoke.*

The most common suggestion was to include more resources, either specific resources from the community they felt were missing or more tips on how to quit. When asked about frequency, 3 patient-participants said they would like to receive texts once a week, 3 said twice a week, and 4 said 3 or more times a week.

Learned About Community Resources

Many patient-participants learned about community resources they were not aware of and found the additional tips on ways to avoid temptation or keep from using helpful. When asked what she thought about the messages, one patient-participant stated:

> *They were helpful. There were not only tips but there were resources and phone numbers.*

Another response suggested adding even more information about resources.

> *Add more about the resources. 2-1-1 and WIC are good, maybe [add] FIA (Family Independence Agency/Department of Human Services) info.*

Information Presented Was Overwhelming

In total, 4 participants indicated that there was too much information presented in the text messages. However, this concern appeared to be related to the way it was presented (1 message for each week of pregnancy, all in one document) and not the amount within each individual text message.

Provider-Participant Focus Groups

Concern About Patient Takeaway

Both doctors and staff reported liking the intervention overall. Many commented on the strength of the message. Some felt there needed to be more emphasis on the dangers of marijuana use and the need to completely quit. Overall, both focus groups understood and appreciated the positive message behind "the less you use the better" for both the woman and the baby. Provider-participants felt the doctor was very relatable and felt like a real doctor rather than an actress.

Integration Into Patient Visit

Doctors and staff agreed that overall the message was important for their patients to hear and that the information provided and actresses chosen would be relevant for their patients. Provider-participants reported that there may be several junctures at which to incorporate the program into the flow of the patient visit. For example, they indicated that the point after the patient is brought to the exam room, but before the patient is seen by the physician, is generally wasted. They also felt that the 26-28 week glucose screen all pregnant patients must complete would be another opportunity to present the intervention (or a follow-up intervention) because the patient is required to wait for 1 hour before the blood is drawn.

Quality of the Presentation

Although provider-participants generally liked the actresses, there was one actress in particular that staff felt was not natural or relatable (the same actress criticized by the patient-participants). Half of the medical staff did not like the voice for the animated character. Other specific suggestions included decreasing the amount of information presented on the screen at one time and cutting back on the length of the introductions and transitions during the intervention. Medical staff suggested changing the order of the text messages to fit better with the stage of pregnancy.

Discussion

Principal Findings

This study sought to evaluate, among pregnant women and prenatal care providers, the acceptability of an electronic brief intervention and text messaging program to reduce marijuana use in pregnancy. Following the technology acceptance model [30], we evaluated participants' acceptance of the interventions, video and text messaging, and their perceived usability and ease of use. Overall, patient-participants rated the brief intervention positively, more specifically as helpful, respectful, interesting, and easy to use. The videos that were embedded within the brief intervention were also seen as unbiased and as presenting helpful information about cessation and the effects of marijuana on the baby. The information was well received, although one actress for the testimonials was consistently given negative feedback. Similarly, the text messaging content was seen as helpful in providing information about community resources and additional tips on how to reduce or cease using marijuana during pregnancy.

Provider-participant feedback was similarly positive regarding the brief intervention content. They saw the information as both relevant and important for their patients to hear. However, questions were raised regarding the strength of the message and whether it should focus on complete cessation or harm reduction. They offered several suggestions regarding integration into existing structure of patient care, including utilizing the time between intake procedure and examination and follow-up or initiation of intervention during the mandatory glucose screening visits.

These findings suggest that the women in this study were open to examining their marijuana use during pregnancy and to doing so via technology. Participants were happy with the unbiased presentation of the effects of marijuana on the baby, found the materials useful and easy to use, and clearly spent time evaluating whether or not they should stop use during pregnancy.

How Feedback Informed the Planned Clinical Trial

On the basis of the feedback from both patient-participants and provider-participants, several changes were made to the intervention materials in preparation for the planned clinical trial. The order of the information presented in the text messages was improved based on the stage of pregnancy (ie, suggestions for dealing with nausea were moved to the beginning when women are more likely to be struggling with morning sickness). Additionally, the amount of text on screen and introduction language used during the intervention was reduced to only the key pieces of information. Not all feedback provided by provider-participants and patient-participants could be incorporated into changes to the intervention. Changes to the voice of the narrator were not possible. However, to facilitate more consistently positive feelings toward the narrator, research staff carefully reviewed the narrator actions changing narrator movements and using less smiling.

On the basis of the feedback from provider-participants, future clinical trial participants will be recruited, screened, and will participate (using a tablet provided by study staff) during natural periods of downtime during their prenatal visit (eg, while the patient is waiting to see the provider, in the exam room, or immediately after the visit is complete). Furthermore, based on the feedback from patient-participants regarding the frequency of receiving text messages, those enrolled in the clinical trial will be offered the choice of receiving text messages on their personal phone once, twice, or thrice per week for the remainder of their pregnancy. Despite some patient-participants reporting little use of texting, all reported the ability to receive text messages. Finally, recruitment rates obtained during feedback suggested that alternate recruitment options (such as partnering with midwifes and ultrasound technicians) should be considered.

Limitations

This study is limited by its relatively small sample size of all African American women from a clinic in the urban Detroit area. However, our aim was not to conduct a fully powered test of an a priori hypothesis, but rather to provide information regarding participant acceptability and usability, which typically involves smaller sample sizes. Additionally, it may have been preferable to present text messages for feedback as a presentation where each text message could be looked at separately. Having a single document with a sample of each week's messages was overwhelming for some participants. Future studies should also consider ways to tailor the text messages for the participants providing feedback. This study was only able to show examples and describe how messages would be tailored and may have missed valuable feedback because of the presentation format.

Conclusions

Technology-based approaches have the potential to access a relatively high proportion of any given at-risk population. If

this potential is realized, it could result in a substantial public health impact even when effects are modest. Furthermore, promotion of help-seeking is also an integral part of the proposed high-reach interventions, making them an ideal complement to more intensive support programs, where available. The acceptability (and even preference in some cases) of these technology-based approaches for marijuana use in pregnancy suggests that further research, particularly evaluation of efficacy, is needed.

Acknowledgments

This research was supported by NIH award DA036788 to SJO. The authors wish to gratefully acknowledge the medical staff from the WSU University Physicians Group Obstetrics & Gynecology Clinic and the invaluable assistance of the pregnant women who participated in the study.

Conflicts of Interest

SJO is part owner of Interva, Inc, the company that markets the intervention authoring tool used to develop the computerized intervention used in the study. Interva, Inc had no role in the study design, data analysis, or write-up of results.

References

1. Gray TR, Eiden RD, Leonard KE, Connors GJ, Shisler S, Huestis MA. Identifying prenatal cannabis exposure and effects of concurrent tobacco exposure on neonatal growth. Clin Chem 2010 Sep;56(9):1442-1450 [FREE Full text] [doi: 10.1373/clinchem.2010.147876] [Medline: 20628142]
2. El MH, Tiemeier H, Steegers EA, Jaddoe VW, Hofman A, Verhulst FC, et al. Intrauterine cannabis exposure affects fetal growth trajectories: the Generation R Study. J Am Acad Child Adolesc Psychiatry 2009 Dec;48(12):1173-1181. [doi: 10.1097/CHI.0b013e3181bfa8ee] [Medline: 19858757]
3. Zuckerman B, Frank DA, Hingson R, Amaro H, Levenson SM, Kayne H, et al. Effects of maternal marijuana and cocaine use on fetal growth. N Engl J Med 1989 Mar 23;320(12):762-768. [doi: 10.1056/NEJM198903233201203] [Medline: 2784193]
4. Varner MW, Silver RM, Rowland Hogue CJ, Willinger M, Parker CB, Thorsten VR, Eunice Kennedy Shriver National Institute of Child Health and Human Development Stillbirth Collaborative Research Network. Association between stillbirth and illicit drug use and smoking during pregnancy. Obstet Gynecol 2014 Jan;123(1):113-125 [FREE Full text] [doi: 10.1097/AOG.0000000000000052] [Medline: 24463671]
5. Szutorisz H, DiNieri JA, Sweet E, Egervari G, Michaelides M, Carter JM, et al. Parental THC exposure leads to compulsive heroin-seeking and altered striatal synaptic plasticity in the subsequent generation. Neuropsychopharmacology 2014 May;39(6):1315-1323 [FREE Full text] [doi: 10.1038/npp.2013.352] [Medline: 24385132]
6. Samhsa. 2016. Key substance use and mental health indicators in the United States: results from the 2015 National Survey on Drug Use and Health URL: https://www.samhsa.gov/data/sites/default/files/NSDUH-FFR1-2015/NSDUH-FFR1-2015/NSDUH-FFR1-2015.htm [accessed 2017-10-31] [WebCite Cache ID 6ucsyM4tM]
7. Samhsa. Rockville, MD; 2016. Results from the 2015 National Survey on Drug Use and Health: detailed tables URL: https://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015.htm [accessed 2017-11-01] [WebCite Cache ID 6ue7tohin]
8. Volkow ND, Compton WM, Wargo EM. The risks of marijuana use during pregnancy. J Am Med Assoc 2017 Jan 10;317(2):129-130. [doi: 10.1001/jama.2016.18612] [Medline: 27992628]
9. Brady TM, Ashley OS. Citeseerx.ist.psu. Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Sciences; 2005. Women in substance abuse treatment: results from the Alcohol and Drug Services Study (ADSS) URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.439.783&rep=rep1&type=pdf
10. Lipari RN, Park-Lee E, Van Horn S. Samhsa. Rockville, MD: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality; 2016 Sep 29. America's need for and receipt of substance use treatment in 2015 URL: https://www.samhsa.gov/data/sites/default/files/report_2716/ShortReport-2716.html [accessed 2017-11-01] [WebCite Cache ID 6ue8GPRcn]
11. Barry KL, Caetano R, Chang G, DeJoseph MC, Miller LA, O'Connor MJ, et al. CDC. Atlanta, GA: CDC; 2009. Reducing alcohol-exposed pregnancies: a report of the National Task Force on fetal alcohol syndrome and fetal alcohol effect URL: https://www.cdc.gov/ncbddd/fasd/documents/redalcohpreg.pdf [accessed 2017-11-01] [WebCite Cache ID 6ue8TuYOZ]
12. Yarnall KS, Pollak KI, Østbye T, Krause KM, Michener JL. Primary care: is there enough time for prevention? Am J Public Health 2003 Apr;93(4):635-641. [Medline: 12660210]
13. Aalto M, Pekuri P, Seppä K. Obstacles to carrying out brief intervention for heavy drinkers in primary health care: a focus group study. Drug Alcohol Rev 2003 Jun;22(2):169-173. [doi: 10.1080/09595230100100606] [Medline: 12850903]
14. Chapin J, Root W. Improving obstetrician-gynecologist implementation of smoking cessation guidelines for pregnant women: an interim report of the American College of Obstetricians and Gynecologists. Nicotine Tob Res 2004 Apr;6(2):253-257. [doi: 10.1080/14622200410001669123]

15. Grimley DM, Bellis JM, Raczynski JM, Henning K. Smoking cessation counseling practices: a survey of Alabama obstetrician-gynecologists. South Med J 2001 Mar;94(3):297-303. [Medline: 11284517]
16. Helwig AL, Swain GR, Gottlieb M. Smoking cessation intervention: the practices of maternity care providers. J Am Board Fam Pract 1998;11(5):336-340. [Medline: 9796762]
17. Melvin CL, Dolan-Mullen P, Windsor RA, Whiteside Jr HP, Goldenberg RL. Recommended cessation counselling for pregnant women who smoke: a review of the evidence. Tob Control 2000;9(Suppl 3):iii80-iii84. [Medline: 10982917]
18. D'Amico EJ, Paddock SM, Burnam A, Kung FY. Identification of and guidance for problem drinking by general medical providers: results from a national survey. Med Care 2005;43(3):229-236. [Medline: 15725979]
19. Moore BA, Fazzino T, Garnet B, Cutter CJ, Barry DT. Computer-based interventions for drug use disorders: a systematic review. J Subst Abuse Treat 2011 Apr;40(3):215-223 [FREE Full text] [doi: 10.1016/j.jsat.2010.11.002] [Medline: 21185683]
20. 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]
21. Beatty JR, Chase SK, Ondersma SJ. A randomized study of the effect of anonymity, quasi-anonymity, and Certificates of Confidentiality on postpartum women's disclosure of sensitive information. Drug Alcohol Depend 2014 Jan 01;134:280-284. [doi: 10.1016/j.drugalcdep.2013.10.016] [Medline: 24246900]
22. Pew Internet & American Life Project. Pewinternet. 2016. Mobile fact sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2017-02-07] [WebCite Cache ID 6o6Pvw4pt]
23. 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]
24. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
25. Mobilesquared.co. 2010. Conversational advertising: a report from mobileSQUARED URL: http://mobilesquared.co.uk/media/27820/Conversational-Advertising_SinglePoint_2010.pdf [accessed 2017-10-31] [WebCite Cache ID 6ucubtdyK]
26. Naughton F, Prevost AT, Gilbert H, Sutton S. Randomized controlled trial evaluation of a tailored leaflet and SMS text message self-help intervention for pregnant smokers (MiQuit). Nicotine Tob Res 2012 May;14(5):569-577. [doi: 10.1093/ntr/ntr254] [Medline: 22311960]
27. Abroms LC, Johnson PR, Heminger CL, Van Alstyne JM, Leavitt LE, Schindler-Ruwisch JM, et al. Quit4baby: results from a pilot test of a mobile smoking cessation program for pregnant women. JMIR Mhealth Uhealth 2015 Jan 23;3(1):e10 [FREE Full text] [doi: 10.2196/mhealth.3846] [Medline: 25650765]
28. Fusch PI, Ness LR. Are we there yet? Data saturation in qualitative research. The Qual Report 2015;20(9):1408-1416 [FREE Full text]
29. Pollick SA, Beatty JR, Sokol RJ, Strickler RC, Chang G, Svikis DS, et al. Acceptability of a computerized brief intervention for alcohol among abstinent but at-risk pregnant women. Subst Abus 2015;36(1):13-20 [FREE Full text] [doi: 10.1080/08897077.2013.857631] [Medline: 24266770]
30. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-339 [FREE Full text] [doi: 10.2307/249008]
31. Russell MA, Stapleton JA, Jackson PH, Hajek P, Belcher M. District programme to reduce smoking: effect of clinic supported brief intervention by general practitioners. Br Med J (Clin Res Ed) 1987 Nov 14;295(6608):1240-1244 [FREE Full text] [Medline: 3120963]
32. Higgins-Biddle JC, Babor TF, Mullahy J, Daniels J, McRee B. Alcohol screening and brief intervention: where research meets practice. Conn Med 1997 Sep;61(9):565-575. [Medline: 9334512]
33. Babor TF, McRee BG, Kassebaum PA, Grimaldi PL, Ahmed K, Bray J. Screening, Brief Intervention, and Referral to Treatment (SBIRT): toward a public health approach to the management of substance abuse. Subst Abus 2007;28(3):7-30. [doi: 10.1300/J465v28n03_03] [Medline: 18077300]
34. Miller WR, Rollnick S. Motivational Interviewing: Preparing People for Change. 2nd ed. New York: The Guilford Press; 2002.
35. Ondersma SJ, Beatty JR, Svikis DS, Strickler RC, Tzilos GK, Chang G, et al. Computer-delivered screening and brief intervention for alcohol use in pregnancy: a pilot randomized trial. Alcohol Clin Exp Res 2015 Jul;39(7):1219-1226 [FREE Full text] [doi: 10.1111/acer.12747] [Medline: 26010235]
36. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckman J, editors. Action Control: From Cognition to Behavior. Berlin Heidelberg: Springer-Verlag; 1985:11-39.
37. 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. [Medline: 11392867]
38. Miles MB, Huberman AM, Saldana J. Chapter 11: Drawing and verifying conclusions. In: Qualitative Data Analysis. A Methods Sourcebook. Los Angeles: SAGE Publications, Inc; 2014:275-322.

Abbreviations

eSBIRT: electronic administration of SBIRT
IRB: institutional review board
PC: personal computer
SBIRT: screening, brief intervention, and referral for treatment

Edited by G Eysenbach; submitted 26.04.17; peer-reviewed by A Berman, P Johnson, R Shrestha; comments to author 29.06.17; revised version received 23.08.17; accepted 10.09.17; published 08.11.17.

Please cite as:
Gray J, Beatty JR, Svikis DS, Puder KS, Resnicow K, Konkel J, Rice S, McGoron L, Ondersma SJ
Electronic Brief Intervention and Text Messaging for Marijuana Use During Pregnancy: Initial Acceptability of Patients and Providers
JMIR Mhealth Uhealth 2017;5(11):e172
URL: http://mhealth.jmir.org/2017/11/e172/
doi:10.2196/mhealth.7927
PMID:29117931

Original Paper

Using Android and Open Data Kit Technology in Data Management for Research in Resource-Limited Settings in the Niger Delta Region of Nigeria: Cross-Sectional Household Survey

Omosivie Maduka[1*], MBBS, FMCPH, FRSPH; Godwin Akpan[2*], BEng, MSc, PhD; Sylvester Maleghemi[2], MBBS, MSc

[1]Department of Preventive and Social Medicine, College of Health Sciences, University of Port Harcourt, Port Harcourt, Nigeria

[2]World Health Organization, Port Harcourt, Nigeria

[*]these authors contributed equally

Corresponding Author:
Omosivie Maduka, MBBS, FMCPH, FRSPH
Department of Preventive and Social Medicine
College of Health Sciences
University of Port Harcourt
Alakahia, Choba
Port Harcourt, 50001
Nigeria
Phone: 234 8033298096
Email: omosivie.maduka@uniport.edu.ng

Abstract

Background: Data collection in Sub-Saharan Africa has traditionally been paper-based. However, the popularization of Android mobile devices and data capture software has brought paperless data management within reach. We used Open Data Kit (ODK) technology on Android mobile devices during a household survey in the Niger Delta region of Nigeria.

Objective: The aim of this study was to describe the pros and cons of deploying ODK for data management.

Methods: A descriptive cross-sectional household survey was carried out by 6 data collectors between April and May 2016. Data were obtained from 1706 persons in 601 households across 6 communities in 3 states in the Niger Delta. The use of Android mobile devices and ODK technology involved form building, testing, collection, aggregation, and download for data analysis. The median duration for data collection per household and per individual was 25.7 and 9.3 min, respectively.

Results: Data entries per device ranged from 33 (33/1706, 1.93%) to 482 (482/1706, 28.25%) individuals between 9 (9/601, 1.5%) and 122 (122/601, 20.3%) households. The most entries (470) were made by data collector 5. Only 2 respondents had data entry errors (2/1706, 0.12%). However, 73 (73/601, 12.1%) households had inaccurate date and time entries for when data collection started and ended. The cost of deploying ODK was estimated at US $206.7 in comparison with the estimated cost of US $466.7 for paper-based data management.

Conclusions: We found the use of mobile data capture technology to be efficient and cost-effective. As Internet services improve in Africa, we advocate their use as effective tools for health information management.

(JMIR Mhealth Uhealth 2017;5(11):e171) doi:10.2196/mhealth.7827

KEYWORDS

mobile phones; technology; Africa

Introduction

Rationale

Data collection may be of a routine nature such as from registers or ad hoc such as from evaluations or research activities. Effective data management is essential for evidence-based public health interventions and contributes to sustainable development. For a data collection effort to meet the needs of policy makers, projects, programs, and other stakeholders, it must be comprehensive, accurate, cost-effective, and timely.

Data collection for research purposes in many locations in Sub-Saharan Africa has been traditionally paper-based [1].

Questions are typed on data processing machines and multiple copies printed and photocopied based on the sample size. These are self-administered or interviewer-administered, and information filled in these data sheets/questionnaires are manually entered into data processing software and analyzed. This strategy is fraught with several shortcomings spanning all stages of data management. During data collection, issues of nonresponse to sections of the tool, data entry errors, and false data entries are possibilities. Data collation may also be subject to errors during transfer to databases. These necessitate that a great deal of person-hours are spent in *data cleaning*, which is the painstaking process of correcting or removing incorrect, duplicate, or corrupt data entries in an attempt to improve the validity of research findings. Data collation and cleaning is especially challenging when dealing with large sample sizes [2].

Mobile-based data capture software attempt to resolve these problems and improve the validity and reliability of health information. Data capture software are linked to databases, servers, or repositories for easy uploads and retrieval to the data analysis software, thus eliminating the need for data entry. Furthermore, many of them also have inbuilt data quality checks that improve data completeness and accuracy. Although various versions of the software have been in existence for decades and are actively used by researchers in developed countries, its usefulness has only recently been explored in resource-limited settings [1,3-10].

The growing popularity of Android mobile devices in Sub-Saharan Africa, the explosion in the telecommunications market, and the development of user-friendly mobile apps have brought mHealth within reach of populations in Sub-Saharan Africa. Nigeria, with a population of over 187 million people, has been a haven for the expansion of telecommunication and mobile phone companies. These companies have flooded the market with affordable brands of Android mobile phones with the cheapest going for as low as US $30.00. Furthermore, these devices can be used multiple times over many years, making it a cost-effective tool for data collection as part of research efforts.

In spite of this, use of mobile technology for research is not very popular among researchers in the country. This is because its use is perceived to be cost-intensive, heavily reliant on information technology expertise, and cumbered by unstable electricity and Internet connectivity. As such, it is often restricted to use by international nongovernmental agencies. These perceived bottlenecks can, however, be quickly overcome so that researchers are able to enjoy their obvious benefits.

Objective

As part of research efforts during the Department for International Development (DfID)–funded Climate Impact Research, Capacity and Leadership Enhancement (CIRCLE) research fellowship, we undertook to use Open Data Kit (ODK) technology. ODK was first developed by researchers at the University of Washington's Department of Computer Science and Engineering in 2008. It has enjoyed a wide application of uses in epidemiology. We implemented its use for data management in a community-based household survey to assess and compare morbidity and mortality patterns between persons residing in communities exposed to gas-flaring and those residing in communities without any oil exploration activities. We sought to characterize the pros and cons of using the ODK software and Android technology for data management in a large study, in terms of personnel, accuracy, time, and cost.

Methods

Study Design and Study Area

A descriptive cross-sectional household community survey was carried out over a three-week period between April and May 2016 among a sample of 1706 persons spanning 600 households in 6 communities in 3 states in the Niger Delta region of Nigeria. The Niger Delta region of Nigeria, as defined by the Nigerian Government, occupies about 70,000 km² and makes up 7.5% of Nigeria's landmass. It comprises the 9 oil-producing states in the country, namely, Bayelsa, Rivers, Delta, Akwa Ibom, Cross River, Edo, Abia, Imo, and Ondo. These states are home to about 31 million people spanning over 40 ethnic groups and 185 Local Government Areas (LGAs), who speak about 250 different dialects. The study sites were in Mbodo-Aluu and Omuhiombia in Ikwerre LGA of Rivers State, Ibada-Elume and Oton in Sapele LGA of Delta State, Sampou in Kolokuma/Opokuma LGA, and Nedugo in Yenagoa LGA in Bayelsa State.

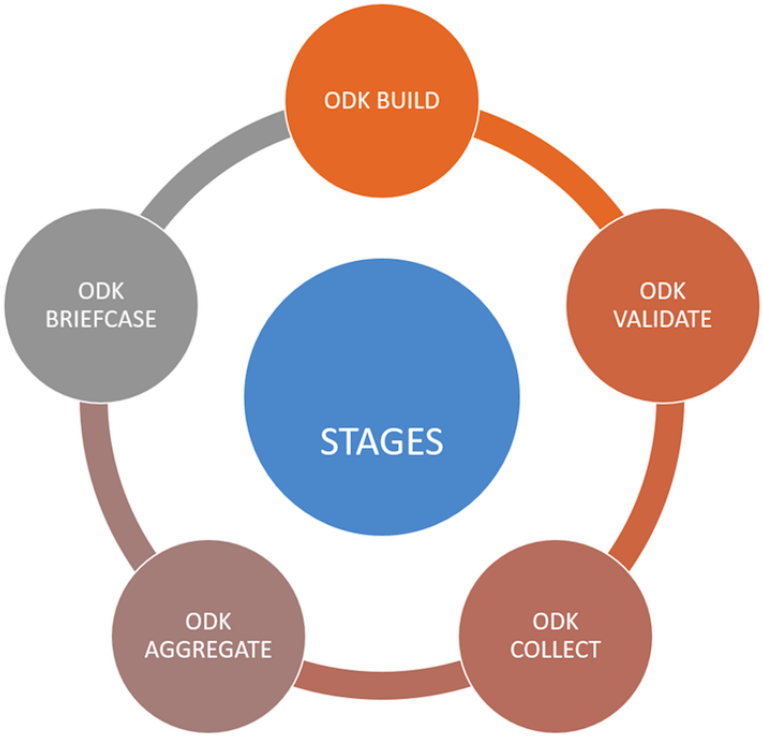
Study Instrument and Data Collection

The use of ODK uploaded on Android mobile devices involved the following steps (Figure 1): form building, validation/testing, training, data collection, collation (aggregate and briefcase), and data analysis. The original household questionnaire was a 15-page, 145-item interviewer-administered tool in 7 subsections. Building of this form on ODK included a consent page for collecting signatures, indicating that informed consent has been given; global positioning system (GPS) capturing of household coordinates; sociodemographic information; biological measurements; and information on household environmental characteristics, morbidity, and mortality.

Six field data collectors participated in a 1-day training on the use of the e-questionnaires that had been uploaded onto the ODK collect application downloaded onto Android mobile phones. All the data collectors were experienced in the use of Android mobile phones. However, only 1 data collector was experienced with using the ODK collect application on Android phones. The training covered familiarization with ODK collect application, the e-questionnaire, how to take GPS coordinates, administration of informed consent, and taking biological measurements. The researchers facilitated various aspects of the 1-day training.

Eight Android mobile phones were granted by the World Health Organization (WHO) state office for data collection. The extra 2 phones served as backup in the event of malfunction or challenges with global system for mobile communication (GSM) mobile and data networks.

Figure 1. Stages in deployment of Open Data Kit for data collection.



Each data collector was assigned a target of 10 households per day per community. Field data collectors were trained to upload completed forms onto a secure server, with back-end access provided to only the research team lead. Data underwent "cleaning" to identify errors. Inaccurate date and time records were identified during data cleaning through a review of the start/end date and time records from the device. The data entries underwent a preliminary analysis on the server platform before downloading in Microsoft Excel format and exporting to SPSS version 21 (IBM Corp) for analysis.

Ethical approval for the research was obtained from the research ethics committee of the University of Port Harcourt (ethics approval number: UPH/CEREMAD/REC/04). In addition, each community gave permission for data collection through the community leaders.

Results

Characteristics of ODK Data Collection

A total of 8 ODK-enabled Android mobile phones were used by 6 data collectors to collect data from 1706 household members in 601 households. Data entries per device ranged from 33 (1.93%, 33/1706) to 482 (28.25%, 482/1706) among individuals and between 9 (1.5%, 9/601) and 122 (20.3%, 122/601) households. The most entries (470) were made by data collector 5 using only 1 device (Tables 1 and 2). In addition, there were only 2 respondents (0.12%, 2/1706) with missing data out of the 1706 persons recruited into the study.

ODK phones obtained GPS coordinates for all 601 households in the study communities. Using this, we were able to plot the geomap illustrating the location of each household across the study communities (Figure 2).

Table 1. Number of data entries for households and household occupants per device (N=1706).

Device ID	Households (N=601), n (%)	Individuals (N=1706), n (%)
3523xxxxxxxxxxx	107 (17.8)	303 (17.76)
3538xxxxxxxxxxx	120 (20.0)	168 (9.85)
3538xxxxxxxxxxx	122 (20.3)	482 (28.3)
3538xxxxxxxxxxx	109 (18.1)	303 (17.76)
3538xxxxxxxxxxx	10 (1.7)	34 (1.99)
3560xxxxxxxxxxx	26 (4.3	76 (4.45)
8628xxxxxxxxxxx	9 (1.5)	33 (1.93)
8636xxxxxxxxxxx	98 (16.3)	307 (18.00)

Table 2. Devices used by data collectors, number of households, and data entries made.

Interviewer and devices used	Number of entries (N=1705)	Number of households (N=601)
Interviewer 1	**N=310**	**N=110**
3523xxxxxxxxxx	303	107
3560xxxxxxxxxx	5	2
8636xxxxxxxxxx	2	1
Interviewer 2	**N=46**	**N=13**
3538xxxxxxxxxx	2	1
3538xxxxxxxxxx	9	2
3560xxxxxxxxxx	2	1
8628xxxxxxxxxx	33	9
Interviewer 3	**N=166**	**N=119**
3538xxxxxxxxxx	166	119
Interviewer 4	**N=340**	**N=120**
3538xxxxxxxxxx	3	1
3538xxxxxxxxxx	303	109
3538xxxxxxxxxx	34	10
Interviewer 5	**N=470**	**N=119**
3538xxxxxxxxxx	470	119
Interviewer 6	**N=374**	**N=120**
3560xxxxxxxxxx	69	23
8636xxxxxxxxxx	305	97

Figure 2. Geomap of all the households across the 6 communities in 3 states of the Niger Delta region of Nigeria.

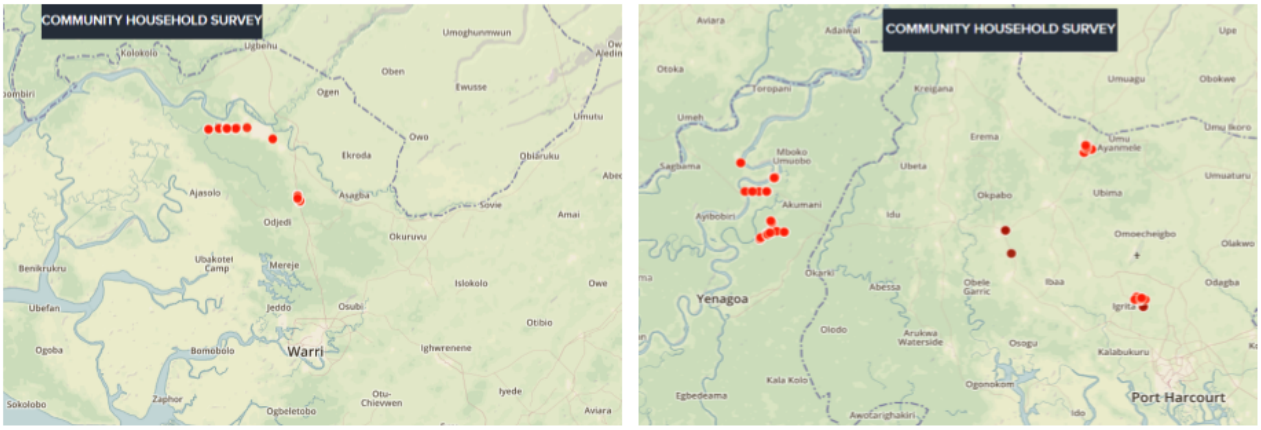


Table 3. Accuracy of date and time for Android mobile phones used for data collection (N=number of households where data were collected).

Device identity	Frequency (N=601)	
	Accurate timing (N=528), n (%)	Errors in date/time (N=73), n (%)
3523xxxxxxxxxx (N=107)	106 (99.1)	1 (0.9)
3538xxxxxxxxxx (N=120)	110 (91.7)	10 (8.3)
3538xxxxxxxxxx (N=122)	102 (83.6)	20 (16.4)
3538xxxxxxxxxx (N=109)	83 (76.1)	26 (23.9)
3538xxxxxxxxxx (N=10)	6 (60)	4 (40)
3560xxxxxxxxxx (N=26)	24 (92)	2 (8)
8628xxxxxxxxxx (N=9)	9 (100)	0 (0)
8636xxxxxxxxxx (N=98)	88 (90)	10 (10)

Date and Time Characteristics

Of the data collected using 8 devices over 601 households, 528 households (87.9%) had accurate date and time recordings for when data collection started and ended, whereas 73 (12.1%) had errors in date or time of data collection (Table 3). Of these, 20 (27.5%) errors were from an error in date settings on 1 mobile device, 2 (2.7%) errors were from an absence in start time, and the remaining 51 (69.8%) errors resulted from inability to obtain GPS coordinates (due to network down time) on one day, necessitating repeat visits the next day. Those forms could not be completed until all the required fields had been filled.

Only households with accurate date and time records were used (n=528) in calculating mean and median duration of data collection. The average duration of data collection for all households analyzed was 55.8 (standard deviation, SD=73.3) min with a median duration of 25.7 min. The shortest average duration of data collection was at Omerelu (25 min), whereas the longest average duration for data collection was at Oton (89.5 min). An average of 29.2 (SD 9.3) min was used for individual interviews with a median duration of 9.3 min (Table 4).

Table 4. Summary statistics for time spent collecting data using Open Data Kit across households and individuals in each community and for each Android mobile device.

Characteristics	Households			Individuals		
	Duration, mean (SD)	Median	Range	Duration, mean (SD)	Median	Range
Communities						
Ibada-Elume (n=99)	41.1 (56.8)	21.1	6.9-395.7	23.8 (35.5)	8.0	3.0-161.9
Mbodo (n=73)	74.9 (94.2)	37.8	11.1-416.7	34.7 (56.8)	14.0	4.3-400.6
Nedugo (n=94)	48.9 (55.9)	25.5	8.5-263.4	26.1 (41.9)	8.0	3.2-202.5
Omerelu (n=85)	25.0 (9.9)	22.7	8.7-57.7	10.2 (5.7)	8.0	3.0-25.7
Oton (n=78)	9.5 (105.6)	30.7	8.1-374.3	50.7 (74.4)	13.3	3.1-374.3
Sampou (n=99)	62.9 (70.9)	28.8	8.9-322.7	33.1 (45.6)	9.2	2.4-193.5
Android devices						
3523xxxxxxxxxx (n=106)	53.2 (31.2)	19.9	6.9-416.7	21.2 (9.2)	9.2	3.2-208.4
3538xxxxxxxxxx (n=110)	100.2 (90.8)	73.5	8.3-400.6	79.1 (74.8)	57.7	8.3-400.6
3538xxxxxxxxxx (n=102)	41.7 (40.6)	28.8	14.0-271.6	13.4 (19.7)	7.2	3.2-135.8
3538xxxxxxxxxx (n=83)	37.9 (39.6)	23.3	10.7-217.1	16.8 (24.6)	8.4	3.9- 158.8
3538xxxxxxxxxx (n=6)	29.3 (9.8)	31.8	14.6-38.7	7.9 (1.9)	8.3	4.9-9.7
3560xxxxxxxxxx (n=24)	17.2 (8.6)	14.2	8.9-47.0	6.4 (3.2)	5.6	2.4-16.40
8628xxxxxxxxxx (n=9)	194.4 (104.0)	181.6	46.6-322.7	55.7 (30.0)	60.5	11.6-91.0
8636xxxxxxxxxx (n=88)	34.8 (58.4)	18.8	8.1-368.0	11.7 (18.4)	6.6	3.0-122.7
All data collection (n=528)	55.8 (73.3)	25.7	6.9-416.7	29.2 (9.3)	9.3	2.4-400.6

Table 5. Implementation cost (comparing cost elements between paper-based and Open Data Kit phone).

Stages of data management	Android-based mobile devices (US $)	Paper-based (US $)	Comments
Questionnaire design and pretesting	0	0.0	This was done by the investigators
Purchase/hire of Android mobile phones	0 (300)[c]	0.0	Mobile phones were granted for data collection by the WHO[a] state program office
Building of forms on ODK[b] platform	66.7	0.0	
Training of data collectors	33.3	33.3	
Mass production of questionnaires of 15 pages (2000 copies)	0.0	100.0	US $0.03 per page for 15 pages and 2000 copies
Data entry and cleaning	66.7	333.3	Data cleaning alone needed for ODK[b]
Data download and cleaning	0.0	0.0	
Airtime and Internet data plans	40.0	0.0	US $5.0 per device
Total estimated costs	206.7 (506.7)[c]	466.7	

[a]WHO: World Health Organization.

[b]ODK: Open Data Kit.

[c]Cost of deployment, should the project have needed to purchase the same quality of Android phones obtained for use from the WHO state office.

Comparing Cost of ODK and Paper-Based Data Collection

A cost analysis conducted showed the costs of deploying ODK on Android mobile phones for data management to be US $206.7 compared with the cost of paper-based data management of US $466.7 using already existing mobile phones. However, when the cost analysis factored in phone purchases, the cost of deploying ODK increased to US $506.7. The major contributors to the cost of using ODK on Android mobile phones, aside from purchase of the phone, included building of the forms, back-end data management, and Internet access. The major cost considerations for paper-based data management are data entry and cleaning and mass production of the data sheets (Table 5).

Discussion

Principal Findings

Our study identified some pros and cons for the use of ODK software on Android mobile devices for data collection in a resource-limited setting. The pros included a comparatively shorter time on the field, improved rigor and accuracy of data collection, effective collection of GPS data and consent signatures, real-time monitoring of data entries, and reduced time lag between data collection and analysis in comparison with paper-based data collection. The cons we identified were related to the quality of the Android phones, errors in date and time recorded for interviews on some Androids, challenges with fluctuations in the GSM mobile and data networks leading to some delays in completing and uploading forms, and the increase in implementation costs if we were to purchase Android phones rather than leverage on existing ones. Our ability to leverage on existing/available Android devices facilitated implementation costs that were lower than what would have been incurred on paper-based data management. We also took advantage of scarce expertise in programming for building the forms and providing technical assistance in the use of the ODK collect and Ona platforms. Another downside of using ODK on Android mobile phones from our experience was that some interviews seemed to take as much as 7 hours to complete. This resulted from the fact that a few data collectors encountered challenges with Internet connectivity necessary for capturing GPS coordinates.

Android-based mobile phones offer additional capabilities, including built-in GPS functionality and other applications that can be integrated into electronic data collection such as digital photography and automated timestamp information. Mobile information technology also enables other data sources, including census and mapping data, and other tools for data visualization, including Google Maps, to be more readily integrated into the process of data collection, reporting, and analysis [3,11,12]. The multiple uses of mobile devices for message and calls, social media interactions, and data capture and management among others makes it a very useful tool for virtually all public health interventions. However, the quality of the Android mobile device comes into play as a factor affecting accuracy of data collection especially with regard to accuracy of GPS coordinates and data and time settings.

Limitations

The challenge related to the cost of the Android mobile device can be surmounted by uploading the ODK software on the Android devices owned by the research team and data collectors, thus eliminating the need to purchase Android phones. Researchers will, however, need to be deliberate about using medium-to-high grade Android phones so as to eliminate errors such as those we experienced with date and time on some of the phones used. Cellular and data service providers often do not have coverage in certain communities, and this can provide

a limitation to data collection with Android phones. Fluctuations in Internet service may also lead to delays in completing data entry and uploading completed forms. This limitation played out in our study such that some Android devices recorded as much as 7 hours of interview for some households because of poor data connection to finalize and upload data forms. We attempted to manage this limitation by providing multiple subscriber identity module (SIM) cards for each data collector depending on the cellular networks available in the community. However, in some situations, we resorted to repeat visits to the households to collect GPS coordinates, whereas in others, we collected the data offline and uploaded it later once data network was restored [5].

There were differences between the number of interviews conducted between devices and data collectors. This can be accounted for by the use of 8 mobile phones by 6 data collectors in the course of data collection. The extra phones came in handy when a phone malfunctioned or when there were network outages affecting one service provider and necessitating a switch to another service provider.

ODK software is one of the several mobile-based data management software and e-platforms for data collection. Other software such as epi-survey and form hub are also viable options for data management in resource-limited settings. With improvements in quality of Android phones, mobile and data service provision, and technical skill, we foresee health development practitioners and researchers in resource-limited settings relying more on paperless data management.

Conclusions

Paperless data management has obvious benefits over paper-based data management and is found to be quite useful even in resource-limited settings with prospects for use in large-scale regional and national surveys, active and passive surveillance, and other epidemiology activities. As cellular networks improve, we are likely to experience an increasing shift from paper-based to paperless health management information systems.

Acknowledgments

This research was supported by funding from the Department for International Development under the Climate Impact Research, Capacity and Leadership Enhancement program.

Conflicts of Interest

None declared.

References

1. King C, Hall J, Banda M, Beard J, Bird J, Kazembe P, et al. Electronic data capture in a rural African setting: evaluating experiences with different systems in Malawi. Glob Health Action 2014;7:25878 [FREE Full text] [Medline: 25363364]
2. Van den Broeck J, Argeseanu CS, Eeckels R, Herbst K. Data cleaning: detecting, diagnosing, and editing data abnormalities. PLoS Medicine 2005 Oct;2(10):e267 [FREE Full text] [doi: 10.1371/journal.pmed.0020267]
3. Tom-Aba D, Olaleye A, Olayinka AT, Nguku P, Waziri N, Adewuyi P, et al. Innovative technological approach to Ebola virus disease outbreak response in Nigeria using the open data kit and form hub technology. PLoS One 2015;10(6):e0131000 [FREE Full text] [doi: 10.1371/journal.pone.0131000] [Medline: 26115402]
4. Idowu B, Adagunodo R, Adedoyin R. Information technology infusion model for health sector in a developing country: Nigeria as a case. Technol Health Care 2006;14(2):69-77. [Medline: 16720950]
5. Tomlinson M, Solomon W, Singh Y, Doherty T, Chopra M, Ijumba P, et al. The use of mobile phones as a data collection tool: a report from a household survey in South Africa. BMC Med Inform Decis Mak 2009;9:51 [FREE Full text] [doi: 10.1186/1472-6947-9-51] [Medline: 20030813]
6. Lori JR, Munro ML, Boyd CJ, Andreatta P. Cell phones to collect pregnancy data from remote areas in Liberia. J Nurs Scholarsh 2012 Sep;44(3):294-301 [FREE Full text] [doi: 10.1111/j.1547-5069.2012.01451.x] [Medline: 22672157]
7. Guo T, Laksanasopin T, Sridhara A, Nayak S, Sia S. Mobile device for disease diagnosis and data tracking in resource-limited settings. Methods Mol Biol 2015;1256:3-14. [doi: 10.1007/978-1-4939-2172-0_1] [Medline: 25626528]
8. Schuttner L, Sindano N, Theis M, Zue C, Joseph J, Chilengi R, et al. A mobile phone-based, community health worker program for referral, follow-up, and service outreach in rural Zambia: outcomes and overview. Telemed J E Health 2014 Aug;20(8):721-728 [FREE Full text] [doi: 10.1089/tmj.2013.0240] [Medline: 24926815]
9. Afarikumah E. Electronic health in Ghana: current status and future prospects. Online J Public Health Inform 2014;5(3):230 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4943] [Medline: 24678382]
10. 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]
11. Raja A, Tridane A, Gaffar A, Lindquist T, Pribadi K. Android and ODK based data collection framework to aid in epidemiological analysis. Online J Public Health Inform 2014;5(3):228 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4996] [Medline: 24678381]

12. Haskew J, Kenyi V, William J, Alum R, Puri A, Mostafa Y, et al. Use of mobile information technology during planning, implementation and evaluation of a polio campaign in South Sudan. PLoS One 2015;10(8):e0135362 [FREE Full text] [doi: 10.1371/journal.pone.0135362] [Medline: 26252383]

Abbreviations

CIRCLE: Climate Impact Research, Capacity and Leadership Enhancement
DfID: Department for International Development
GPS: global positioning system
GSM: global system for mobile communication
LGAs: Local Government Areas
ODK: Open Data Kit
SD: standard deviation
SIM: subscriber identity module
WHO: World Health Organization

Edited by G Eysenbach; submitted 05.04.17; peer-reviewed by D Tom-Aba, C King; comments to author 24.05.17; revised version received 19.06.17; accepted 19.06.17; published 30.11.17.

Please cite as:
Maduka O, Akpan G, Maleghemi S
Using Android and Open Data Kit Technology in Data Management for Research in Resource-Limited Settings in the Niger Delta Region of Nigeria: Cross-Sectional Household Survey
JMIR Mhealth Uhealth 2017;5(11):e171
URL: http://mhealth.jmir.org/2017/11/e171/
doi:10.2196/mhealth.7827
PMID:29191798

Original Paper

Mobile Phone Use in Psychiatry Residents in the United States: Multisite Cross-Sectional Survey Study

Shih Gipson[1,2], MD; John Torous[2,3], MD; Robert Boland[2,4], MD; Erich Conrad[5], MD

[1]Department of Psychiatry, Boston Children's Hospital, Boston, MA, United States
[2]Department of Psychiatry, Harvard Medical School, Boston, MA, United States
[3]Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA, United States
[4]Department of Psychiatry, Brigham and Women's Hospital, Boston, MA, United States
[5]Department of Psychiatry, Louisiana State University Health Sciences Center, New Orleans, LA, United States

Corresponding Author:
Shih Gipson, MD
Boston Children's Hospital
Department of Psychiatry
300 Longwood Ave
Boston, MA, 02115
United States
Phone: 1 6173556680
Fax: 1 6177300319
Email: mariegipsonmd@gmail.com

Abstract

Background: Mobile technology ownership in the general US population and medical professionals is increasing, leading to increased use in clinical settings. However, data on use of mobile technology by psychiatry residents remain unclear.

Objective: In this study, our aim was to provide data on how psychiatric residents use mobile phones in their clinical education as well as barriers relating to technology use.

Methods: An anonymous, multisite survey was given to psychiatry residents in 2 regions in the United States, including New Orleans and Boston, to understand their technology use.

Results: All participants owned mobile phones, and 79% (54/68) used them to access patient information. The majority do not use mobile phones to implement pharmacotherapy (62%, 42/68) or psychotherapy plans (90%, 61/68). The top 3 barriers to using mobile technology in clinical care were privacy concerns (56%, 38/68), lack of clinical guidance (40%, 27/68), and lack of evidence (29%, 20/68).

Conclusions: We conclude that developing a technology curriculum and engaging in research could address these barriers to using mobile phones in clinical practice.

(JMIR Mhealth Uhealth 2017;5(11):e160) doi:10.2196/mhealth.7146

KEYWORDS

technology; graduate medical education; mobile phone; psychiatry

Introduction

Mobile technology ownership is common in the United States with 77% of the population owning a mobile phone [1]. In learning and in practice, medical students and trainees also have adopted using mobile devices. However, little is known about how psychiatry residents use mobile technology, particularly mobile phones, in their training. The current data on how medical trainees use mobile devices as a part of their education are limited. A study from Canada has found that medical students receive instructions on searching and assessing primary medical literature, but few receive formal instruction on non-traditional tools such as general use search engines, Wikipedia, or social media [2]. The authors suggest that teaching about nontraditional sources may enhance the current curriculum. This was underscored by another survey of 62 psychiatry residents which showed that 68% primarily use online resources for education rather than printed materials [3]. Given that majority of the population use mobile apps to access digital media, it seems likely that medical students and residents also

access clinical resources via their phones and apps [4]. However, it is unclear how residents and fellows in psychiatry use their mobile devices as a clinical resource. In 2014, Boruff et al [5] conducted a study to understand how Canadian medical students, residents, and faculty utilized their mobile phones to answer clinical questions and medical information. They found that mobile phones and tablets were broadly used in clinical settings and medical students were more likely to use these technologies. This is not a surprising finding due to the convenience and efficiency of mobile devices when retrieving information from the Internet. In a 2016 study, Gagnon et al [6] performed a systematic review of available literature to understand factors that influenced health care professionals in adapting mobile devices in their practice. They found that perceived usefulness and ease of use were some of the major factors in adapting mobile technologies among clinicians.

Although residents and fellows (trainees) are taught about protecting patient information in electronic medical records (EMRs), mobile technologies present new challenges that are distinct from those related to EMRs. Mobile devices like smartphones can be used interchangeably for professional and personal use including looking up patient information, sharing a photo on social media, and searching clinical decision support information. The use of these devices presents a new challenge to clinicians in integrating mobile phones in clinical use. These challenges include boundary issues as well as maintaining patient data safety. Although younger generations of psychiatry residents and fellows may often be considered tech savvy, their mobile phone use in clinical settings remains unclear.

To date, the most comprehensive data on mobile phone use in clinical practice are from a broad survey of physicians and nurses in the United Kingdom and psychiatry-specific data have not been reported [7]. Thus, in this study, we aim to provide an understanding on how psychiatric residents and fellows are using mobile phones in their clinical education and what barriers they perceive to this use.

Methods

To understand how psychiatry residents and fellows use mobile technology in clinical practice, we performed a multisite survey in 2 psychiatry residency programs in different regions of the United States, which included Louisiana State University Health Sciences Center New Orleans, Louisiana, and Longwood psychiatry residency in Boston, Massachusetts. These surveys were approved by the institutional review board of each site, and the surveys were conducted anonymously and voluntarily. No compensation was provided to participate in the survey.

Paper surveys were provided to psychiatry trainees during grand rounds for a 4-week period. Surveys in New Orleans, Louisiana, were collected in May-June 2015, and then collected in Boston, Massachusetts, in November 2015. An electronic version of the survey on RedCap was available in New Orleans, Louisiana. There were a total of 26 questions on this survey, developed by the authors, which inquired about their mobile technology use in administrative tasks, communication, and treatment planning. Participants were also provided with text boxes to elaborate on apps they used in administrative tasks, communication, and treatment planning. We also surveyed residents on the possible barriers in implementing mobile technology in clinical practice through multiple choice as well as an option to write in barriers that were not included in the available choices. Barriers provided were based on a previous study that identified barriers in using electronic resources among medical students on their psychiatry clerkship [3]. The survey questions have been listed in Table 1. The completed surveys were subsequently entered and analyzed using Microsoft Excel on an encrypted and password-protected computer.

Results

In New Orleans, 25 out of 41 trainees participated in this survey, and in Boston, 43 out of 52 trainees participated, which gave a total of 68 participants.

All the 68 participants owned a mobile phone or tablet. Moreover, 22 (32%) trainees practiced mostly in an outpatient setting and 41 (60%) practiced mostly in an inpatient setting. In addition, 5 (7%) trainees reported that they practice in a setting other than inpatient and outpatient psychiatry. A total of 54 (79%) of our participants reported that they use a mobile device to access protected patient information. Furthermore, 13 (19%) participants reported that they did not use a mobile device to access protected patient information, and 1 (1%) trainee did not respond to this question. Of the participants, 29 (43%) trainees reported using mental health-related apps to access protected patient information on their mobile device. Moreover, 40 (59%) trainees reported using an Internet site to access protected patient information.

A total of 16 (24%) trainees reported using a mobile device to communicate with patients, and 38 (56%) trainees reported that they did not use a mobile device to communicate with patients. Moreover, 14 (21%) participants did not respond when asked whether they used a mobile device to communicate with patients, and 38 (56%) trainees reported using a desktop computer to communicate with patients. Furthermore, 28 (41%) residents and fellows reported that they did not use a desktop computer to communicate with patients, and2 (3%) participants did not respond to this question. In addition, 26 (38%) trainees used a mobile device when implementing medication regimens, and 42 (62%) trainees did not use a mobile device when implementing medication regimens. A total of 16 (24%) participants reported using apps to implement medication regimens, and 15 (22%) trainees reported using an Internet site to implement medication regimens. Only 6 (9%) trainees reported using a mobile device to implement psychotherapy plans, and 61 (90%) trainees did not use a mobile device when implementing psychotherapy plans. Moreover, 1 (1%) trainee did not respond to this question. A total of 3 (4%) trainees used apps, and 2 (3%) trainees used Internet sites to implement psychotherapy plans.

Furthermore, 30 (44.1%) trainees used a mobile device to manage their clinic schedule; 38 (56%) trainees did not use a mobile device to manage their clinic schedule; 30 (44%) trainees intended to use a mobile device to manage their clinic schedule; and 38 (56%) trainees did not intend to use a mobile device to manage their clinic schedule.

Table 1. Survey questions.

Survey questions	Answer choices
How many years old are you?	_____ years old
Which one of the following genders do you most closely identify?	Male
	Female
	Transgender
	Prefer not to specify
Which one label do you most closely identify?	Attending physician
	Resident physician
	Medical student
	Other (please specify): _____
In which one practice setting do you spend most of your clinical hours?	Outpatient
	Inpatient
	Other
Do you own a smartphone or tablet?	Yes
	No
Do you intend to purchase a smartphone or tablet in the next 6 months?	Yes
	No
Do you use a mobile device (such as smartphone or tablet) to access protected patient information, such as their chart or their e-mail messages to you?	Yes
	No
Which methods do you use to access protected patient information on your mobile device? (Check all that apply)	Apps
	Internet Sites
	Other
Do you intend to use a mobile device to access protected patient information, such as their chart or their e-mail messages to you?	Yes
	No
Do you use a mobile device to communicate with patients?	Yes
	No
Do you use a desktop computer to communicate with patients?	Yes
	No
Do you use a mobile device when implementing medication regimens?	Yes
	No
Which methods do you use to implement medication regimens for your patient? (Check all that apply)	Apps
	Internet Sites
	Other
Do you use a mobile device when implementing psychotherapy plans?	Yes
	No
Which methods do you use to implement psychotherapy regimens for your patient? (Check all that apply)	Apps
	Internet Sites
	Other
Do you use a mobile device to manage your clinic schedule?	Yes
	No
Do you intend to use a mobile device to manage your clinic schedule?	Yes
	No

Survey questions	Answer choices
Do you use a mobile device to communicate with other staff?	Yes
	No
Which modality do you communicate with your mobile device? (Check all that apply)	Email
	Instant Messaging
	Text Messaging
	Call
	Other
In the last 3 months, have you recommended patients to use supplemental apps to their current medication management?	Yes
	No
Why or why not?	Text box
In the last 3 months, have you recommended patients to use supplemental apps to their current psychotherapy plan?	Yes
	No
Why or why not?	Text box
In the last 3 months, have you recommended online resources to patients?	Yes
	No
Why or why not?	Text box
What do you feel are the greatest barriers for using mobile devices in the care of patients? (Select up to three)	Privacy
	Safety
	Liability
	Cost
	Too much data
	Lack of evidence
	Lack of reimbursement
	Lack of clinical guidance

A total of 61 (90%) trainees used a mobile device to communicate with clinic staff and only 6 (9%) trainees did not. Moreover, 1 (1%) trainee did not answer this question. When asked about the modality used to communicate with clinic staff, 56 (84%) trainees reported using email, 10 (15%) trainees reported using an instant messaging app, 44 (66%) trainees reported using text, and 43 (64%) trainees reported calling clinic staff with their mobile devices.

In the last 3 months of completing the survey, 18 (27%) trainees reported recommending supplemental apps to patients in their current medication management, 48 (71%) trainees reported they do not, and 2 (3%) trainees did not respond to this question. Moreover, 19 (28%) trainees recommended supplemental apps to patients in their current psychotherapy plan, 48 (71%) trainees did not recommend any supplemental apps, and 1 (1%) trainee did not respond to this question. In the last 3 months of completing the survey, 38 (56%) trainees recommended online resources to patients, 27 (40%) trainees did not recommend online resources, and 3 (4%) trainees did not respond to this question.

When surveyed on perceived greatest barriers for using mobile devices in the care of patients, 38 (56%) trainees selected privacy, 27 (40%) trainees selected lack of clinical guidance, 20 (29%) trainees selected lack of evidence, 15 (22%) trainees selected liability, 12 (18%) trainees selected too much data, 11 (16%) trainees selected lack of reimbursement, 5 (7%) trainees selected cost, and 5 (7%) trainees selected safety.

Although the authors included questions on gender and age, there were a limited number of responses between both programs, and these questions were omitted in this study. Participants were provided with textboxes for various questions as listed in Table 1, but there were limited number of responses and they were omitted in this study. The above-described data are also provided in Tables 2 and 3.

Table 2. Survey responses.

Combined survey questions	n (%)
In which *one* practice setting do you spend most of your clinical hours?	
Outpatient	22 (32.35)
Inpatient	41 (60.3)
Other	5 (7.4)
Do you own a smartphone or tablet?	
Yes	68 (100)
No	0 (0)
Do you use a mobile device (such as a smartphone or tablet) to access protected patient information, such as their chart or their email messages to you?	
Yes	54 (79.41)
No	13 (19.1)
Did not answer	1 (1.5)
Which modality do you communicate with your mobile device?	
Email	56 (83.5)
Instant messaging app	10 (14.9)
Text	44 (65.7)
Call	43 (64.2)
In the last 3 months, have you recommended patients to use supplemental apps to their current medication management?	
Yes	18 (26.5)
No	48 (70.6)
Did not answer	2 (2.9)
In the last 3 months, have you recommended patients to use supplemental apps to their current psychotherapy plan?	
Yes	19 (27.9)
No	48 (70.6)
Did not answer	1 (1.5)
In the last 3 months, have you recommended online resources to patients?	
Yes	38 (55.9)
No	27 (39.7)
Did not answer	3 (4.4)

Table 3. Perceived barriers by trainees.

Question	Privacy	Safety	Liability	Cost	Lack of evidence	Too much data	Lack of reimbursement	Lack of clinical guidance
What do you feel are the greatest barriers for using mobile devices in the care of patients? (Select up to three), n (%)	38 (55.9)	5 (7.4)	15 (22)	5 (7.4)	20 (29.4)	12 (17.6)	11 (16.2)	27 (39.7)

Discussion

This multisite study provides the first results on both psychiatry trainee ownership and their use of mobile phones for clinical education and patient care. We found that all psychiatry residents who participated in this survey owned mobile phones, which exceeds the mobile phone usage in the general population. Further, the selected programs cover multiple hospitals within their city and are exposed to various types of EMRs. Although psychiatry residents own mobile phones, their reported use in our survey suggests that it is limited. The most common reported clinical use is communicating with clinical staff, especially for scheduling, and the second most common use is to access patient information. Although these uses may seem simple in that

mobile phones are being used for communication, each also raises educational opportunities for educators to consider. There have been recent efforts to ensure residents are educated about best practices with social media and websites such as Facebook, and in 2014, Dejong and Gorrindo [8] have discussed about texting and the professionalism principles when communicating with patients using this method. Among internal medicine residents, the use of short message service (SMS) texting as a means of communication has also raised the ethical question of whether the ease of use is a potential breach of patient privacy [9]. Overall, there remains a limited amount of literature on how we teach psychiatry residents best practices for using mobile devices in clinical care roles. If residents are using them primarily for scheduling purposes, it is not difficult to imagine that soon they will also be using mobile phones for direct clinical care roles as well – and indeed our survey suggests that some already are.

Several limitations of this study should be noted. All data were self-reported, and we did not verify through direct observation how residents are using mobile phones and apps in clinical care. Although our study is the first to examine this topic, our questions were not exhaustive and did not include details on app use when using with patients and communicating with other staff members. We attempted to collect demographic data including gender and age, but due to low compliance, the limited data gathered could not be analyzed. We also did not obtain details on how residents evaluate and consider whether to use or not use an app. We designed a quantitative design with yes/no questions to understand the basics of mobile device usage among trainees. Future iterations of our study with Likert scales and more qualitative metrics would be important to complete the understanding of mobile device use among trainees. Due to concerns of lack of participation, we also designed our survey with yes/no questions to help increase compliance. It is important to note that our study was limited to 2 sites, and it is difficult to assess how generalizable our results are outside of these 2 study sites.

Given this was a convenience sample and we have no information from the residents who did not choose to participate, it is, however, possible that the residents more interested in technology were more likely to participate. On the basis of our survey, future studies can consider larger trainee population of various specialties, while providing more flexibility in the survey including clarifying practice setting and providing more write in options. This can help provide further data and understanding on which areas of practice should be targeted through education as well as enhancing work flow using technology.

Despite a robust app development industry, clinical evidence for the effectiveness of apps in psychiatry remains nascent. Further, evaluating the safety and quality of apps is challenging. Thus, the fact that some residents are venturing into this largely unknown space of apps is notable. Generally, our survey suggests that trainees are at times hesitant in using mobile devices. However, a partial integration in their clinical practice is evident in that over 89% of trainees use their mobile devices to communicate with clinical staff. In addition, over 55% report using desktop computers to communicate with their patients, and the authors expect trainees to transition to using mobile devices in their patient communication due to increased availability and convenience of mobile devices. This reflects the need for educators and supervisors to consider at least inquiring if residents are using these tools and to prepare our trainees to potentially utilize these tools in practice.

Table 4. Suggested curriculum – Technology in Psychiatry Seminar (TIPS).

Topic	Description
"Anatomy" and "Physiology"	To understand basic inner works of technologies available to medical professionals and patients – includes understanding of basic technology terminology, how technologies are developed, and types of technology that are currently used in psychiatry
Telepsychiatry 1	Review literature in telepsychiatry Understand pros and cons of telepsychiatry
Telepsychiatry 2	Review of cases in telepsychiatry including tips to be used in practice Review basic video conferencing etiquette
Mobile technology	Review of current literature on mobile technology use in various psychopathology – mood disorder, anxiety, sleep, etc Discuss current applications of mobile technology
Professionalism, ethics, and privacy	To discuss important boundaries between the digital doctor–patient relationship, including texting/emailing etiquette, professional social media use, and maintaining a professional presence online Discuss maintenance of patient privacy when using technology
Research	Understand basic research study design in technology and review of current literature
Technology and psychopathology	Psychopathology relating to technology: Internet gaming disorder, Internet addiction disorder, Online gaming industry etc Using technology to treat psychopathology – mindfulness, CBT, and CBTi
Counseling patients on safe technology use	How to counsel patients on safe technology use – sleep hygiene, self-evaluation of apps, making sure it is a licensed clinician on telepsychiatry platform, etc

Along these lines, our survey results suggest several opportunities for psychiatric educators to offer guidance and support for trainees regarding the use of mobile technology. Residents raised several important concerns about apps including privacy, lack of clinical guidance, lack of evidence, and liability, among others. However, each of these topics represents a complex area as mobile technology for health care is advancing more rapidly than legislation, digital security, clinical trials, and clinical teaching can equal. Another challenge is that many educators may be less familiar or excited by mobile phone apps and other new technologies and thus lack the necessary experiences to educate residents. Given that "lack of clinical guidance" was selected as the number 2 barrier to using mobile technology in patient care, this suggests there is a need for a curriculum focused on digital and mobile technologies. Creating a curriculum in technology during graduate medical training could provide a platform for introducing trainees to basic concepts about current technologies including mobile technology and technology use in patient care [10]. In a technology curriculum, the core concept is not only to include basic concepts but also to provide principles for evaluating technology from a clinical perspective. Therefore, the authors propose the development of a technology curriculum that educates trainees in various digital psychiatry tools including EMR, telepsychiatry, mobile device use, and wearables. This potential technology curriculum is listed in Table 4. Further research and experience would be needed to find the most effective curriculum. Fundamentally, having a technology curriculum can allow trainees to better understand technologies relating to patient care without over incorporation in daily practice.

From this study, trainees are interested in using and have used mobile technology in practice. However, their education on how to use it remains lacking, and educators should consider including mobile technology education as a part of the residents' curriculum. Encouraging mobile technology research is also important. Until further research is conducted and empirical data are collected, basic clinical pearls and core concepts for mobile technology such as understanding the clinical evidence and privacy regulations could be the foundation of a technology curriculum. Although research- and evidence-based use of mobile technology is still in development, educators can teach residents how to dissect an app beyond its appearance and usability. Currently, The American Psychiatric Association's App Evaluation Task Force Committee is trying to understand mobile technology and create guidelines for clinicians to better evaluate apps for themselves and for patients [11].

Psychiatry trainees use mobile phones for their work, mainly for scheduling and administrative tasks. Though they remain hesitant to incorporate mobile health into their direct clinical care, partial integration of a mobile device is evident. Part of the hesitance among trainees appears to be due to the lack of education and guidance during their current training. It is imperative to prepare trainees to practice in the 21st century where mobile phone use is part of daily life so that they will be knowledgeable in helping patients understand potential impacts of mobile technology. Such preparation and training may require the development of a new curriculum and educational efforts.

Conflicts of Interest

None declared.

References

1. Pewinternet. 2017. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2017-09-20] [WebCite Cache ID 6tcdiS0hD]
2. Tran K, Morra D, Lo V, Quan SD, Abrams H, Wu RC. Medical students and personal smartphones in the clinical environment: the impact on confidentiality of personal health information and professionalism. J Med Internet Res 2014 May 22;16(5):e132 [FREE Full text] [doi: 10.2196/jmir.3138] [Medline: 24855046]
3. Snow C, Torous J, Gordon-Elliott J, Penzner J, Meyer F, Boland R. Use of electronic resources for psychiatry clerkship learning: a medical student survey. Acad Psychiatry 2016 Dec 19 Epub ahead of print(forthcoming). [doi: 10.1007/s40596-016-0647-3] [Medline: 27995539]
4. Perez S. TechCrunch. 2014 Aug 21. Majority Of Digital Media Consumption Now Takes Place In Mobile Apps URL: https://techcrunch.com/2014/08/21/majority-of-digital-media-consumption-now-takes-place-in-mobile-apps/ [accessed 2017-09-20] [WebCite Cache ID 6tce6vnPL]
5. Boruff J, Storie D. Mobile devices in medicine: a survey of how medical students, residents, and faculty use smartphones and other mobile devices to find information. J Med Libr Assoc 2014 Jan;102(1):22-30 [FREE Full text] [doi: 10.3163/1536-5050.102.1.006] [Medline: 24415916]
6. Gagnon MP, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
7. Mobasheri M, King D, Johnston M, Gautama S, Purkayastha S, Darzi A. The ownership and clinical use of smartphones by doctors and nurses in the UK: a multicentre survey study. BMJ Innov 2015 Oct 07;1(4):174-181. [doi: 10.1136/bmjinnov-2015-000062]
8. DeJong S, Gorrindo T. To text or not to text: applying clinical and professionalism principles to decisions about text messaging with patients. J Am Acad Child Adolesc Psychiatry 2014 Jul;53(7):713-715. [doi: 10.1016/j.jaac.2014.05.002] [Medline: 24954818]

9. Prochaska M, Bird A, Chadaga A, Arora V. Resident use of text messaging for patient care: ease of use or breach of privacy? JMIR Med Inform 2015 Nov 26;3(4):e37 [FREE Full text] [doi: 10.2196/medinform.4797] [Medline: 26611620]
10. Gipson SYMT, Kim JW, Shin AL, Kitts R, Maneta E. Teaching child and adolescent psychiatry in the twenty-first century. Child Adolesc Psychiatr Clin N Am 2017 Jan;26(1):93-103. [doi: 10.1016/j.chc.2016.07.004]
11. Torous J, Chan S, Gipson S. Psychnews. 2016 Jan 29. APA Task Force to Develop Guidelines to Assess Apps, Wearable Sensors URL: http://psychnews.psychiatryonline.org/doi/10.1176/appi.pn.2016.2a10 [accessed 2017-09-25] [WebCite Cache ID 6tk2izAWJ]

Abbreviations

CBT: Cognitive Behavioral Therapy
CBTi: Cognitive Behavioral Therapy for insomnia
EMRs: electronic medical records
SMS: electronic medical records
TIPS: Technology in Psychiatry Seminar

Edited by G Eysenbach; submitted 12.12.16; peer-reviewed by N Berry, A Baumel, F Lobban, P Yellowlees, M Nahum, M Larsen; comments to author 11.02.17; revised version received 26.05.17; accepted 04.08.17; published 01.11.17.

Please cite as:
Gipson S, Torous J, Boland R, Conrad E
Mobile Phone Use in Psychiatry Residents in the United States: Multisite Cross-Sectional Survey Study
JMIR Mhealth Uhealth 2017;5(11):e160
URL: http://mhealth.jmir.org/2017/11/e160/
doi:10.2196/mhealth.7146
PMID:29092807

Original Paper

Obstructive Sleep Apnea in Women: Study of Speech and Craniofacial Characteristics

Marina Tyan[1], MSc; Fernando Espinoza-Cuadros[1], PhD; Rubén Fernández Pozo[1], PhD; Doroteo Toledano[2], PhD; Eduardo Lopez Gonzalo[1], PhD; Jose Daniel Alcazar Ramirez[3], MD; Luis Alfonso Hernandez Gomez[1], PhD

[1]Signal Processing Applications Group, Signal, Systems and Radiocommunications Department, Universidad Politécnica de Madrid, Madrid, Spain
[2]Audio, Data Intelligence and Speech Group, Universidad Autónoma de Madrid, Madrid, Spain
[3]Respiratory Department, Sleep Unit, Hospital Quirón Salud de Málaga, Málaga, Spain

Corresponding Author:
Fernando Espinoza-Cuadros, PhD
Signal Processing Applications Group
Signal, Systems and Radiocommunications Department
Universidad Politécnica de Madrid
Avenida Complutense, 30
Madrid, 28040
Spain
Phone: 34 915495700 ext 212
Email: fernando@gaps.ssr.upm.es

Abstract

Background: Obstructive sleep apnea (OSA) is a common sleep disorder characterized by frequent cessation of breathing lasting 10 seconds or longer. The diagnosis of OSA is performed through an expensive procedure, which requires an overnight stay at the hospital. This has led to several proposals based on the analysis of patients' facial images and speech recordings as an attempt to develop simpler and cheaper methods to diagnose OSA.

Objective: The objective of this study was to analyze possible relationships between OSA and speech and facial features on a female population and whether these possible connections may be affected by the specific clinical characteristics in OSA population and, more specifically, to explore how the connection between OSA and speech and facial features can be affected by gender.

Methods: All the subjects are Spanish subjects suspected to suffer from OSA and referred to a sleep disorders unit. Voice recordings and photographs were collected in a supervised but not highly controlled way, trying to test a scenario close to a realistic clinical practice scenario where OSA is assessed using an app running on a mobile device. Furthermore, clinical variables such as weight, height, age, and cervical perimeter, which are usually reported as predictors of OSA, were also gathered. Acoustic analysis is centered in sustained vowels. Facial analysis consists of a set of local craniofacial features related to OSA, which were extracted from images after detecting facial landmarks by using the active appearance models. To study the probable OSA connection with speech and craniofacial features, correlations among apnea-hypopnea index (AHI), clinical variables, and acoustic and facial measurements were analyzed.

Results: The results obtained for female population indicate mainly weak correlations (r values between .20 and .39). Correlations between AHI, clinical variables, and speech features show the prevalence of formant frequencies over bandwidths, with F2/i/ being the most appropriate formant frequency for OSA prediction in women. Results obtained for male population indicate mainly very weak correlations (r values between .01 and .19). In this case, bandwidths prevail over formant frequencies. Correlations between AHI, clinical variables, and craniofacial measurements are very weak.

Conclusions: In accordance with previous studies, some clinical variables are found to be good predictors of OSA. Besides, strong correlations are found between AHI and some clinical variables with speech and facial features. Regarding speech feature, the results show the prevalence of formant frequency F2/i/ over the rest of features for the female population as OSA predictive feature. Although the correlation reported is weak, this study aims to find some traces that could explain the possible connection between OSA and speech in women. In the case of craniofacial measurements, results evidence that some features that can be used for predicting OSA in male patients are not suitable for testing female population.

(JMIR Mhealth Uhealth 2017;5(11):e169) doi:10.2196/mhealth.8238

KEYWORDS

obstructive sleep apnea; acoustics, speech; image processing, computer-assisted

Introduction

Sleep disorders are receiving increased attention as a cause of daytime sleepiness, impaired work, and traffic accidents and are associated with hypertension, heart failure, arrhythmia, and diabetes. The most common form of sleep-disordered breathing is the obstructive sleep apnea (OSA) syndrome, and it is characterized by an obstruction of the upper airway (UA) during sleep at the level of the pharynx, yielding partial (hypopnea) or total (apnea) breathing cessation episodes longer than 10 s at a time [1].

The gold standard for the diagnosis of OSA is a full overnight polysomnography (PSG) test [2] performed in an attended laboratory setting. PSG monitors electrophysiologic variables to score sleep stages and detect arousals and cardiorespiratory variables to detect complete (apnea) or near-complete (hypopnea) cessation of airflow. The OSA severity is determined based on the number of apnea and hypopnea episodes per hour of sleep or apnea-hypopnea index (AHI; mild defined as an AHI of 5-15, moderate as 15-30, and severe as ≥30).

However, PSG is expensive and time-consuming, and, furthermore, the recordings are performed in an unfamiliar environment for the patient. Therefore, faster, noninvasive, and less costly alternatives have been proposed for early OSA detection and severity assessment, such as unattended domiciliary sleep studies.

Although overweight and an excess of regional adipose tissue are considered major risk factors for OSA, there are also other interacting elements in OSA pathogenesis, such as craniofacial abnormalities and an altered UA structure, being approached by several studies since the early approaches by means of the analysis of magnetic resonance imaging [3] until the photometry over digital photographs of head [4,5]. Among OSA phenotype-related characteristics are dental occlusion, longer distance between the hyoid bone and the mandibular plane as described by Lowe and coworkers [6], and relaxed pharyngeal soft tissues and large tongue base as described by Schwab and coworkers [3], which generally cause a longer and more collapsible UA. Consequently, abnormal or particular speech in OSA patients may also be expected from the altered structure or function of their UA.

Therefore, several approaches to speech-based OSA detection have been developed since the acoustic perceptive analysis [7,8] until the most recent proposals for using automatic speech-processing techniques in OSA detection [9]. However, most of the previous mentioned publications have only focused on male subjects. To the best of our knowledge, there are no similar studies that concentrated on female OSA patients, and very few publications are available that discuss this issue [10,11].

Consequently, the main purpose of this paper was to study the potential connection between AHI and speech and facial features, focusing on a female population. Furthermore, we have also considered that it might be interesting to compare our results on male versus female patients. In that way, we can observe how the connection between OSA and speech and facial features can be affected by gender.

For an easy interpretation of our results, similar to [12], acoustic analysis is performed by evaluating formant frequencies and bandwidths on sustained phonations of vowel sounds. Facial features are extracted by identifying a set of relevant landmarks on subjects' images, following also a rather simple procedure similar to the one we presented in [9]. Statistical analysis using correlation coefficients is employed to evaluate the connection between speech and facial features with AHI. To gain a better understanding of this connection, we have used statistical contrasts (Mann-Whitney *U* tests) among OSA severity groups.

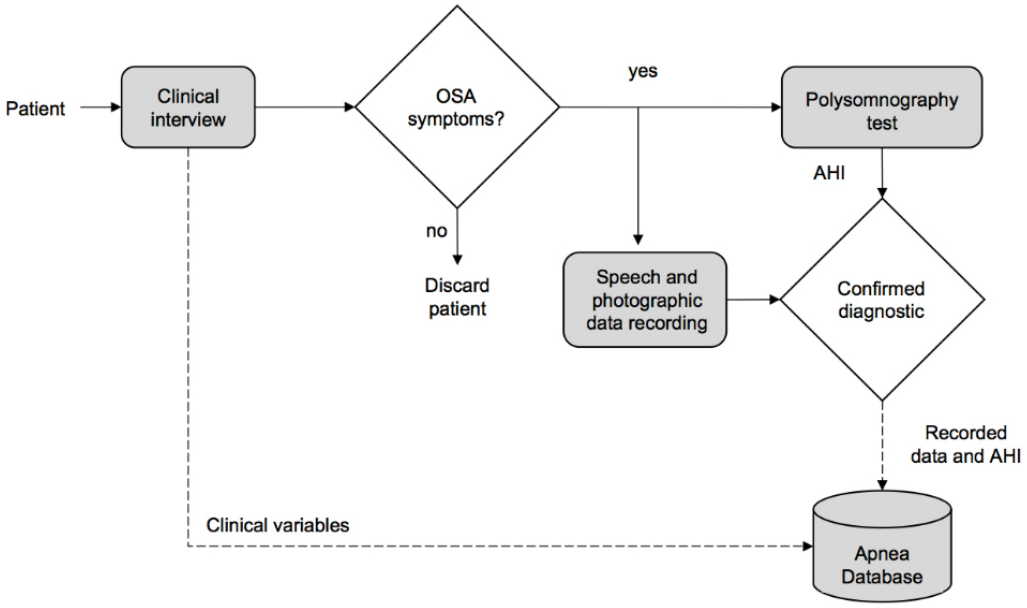
Methods

Subjects and Recording Procedure

Patients were provided by the Hospital Quirón Salud de Málaga (Spain). The subjects referred for PSG previously reported symptoms of OSA during a preliminary interview with a pneumonologist, such as excessive daytime sleepiness, snoring, choking during sleep, or somnolent driving. By means of this interview, the subjects' clinical history was obtained, and an exhaustive physical examination focusing on sleep-related symptoms, associated conditions, comorbidities, and anthropometrics measures was conducted and data collected. Subjects' weight and height were recorded when wearing light clothes. Body mass index (BMI) was calculated as the ratio of body weight (in kg) and the height (in m^2). Cervical perimeter (in cm) was also measured at the level of cricothyroid membrane. Most of the subjects are from Andalusia (southern Spain). The majority of subjects were white, with the exception of 1 Chinese. Exclusion criteria included subjects with no Andalusian dialect, subjects with a known history of syndromal craniofacial abnormalities, subjects who have had craniofacial surgery, ethnicity, and subjects with excessive facial hair that significantly obscured facial landmarks, as well as subjects with photograph capture errors (eg, inclination, bad position).

The diagnosis for each patient was confirmed by specialized medical stuff through standard overnight PSG test, obtaining the AHI on the basis of the number of apnea and hypopnea episodes. According to subjects' AHI, we defined three groups of OSA severity: low AHI (<10) indicates a healthy subject, AHI between 10 and 30 indicates mild OSA patient, whereas AHI above 30 is associated with severe OSA. These thresholds were defined to get balanced number of samples for our statistical contrast analysis. Figure 1 illustrates the data collection process.

Figure 1. Flowchart of recording data for apnea database. AHI: apnea-hypopnea index; OSA: obstructive sleep apnea.



Before the PSG test, all patients were taken to a separate room with adequate acoustic condition and the recording equipment for collecting speech and photographic data, after obtaining patients' consent. Speech and photographic data are explained as follows:

- Acoustic data: Sustained phonations of each Spanish vowel /a/, /e/, /i/, /o/, and /u/ were recorded from every subject at an upright or seated position and with a comfortable speech level in a quiet room. Recording equipment was a standard laptop computer equipped with an SP500 Plantronics headset microphone. Speech was recorded at a sampling frequency of 50 kHz and encoded in 16 bits. Afterwards, it was downsampled to 16 kHz before processing.
- Photographic data: Frontal and profile digital photographs of the head were obtained before the speech recordings, also at the same normal hospital room without any particular illumination condition. In contrast to the studies by Lee and coworkers [4,5], no special actions were taken beyond a simple control for patients' front and profile photographs and some instructions to guarantee that the neck area is visible in the profile image. No calibration action for allowing the conversion from pixel measurements to metric dimensions (eg, measuring the distance from the camera) was taken, and manual identification by palpation of facial landmarks was also avoided. A standard Logitech QuickCam Pro 5000 webcam was used to collect images with a size of 640 × 480 pixels and a color depth of 24 bits.

It is important to point out that the recording protocol was approved by the Institutional Review Committee of the Hospital Quirón Salud de Málaga and performed strictly following the ethical consideration of the medical center. The participants were notified about the research and their signed agreement was obtained.

After applying exclusion criteria, a total of 383 subjects (129 women and 254 men) were included in our study. The female population comprised 64 subjects in OSA group (AHI>10) and 65 in control group (AHI≥10). The male population comprised 168 subjects in OSA group (AHI>10) and 86 in control group (AHI≥10). Descriptive statistics of subjects under study are summarized in Table 1.

Acoustic Features

We focused on formant central frequencies and bandwidths because evidence on the influence of sleep apnea on them has been previously reported by Rob and coworkers [13]. Formants represent resonances of the vocal tract and depend on the UA properties, including its compliance, shape, and dimensions. Hence, these may embed information from specific physiological characteristics in OSA patients, although results shall vary from one sound to another [14]. As mentioned previously, in this contribution, we focused on sustained phonations, which is the common approach for pathologic voices, and apnea may essentially be regarded as one.

Despite these elementary considerations, measuring formant frequencies can be extremely difficult as it is highly influenced by multiple factors, including the method of analysis that is chosen and the analysis settings. Moreover, higher resonances are much more difficult to determine than lower ones because of natural energy losses. Our evaluation on acoustic measurements has shown that, for formants F4 and above, no reliable information could be extracted, and therefore, we restricted our analysis to the first 3 formants. To extract a consistent set of measures on formants' central frequencies and bandwidths, we followed a specific protocol. First, we computed the values for the first 3 formant central frequencies and bandwidths using 2 different freely available software: the Praat Version 6.0.30 (Praat software, Amsterdam) [15] and the Snack Toolkit Version 2.2.8 (Snack Sound Toolkit, Sweden) [16]. Formant frequencies and bandwidths were estimated every 5 ms using 25-ms long analysis windows. Their values were finally obtained by averaging along the most stable regions of the sustained phonations of each vowel, selecting a steady-state segment of 800 ms where the standard deviation of formant contours was the lowest, excluding initial and ending silences in each utterance.

Table 1. Descriptive statistics on Spanish female and male subjects.

Clinical variables	Female (n=129)		Male (n=254)	
	Mean (SD)	Range	Mean (SD)	Range
Apnea-hypopnea index	14.6 (17.0)	0-108.4	22.3 (19.0)	0-87
Weight, in kg	78.0 (18.0)	45-165	92.4 (16.6)	61-162
Height, in cm	161.1 (6.4)	148-178	175.8 (6.9)	160-194
Body mass index, in kg/m^2	30.1 (6.9)	18.6-63.7	29.9 (4.9)	20.1-52.3
Age, in years	50.9 (11.6)	25-88	48.2 (11.9)	21-78
Cervical perimeter, in cm	36.7 (2.9)	30-45	42.5 (3.2)	34-53

To guarantee a reliable estimation, we measured the absolute differences between estimated values obtained from Praat and Snack for each formant F1-F3. We then manually reviewed those cases for which differences exceeded 70 Hz for F1 and F2 and 150 Hz for F3. These thresholds match the level of accuracy in the reference study by Robb and coworkers [13] and seem consistent with values seen in studies that compare results from Praat with those from Snack [17]. In most cases for which deviations exceeded the prespecified thresholds, one of the two values that had been computed (the one from either Snack or Praat) was found to be incorrect (most often when a formant was skipped). In these cases, the erroneously estimated value was subsequently removed, and the value provided by the other software was retained. In some other cases, both Snack and Praat failed in providing precise results. In those cases, values for formant central frequencies and bandwidths had to be manually selected using spectrograms and linear predictive coding (LPC) analysis. The decision on the number of poles for an optimal fitting of the LPC envelope was based on the general knowledge about the formant structure of each vowel. Values for formants' central frequencies were obtained as maxima values of the LPC spectral slope, whereas their associated bandwidths were computed by measuring the frequency region around formants' central frequency within which the spectral envelope amplitude differs −3 dB from the maxima values.

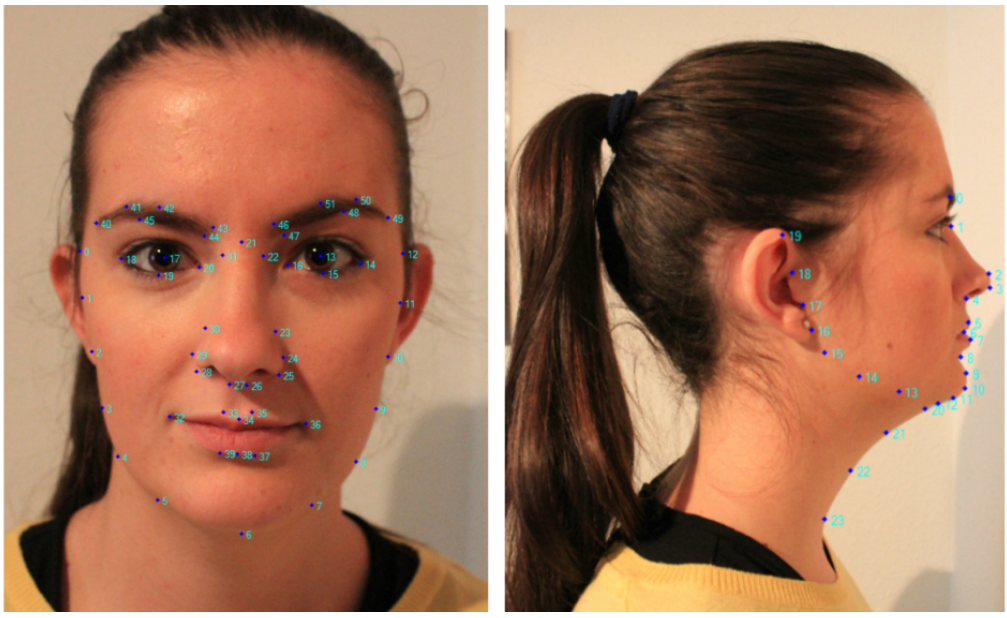
Facial Features

Facial features were similar than those studied by Lee and coworkers [4,5], including local measurements (ie, areas, distances, angles) extracted from landmarkings on photographs. Major differences in our approach when compared with that of Lee and coworkers [4,5] are the use of supervised automatic image processing and the definition of more robust craniofacial measurements adapted to our less controlled photography capture process.

Manual annotation of all images can be tedious, and, even when done by skilled personnel, it is prone to errors because of subjectivity. Consequently, we decided to use a widely used automatic landmarking method, first introduced by Cootes and coworkers [18], based on active appearance model (AAM). On the basis of a priori knowledge of landmark positions, AAM combines a statistical model, which represents the variation of shape and texture of the object, with a gradient-descent fitting algorithm. As depicted in Figure 2, in AAMs for frontal and profile photographs, we used a grid of 52 landmarks taken from a general face identification system and a set of 24 landmarks including specific marks for the neck area, respectively.

During the training stage, frontal and profile AAMs were built from a set of manually annotated photographs using the aam_tools Version 3.0 (aam_tools software, Manchester) [19].

Figure 2. Landmarks on frontal and profile views.



During the fitting stage, starting from a reasonable landmark initialization, the AAM algorithm iteratively corrects the appearance parameters by minimizing the squared error to represent the texture of the target face. Although the AAM performs well for representing shape and appearance variations of an object, the model is location-sensitive to the face's position. In this study, this effect is increased because photographs were not taken following a highly controlled procedure (illumination conditions, control of distance from the camera, and control of frontal and profile positions). Hence a human-supervised stage was found necessary in order to supervise and, if necessary, correct some large deviations in the automatically generated landmarks.

Once landmarks were generated, we proceeded to extract a set of local features, similar to those studied by Lee and coworkers [4,5] but adapted to our less controlled photographic process. These measurements are described in the following sections:

- Cervicomental contour ratio. One of the anatomical risk factors for OSA is the fat deposition on the anterior neck [20]. This risk factor is captured by a measurement proposed by Lee and coworkers [4,5], that is the cervicomental angle, which is formed by the horizontal plane of the submental region and the vertical plane of the neck. The fat deposition on the anterior neck will cause an increase of this angle. However, considering our limited photography capture process, it is extremely difficult to detect points such as cervical point, thyroid, cricoid, neck plane, or sternal notch involved in the cervicomental region. Consequently, we defined an alternative measurement, more robust to both our image capture and automatic landmarking processes. This measurement was defined using a contour in the cervicomental region traced by 6 landmarks placed equidistantly (ie, landmarks 11, 12, and 20-23 in Figure 3), which were annotated with high reliability following our semiautomatic AAM method. Therefore, the relative measurement of fat deposition on the anterior neck was calculated as the ratio of cervicomental-related area within the rectangular region (ie, yellow solid line defined by landmarks 11, 12, and 20-23, and the bottom right vertex landmark V of the rectangle as depicted in Figure 2) and the area of the rectangular region (ie, black dashed line defined by bottom left landmark 23 and upper right landmark 11 as depicted in Figure 3). This results in an uncalibrated measurement with a value that decreases as the fat deposition on the anterior neck increases.
- Face-width ratio. Lee and coworkers studied the relationship between surface facial dimensions and UA structure in subjects with OSA by means of analysis of magnetic resonance images [21]. Significant positive correlations were detected between surface facial dimensions and UA structures, in particular midface width and interocular width. On the basis of these results, we used these 2 facial dimensions to define a face-width uncalibrated measurement as the midface width to interocular width ratio. The corresponding landmarks and measurements are depicted in Figure 4.
- Tragion-ramus-stomion angle. Lowe and coworkers [6] reported that patients with OSA had retracted mandibles, which is related to the inclination of the occlusal plane and the angle between the relative position of the maxilla to mandible. On the basis of [6], we proposed an uncalibrated measure (ie, an angle) intended to capture, to some extent, the characteristic mandible position or mandibular retraction in OSA individuals. To define this angle, we selected a set of landmarks that not only are related to the posterior displacement of the mandible but also could be accurately detected by our automatic landmarking process on the photographs without need of prior marking. The proposed measurement, as depicted in Figure 5, is the angle between the line ramus-stomion (landmarks 16 and 6) and the ramus-tragion (landmarks 16 and 18).

Figure 3. Measurements used for the cervicomental contour ratio.

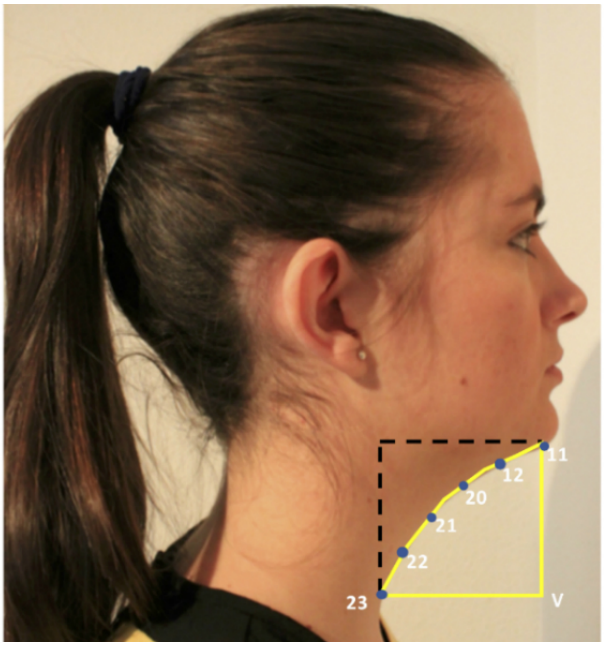


Figure 4. Measurements used for the face-width ratio.

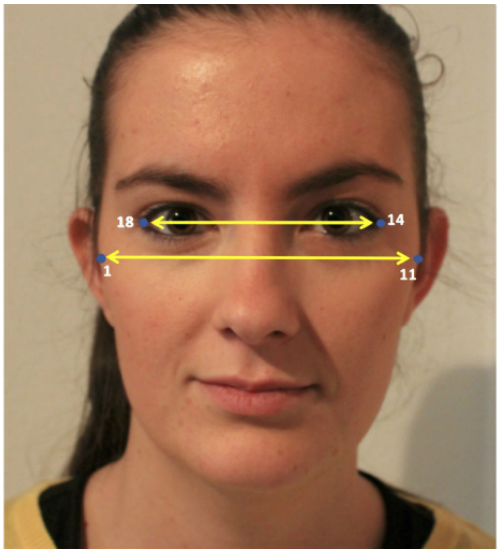
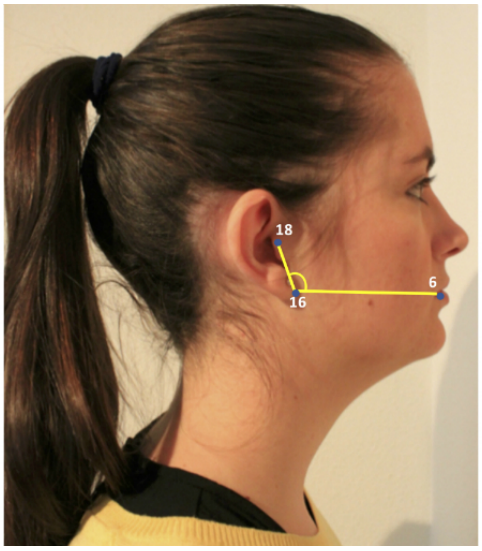


Figure 5. Tragion-ramus-stomion angle.



Statistical Analysis

To describe our results, we used the strength of the Spearman correlation coefficient as described by Fowler and coworkers [22], that is, values between .01 and .19 are regarded as very weak, .2 and .39 as weak, .40 and .69 as modest, .70 and .89 as strong correlation, and in the range of .90 to .99 as very strong. Values are reported hereafter as mean (SD) and range.

Mann-Whitney *U* test (Wilcoxon rank-sum test) was used to assess significant differences between control and OSA groups because data were not normally distributed.

We conducted our statistical analysis using the Statistic and Machine Learning Toolbox of Matlab.

Results

Due to the possible effect of clinical variables on correlation between AHI and speech and craniofacial characteristics, we first analyzed the correlation between clinical variables, speech features, craniofacial features, and AHI. Moreover, in order to observe how the connection between OSA and speech can be affected by gender, we also compared correlations between both genders.

Clinical Variables Analysis

Table 2 presents the Spearman correlation coefficient between clinical variables and AHI for both genders.

As can be seen in Table 2, the strongest correlation for female population found was between age and AHI. Correlations between cervical perimeter, BMI, and AHI are also significant but weak, as well as height, in which case the detected weak correlation is negative. In contrast, in male population, the second strongest correlation found was between weight and AHI, although weak at Fowler scale.

In a comparison by gender, the strongest correlation with AHI is different for each gender: age in the case of women (r=.52, P=.001) and cervical perimeter in the case of men (r=.42, P=.001). That is, generally, for both genders, AHI presents significant correlations with age and parameters strongly related to obesity, such as weight, BMI, and cervical perimeter, which are known as risk factors for OSA.

Table 2. Spearman correlations between clinical variables and apnea-hypopnea index (AHI) on the female (n=129) and male (n=254) population (for clarity, nonsignificant correlation values are omitted) .

Gender	Weight	Height	Age	Cervical perimeter	BMI[a]
Female		−.24[b]	.52[b]	.27[b]	.22[b]
Male	.32[b]		.15[c]	.42[b]	.37[b]

[a]BMI: body mass index.

[b]Correlation is significant at the .01 level (two-tailed).

[c]Correlation is significant at the .05 level (two-tailed).

Table 3. Descriptive statistics of the formant frequencies and bandwidth of vowels on the female population (n=129).

Vowel	Formant	Mean (SD)	Range	Bandwidth	Mean (SD)	Range
/a/	F1	859.6 (89.4)	624.4-1070.4	BW1	226.7 (78.7)	57.5-486.5
	F2	1454.0 (105.0)	1174.7-1687.7	BW2	196.6 (82.0)	76-532.6
	F3	2837.2 (221.3)	2186.9-3368.6	BW3	223.2 (97.4)	83.3-576.1
/e/	F1	489.7 (45.8)	377.8-624.3	BW1	98.8 (46.3)	22-266.7
	F2	2268.4 (141.7)	1880.9-2599.9	BW2	143.1 (55.4)	34.0-310.4
	F3	2917.7(159.3)	2597.2-3328.2	BW3	229.8 (81.1)	73.7-473.9
/i/	F1	368.1 (42.2)	243.1-481.5	BW1	68.6 (36.7)	12.2-219.6
	F2	2620.0 (150.9)	2178-2993.2	BW2	131.1 (55.3)	40.3-348.5
	F3	3170.7 (207.3)	2665.6-3645.9	BW3	236.7 (74.7)	86.8-441.4
/o/	F1	537.6 (50.8)	410.4-664.4	BW1	127.7 (61.6)	24.9-360.5
	F2	982.2 (90.5)	758.2-1243.4	BW2	141.2 (67.5)	15.4-395.1
	F3	2881.8 (215.8)	2401.8-3444.9	BW3	155.1 (67.1)	32.9-339.7
/u/	F1	379.3 (49.6)	254.6-509.6	BW1	70.7 (35.6)	8.2-199.2
	F2	823.1 (101.6)	596-1157.9	BW2	152.6 (122.0)	9.0-569.9
	F3	2824.7 (243.1)	2285.7-3790.5	BW3	213.8 (121.5)	53.4-636.2

Acoustic Features Analysis

Table 3 presents the mean, standard deviation, and the range of formant frequencies and bandwidth for vowels /a/, /e/, /i/, /o/, and /u/ for the female population.

Because of the association between the blockage of the UA and OSA, abnormal or particular speech may be expected in subjects with OSA due to the altered structure of their UA. Likewise, the association between clinical variables (ie, height and weight) and speech [12] is known; thus, indirect association might be expected between speech and OSA. Accordingly, correlations between formant frequencies, bandwidths, and clinical variables are presented in Table 4.

Focusing on formant frequencies, Table 4 shows that the highest, though weak, correlations are found with AHI, age, and cervical perimeter. Surprisingly, none of these formants are correlated with weight, BMI, or height. Moreover, results show that there are 3 formants (F1/a/, F2/e/, and F2/i/), which present weak negative correlation with AHI (r=−.26, P=.001; r=−.24, P=.01; r=−.26, P=.001; respectively). It should be noted that F2/i/ is the only formant correlated with AHI but not correlated with other clinical variables. Likewise, most of the significant correlations with formants were for age, with up to 8 formants. It is known that human voice changes with age [23], which leads us to think that age may cause indirect influence on a relationship between formant frequencies and AHI.

When considering the results for bandwidths in Table 4, only very weak correlations appear: weight negatively correlated with BW1/a/, height with BW3/o/, and age with BW2/a/ and BW2/e/, but no significant correlation was obtained between bandwidths and AHI.

To analyze the gender influence, correlation results in Table 4 were compared with those of a male population, published in our previous study [12] (Table 5). Those results include most of male subjects of the population used in this paper. Given that the difference is very small, we have preferred to use the already published tables with 241 subjects instead of publishing a slightly different one with 254 male subjects.

Table 4. Statistically significant Spearman correlation between formant frequencies, bandwidths, and clinical variables on the female (n=129) population (for clarity, nonsignificant correlation values are omitted).

Feature	Apnea-hypopnea index	Weight	Height	Age	Body mass index	Cervical perimeter
Vowel /a/						
Formant, F1	−.26[a]			−.25[a]		−.24[a]
F2				−.20[b]		
F3				−.25[a]		−.21[b]
Bandwidth, BW1		−.19[b]			−.19[b]	
BW2				−.17[b]		
Vowel /e/						
F2	−.24[a]			−.19[b]		
F3				−.22[a]		
BW2				−.21[a]		
Vowel /i/						
F2	−.26[a]					
Vowel /o/						
F2				−.20[b]		−.20[b]
F3				−.21[b]		−.17[b]
BW3			−.27[a]			
Vowel /u/						
F2				−.18[b]		

[a]Correlation is significant at the .01 level (two-tailed).

[b]Correlation is significant at the .05 level (two-tailed).

According to Table 5, contrary to the results for the female population, only bandwidths present correlation with AHI—BW2/a/ (r=.13, P=.05) and BW3/e/ (r=−.17, P=.01)—but formants do not. The overall results of speech features show that negative correlation coefficients are common between formants, bandwidths, and age. Furthermore, generally those values are smaller (weak at Fowler scale) in both genders.

This finding on the female population showed that 2 of the 3 formant frequencies correlated with AHI also have significant correlation with age (F1/a/ and F2/e/), which leads us to think that age may cause indirect influence on a relationship between formant frequencies and AHI. Similarly, in male population, BW3/e/ is also correlated with weight and BMI, which may indicate an indirect correlation with AHI.

To analyze in detail the influence of each clinical variable on correlation between speech features and OSA, a general review is provided for both genders.

First, we can see that both for male and female populations, most of the significant correlations between acoustic features and clinical variables are linked to age. This is in agreement to several studies on age-related acoustic characteristics, in which different speech features have been reported to correlate with age and have been linked to changes in anatomy and physiology of the speech production system [23]. Some specific studies have reported age-related changes to formants, particularly in the production of vowels. According to these studies, a negative correlation among formants and age, as is also found in our study, can be expected. This lowering of vowel formants with age can presumably be a by-product of the lowering of the vocal folds over the life span in an adult subject, which results in a longer vocal tract [24,25], and with a trend to vowel centralization in older subjects [17,26]. In some cases, these changes have been found to occur only on some particular vowels [25,26]. It should be noted that all the mentioned studies about this issue were performed for both genders.

Table 5. Statistically significant Spearman correlation between formant frequencies and clinical variables on the male (n=241) population (for clarity, nonsignificant correlation values are omitted).

Feature	Apnea-hypopnea index	Weight	Height	Age	Body mass index	Cervical perimeter
Vowel /a/						
Formant, F1				−.14[a]		
F2			−.13[a]	−.13[a]		
Bandwidth, BW1				−.21[b]		
BW2	.13[a]					
Vowel /e/						
F1			−.12[a]	−.12		−.17[b]
F2		−.15[a]	−.20[b]			−.16[a]
F3			−.21[b]			
BW1				−.17[b]		−.17[b]
BW2						−.16[a]
BW3	−.17[b]	−.14[a]			−.13[a]	
Vowel /i/						
F1			−.15[a]	.16[a]		
F2			−.21[b]			
F3						−.20[b]
BW1				−.13[a]		
BW3		−.15[a]	−.14[a]			
Vowel /o/						
F1						−.13[a]
F2				−.27[b]		
F3			−.17[b]			
BW1						−.13[a]
BW2		−.16[a]		.15[a]	−.14[a]	
Vowel /u/						
F1			−.14[a]			
F2				−.24[b]		
F3		−.14[a]	−.20[b]			

[a]Correlation is significant at the .05 level (two-tailed).

[b]Correlation is significant at the .01 level (two-tailed).

Considering weight and height, no significant correlations were found for female subjects. These results do not agree with those reported by González [27], where weak and modest correlations with weight and height were found: F2/e/ and height (r=−.51), and F2/e/ and weight (r=−.50). According to González [27], it seems that the most informative parameters for female height and weight were the second and the third formants from the /a/, /e/, and /i/ vowels. In the case of male population, there are no similarities with that study either. However, unlike women, there are several speech variables with significant correlations with height and weight (see Table 5). In the research by González, stronger correlations were reported for male subjects, mainly between F2/e/ and height (r=−.57) and F4/o/ and weight (r=−.48), whereas in the OSA male population in [12] the higher correlation coefficient values were obtained between F3/e/, F2/i/, and height (r=−.21, P=.001, both), and between BW1/a/ and age (r=−.21, P=.001). Likewise, in case of BMI, no significant correlation with formants was found. One may expect

formants' bandwidths to be larger for OSA patients as an increase in both velar and pharyngeal compliance could result in increased sound damping within the vocal tract [13]. However, only one significant negative correlation was detected between BMI and BW1/a/ (r=−.19, P=.03) for female patients. A similar situation was found for male patients (BW3/e/ with r=−.13, P=.05 and BW2/o/ with r=−.14, P=.03). Despite these clear differences in our studies, both point toward a similar direction: formants seem to be weak predictors of body size in both women and men. Just as in our previous discussion regarding age, it is possible to hypothesize that these significant though weak correlations with height or weight may interfere with specific acoustic characteristics related to OSA.

Finally, cervical perimeter is another feature that is commonly used in discriminating between healthy subjects and OSA patients. More specific than BMI, neck circumference can describe how excessive weight may increase tissue bulk in the neck, which will also increase the dynamic loading of the airway, thus contributing to the pathogenesis of OSA [28]. In the female OSA population under study and similar to what we have found for the other body size measurements, only few significant and weak correlations appeared between cervical perimeter and speech: with F1/a/ and F3/a/ (r=−.24, P=.01 and r=−.21, P=.02, respectively), and F2/o/ (r=−.20, P=.02). Analogous results were found for male subjects (see Table 5). Several previous studies have similarly failed to find modest relationships between voice acoustics and body size effects measured through BMI [28], body mass composition [29] or weight, and neck circumference [30].

Craniofacial Features Analysis

In this section, descriptive statistics on the female (129) and male (254) subjects under study are shown as well as correlation analysis between craniofacial features and OSA through the AHI. The craniofacial analysis comprises the 3 craniofacial measurements extracted from the landmarks, previously annotated, on patient photographs. Similar to acoustic features, differences by gender were also analyzed. Table 6 presents the mean, SD, and the range of craniofacial measurements. In Table 7, correlations between craniofacial measurements and clinical variables are presented for both genders.

In case of female population, as described in Table 7, all 3 craniofacial measurements present significant but weak correlation with AHI. Cervicomental contour ratio is also modestly correlated with BMI, weight, and cervical perimeter. As regards the face-width ratio, there is a weak negative correlation with height and positive correlation with BMI.

In case of male population, all 3 craniofacial measurements also present correlation with AHI. Furthermore, the strongest correlations are modest, negative, and correspond to BMI, cervical perimeter, and weight, both in men and women.

Furthermore, both genders report significant correlations between all 3 craniofacial measurements and AHI: positive in the case of face-width ratio (r=.18, P=.04 for women; r=.23, P=.001 for men), negative for cervicomental contour ratio (r=−.23, P=.01 for women; r=−.37, P=.001 for men), and TRG angle (r=−.19, P=.03 for women; r=−.12, P=.05 for men).

Table 6. Descriptive analysis of the craniofacial measurements on Spanish female and male subjects.

Craniofacial measurements	Female n=129		Male n=254	
	Mean (SD)	Range	Mean (SD)	Range
Cervicomental contour ratio	0.6 (0.1)	0.4-0.9	0.6 (0.09)	0.3-0.8
Face-width ratio	1.4 (0.1)	0.3-1.6	1.5 (0.06)	1.2-1.7
Tragion-ramus-stomion (TRG) angle, in degrees	113.2 (5.7)	97.7-123.7	115.5 (6.22)	98.4-130.4

Table 7. Statistically significant Spearman correlation between craniofacial measurements and clinical variables on the female (n=129) and male (n=254) population (for clarity, nonsignificant correlation values are omitted) .

Craniofacial measurements	Apnea-hypopnea index	Weight	Height	Age	Cervical perimeter	Body mass index
Female						
Cervicomental contour ratio	−.23[a]	−.65[a]			−.58[a]	−0.66[a]
Face-width ratio	0.18[b]		−.21[b]			0.22[a]
Tragion-ramus-stomion (TRG) angle, in degrees	−.19[b]			−.24[a]		
Male						
Cervicomental contour ratio	−.37[a]	−.49[a]		−.17[b]	−.57[a]	−0.59[a]
Face-width ratio	0.23[a]	0.21[a]			0.17[b]	0.25[a]
TRG angle, in degrees	−.12[b]					

[a]Correlation is significant at the .01 level (two-tailed).

[b]Correlation is significant at the .05 level (two-tailed).

In general, male population presents stronger values. Indeed, cervicomental contour ratio has the strongest correlation with AHI in both groups. However, as it was pointed out before, this craniofacial measurement also has modest correlation with BMI, weight, and cervical perimeter. Hence, an underlying connection between AHI and cervicomental contour ratio through these clinical variables may exist.

Similar to what was considered for the acoustic feature analysis, we now analyze the influence of each clinical variable on the correlation between craniofacial features and OSA for both genders.

In the case of age, despite the changes in the facial skeleton that occur with aging, only one significant negative correlation between age and craniofacial measurements was found for both men and women: a very weak correlation with cervicomental contour ratio (r=−.17, P=.01) in the case of male patients and a weak one with TRG angle (r=−.24, P=.001) in the case of female patients.

As for height, there is only one significant weak correlation with face-width ratio (r=−.21, P=.02) in female subjects, and no significant correlation was found in male subjects. Regarding this item, there are some controversial conclusions within the scientific community; some of the researches reported strong relationship between craniofacial parameters and stature [31], whereas some of them have not [32,33].

Considering now BMI, weight, and cervical perimeter, in female subjects (Table 7) the more relevant correlations correspond to BMI (r=−.66, P=.001), weight (r=−.65, P=.001), and cervical perimeter (r=−.58, P=.001) with cervicomental contour ratio. In male subjects, higher significant correlations are also related to cervicomental contour ratio with the same clinical parameters (BMI: r=−.59, P=.001; weight: r=−.49, P=.001; and cervical perimeter: r=−.59, P=.001). These results point to cervicomental contour ratio related to the neck and under-the-chin fat depositions as the most likely of facial measurements to be a possible risk factor for OSA.

Statistical Contrasts Among OSA Severity Groups

In the previous sections, we have studied the correlation between the full AHI range and the set of speech/craniofacial features. In this section, we analyze whether or not these features can be discriminative between two female populations: a control group, defined for AHI<10 and an OSA group for AHI≥10. Similar analyses for male populations were presented by Robb and coworkers in [13] and by ourselves in [12].

Statistical contrasts using Mann-Whitney U test among control and OSA groups are presented in Table 8. Looking at the results in Table 8, it can be seen that most of the discriminative speech features reported by Robb are not detected. Only a significant difference in F2/i/ is present, whereas a few novel differences arise for F2/e/, F3 /i/, and BW2/e/.

Table 8. Contrast among control and obstructive sleep apnea (OSA) severity groups on the female population (N=129).

Feature	Control (apnea-hypopnea index, AHI <10), n=65	Obstructive sleep apnea (AHI≥10), n=64	P value
	Mean (SD)	Mean (SD)	
Clinical variables			
AHI	4.0 (3.1)	25.4 (18.5)	.001[a]
Weight	76.4 (18.8)	79.7 (17.2)	.31
Height	162.3 (6.4)	150.9 (6.2)	.04[a]
Age	45.4 (10.4)	56.5 (10.0)	.001[a]
Body mass index	29.0 (7.1)	31.2 (6.6)	.04[a]
Cervical perimeter	36.0 (2.8)	37.5 (2.9)	.01[a]
Speech features			
Formant, F2/e/	2303.5 (129.2)	2232.7 (145.8)	.01[a]
F2/i/	2664.0 (150.7)	2575.3 (138.5)	.001[a]
F3/i/	3210.4 (203.5)	3130 (204.9)	.03[a]
Bandwidth, BW2/e/	152.3 (52.5)	133.9 (57.0)	.02[a]
Craniofacial features			
Cervicomental contour ratio	0.6 (0.1)	0.6 (0.1)	.07
Face-width ratio	1.4 (0.05)	1.4 (0.06)	.07
Tragion-ramus-stomion (TRG) angle	114.2 (5.0)	112.0 (6.2)	.06

[a]There are significant differences between OSA groups at the .05 level.

We have reported results for a similar contrast among control and OSA groups for men in [12] (Table 9). This allows us to compare results for female and male populations (Tables 8 and 9, respectively) and see that only F3/i/ appears in both populations. It is also interesting to notice that for males, the remainder significant differences appear only in bandwidths BW1/o/, BW3/o/, BW2/a/, and BW3/e/.

If we analyze now the statistical differences among control and OSA groups for the clinical variables (also shown in Tables 8 and 9), we can see that only weight in females and height in males present no statistical differences. Consequently, it must be concluded the presence of indirect influences of speech and AHI mediated through the rest of clinical variables.

A similar statistical contrast between control and OSA groups was made for craniofacial features. Results showed no significant differences between groups for the female population (see Table 8). Results for our male population are presented in Table 10. These results show significant statistical differences in cervicomental contour ratio and face-width ratio. This points out that the studied facial measurements are more suitable for estimating the AHI in male subjects.

Matched Groups

As discussed before, our results indicate that there can be an indirect relationship between AHI and both speech and craniofacial features mediated through the clinical variables (age, weight, height, BMI, and cervical perimeter). To evaluate this indirect effect, statistical contrasts are again presented for control and OSA groups but now selected to exhibit no statistical differences among the clinical variables. Thus, the objective was to test whether or not statistical differences previously observed in Tables 8 and 9 (with unmatched values in clinical variables) remain in a matched condition (ie, when there are no statistical differences in clinical variables among control and OSA groups).

Results on matched groups for female population are presented in Table 11, which correspond to control and OSA groups including subjects in the age range of 41 to 55 years and BMI≥25 so that no statistical differences in clinical variables appear.

Table 9. Contrast among control and obstructive sleep apnea (OSA) severity groups on a male (N=241) population.

Feature	Control (apnea-hypopnea index, AHI<10), n=81	Mild obstructive sleep apnea, OSA (AHI 10-30), n=87		Severe OSA (AHI>30), n=73	
	Median	Median	*P* value	Median	*P* value
Clinical variables					
Weight	86.0	90.0	.24	99.0	.001[a]
Body mass index	27.3	28.9	.01[a]	31.4	.001[a]
Age	42.0	51.0	.001[a]	48.0	.02[a]
Cervical perimeter	41.0	42.0	.001[a]	44.0	.001[a]
Speech features					
Formant, F3/i/	2707.0	2682.0	.05[a]	2642.0	.03[a]
Bandwidth, BW1/o/	94.0	79.0	.001[a]	85.0	.10
BW3/o/	98.0	136.0	.01[a]	107.0	.10
BW2/a/	118.0	125.0	.08	148.0	.01[a]
BW3/e/	170.0	201.0	.10	140.0	.03[a]

[a]There are significant differences between OSA groups at the .05 level.

Table 10. Contrast among control and obstructive sleep apnea (OSA) groups on the male (N=254) population.

Craniofacial measurements	Control (apnea-hypopnea index, AHI<10), n=86	Obstructive sleep apnea, OSA (AHI≥10), n=168	*P* value
	Mean (SD)	Mean (SD)	
Cervicomental contour ratio	0.6 (0.1)	0.5 (0.1)	.00[a]
Face-width ratio	1.4 (0.1)	1.5 (0.1)	.00[a]
Tragion-ramus-stomion (TRG) angle	116 (6.5)	115.2 (6.1)	.33

[a]There are significant differences between OSA groups at the .05 level.

Table 11. Contrast between control and obstructive sleep apnea groups on a subset without differences either on age (41-55 years) or on body mass index (≥25) on the female population.

Feature	Control (apnea-hypopnea index, AHI<10), n=22	Obstructive sleep apnea, OSA (AHI≥10), n=19	*P* value
	Mean (SD)	Mean (SD)	
Clinical variables			
AHI	4.8 (3.0)	26.9 (20.5)	.001[a]
Weight	86.4 (22.0)	89 (16.4)	.67
Height	161.2 (6.3)	161.3 (5.1)	.94
Age	48.9 (4.1)	50.2 (4.2)	.35
Body mass index	33.3 (8.4)	34.2 (5.9)	.68
Cervical perimeter	37.1 (2.2)	38 (2.1)	.18
Speech features			
Formant, F2/i/	2670.5 (139.3)	2571.4 (145.8)	.05
Craniofacial features			
Cervicomental contour ratio	0.6 (0.1)	0.6 (0.1)	.67
Face-width ratio	1.4 (0.04)	1.4 (0.050	.96
Tragion-ramus-stomion (TRG) angle	114.1 (5.0)	113 (5.1)	.71

[a]There are significant differences between OSA groups at the .05 level.

Table 12. Contrast between control and obstructive sleep apnea (OSA) groups on a subset without differences either on age (≤46) or on body mass index (≤30) on the male population.

Feature	Control (apnea-hypopnea index, AHI<10), n=29	Obstructive sleep apnea, OSA (AHI≥10), n=23	*P* value
	Mean (SD)	Mean (SD)	
Clinical variables			
AHI	5.3 (3.0)	27.0 (12.8)	.001[a]
Weight	87.1 (5.3)	87.1 (8.1)	.99
Height	178.9 (5.0)	177.6 (5.6)	.35
Age	38.1 (4.3)	39.0 (5.2)	.47
Body mass index	27.2 (1.3)	27.5 (1.5)	.35
Cervical perimeter	40.3 (1.3)	40.7 (1.7)	.38
Craniofacial features			
Cervicomental contour ratio	0.6 (0.07)	0.6 (0.1)	.02[a]
Face-width ratio	1.4 (0.06)	1.5 (0.05)	.23
Tragion-ramus-stomion (TRG) angle	116.2 (6.8)	115.7 (5.6)	.73

[a]There are significant differences between OSA groups at the .05 level.

As Table 11 illustrates, once the possible effect of age and BMI is minimized, only significant difference in F2/i/ remains, whereas the differences for F2/e/, F3/i/, and BW2/e/ disappear. This result is coherent with correlations in Table 4, where it can be noted that F2/i/ is only correlated with AHI, whereas correlation with some clinical variables appear for F2/e/ and BW2/e/. Also by comparing Table 8 with Table 11, it can be observed that, as it is reasonable, there are no significant differences in craniofacial measurements in both tables.

Matched results for the male population selecting individuals with age ≤46 and BMI in the range of 25 to 30 are presented in Table 12. Results in this table indicate that only the significant difference in cervicomental contour ratio remains, which indicates that the neck fat deposition is a possible risk factor for OSA in male population, as it was pointed before because higher significant correlation was related to this craniofacial feature (see Table 7).

Discussion

Principal Findings

The results of this investigation indicate that acoustic and facial measurements in a female population have weaker correlation with AHI than with clinical variables. Significant correlations for female individuals (mainly weak correlations) are somewhat stronger than those for male subjects (mainly very weak).

In the studied female population, formant frequencies seem to prevail over bandwidths. Specifically, F2/i/ is the speech variable that showed to be a good predictor of OSA syndrome, as it is the only acoustic measurement that remains after contrasting OSA and non-OSA individuals, both unmatched (Table 8) and matched (Table 11), with clinical variables. Regarding craniofacial parameters, according to the results, the particular facial features that we have studied are not suitable to distinguish between OSA and non-OSA female subjects.

In the case of male population, bandwidths seem to prevail in their correlation to AHI over formant frequencies. BW2/a/ and BW3/e/ are the only ones that remain after the same contrast analysis using groups matched in clinical variables (Table 13). Considering craniofacial measurements, cervicomental contour ratio is the variable that is still present after the contrast analysis using matched groups (see Table 12). This outcome suggests that the use of craniofacial measurements is more appropriate to differentiate OSA-affected male patients.

Limitations

We are aware that our research has several limitations. The first one is that the results presented in this study are limited to Spanish subjects, most of whom are speakers of a single Spanish dialect, the Andalusian. Consequently, a cross-language comparison should be made. Another limitation is that measuring formant frequencies and bandwidths is technically problematic, and it always achieves limited precision. To obtain results with higher accuracy and reliability, future studies will need to examine possible impact of different factors, such as patient's position or time of the day, during the data collection process, as acoustic differences may be expected.

With regard to craniofacial features, we have only explored uncalibrated craniofacial measurements because we have limited our study to simulate an OSA assessment app running on a mobile device. Our research may also be limited by the precision of the measurements, particularly in the case of the craniofacial measurements.

Conclusions

An important outcome of our investigation is that there may be a possible underlying impact of clinical variables on the correlations between voice features and OSA. Thus, future research should consider new speech analysis techniques capable of properly compensating unwanted variability due to clinical variables. In the case of craniofacial measurements, the results suggest that the features used in this study are more suitable for male patients than for female patients. Therefore, searching for those specific features that are more convenient for female subjects would be interesting to try to improve the assessment techniques of OSA in women.

Moreover, besides the known OSA risk factors, there are other disorders that can cause OSA, such as hypothyroidism [34] and acromegaly [35] disorders, which can give different craniofacial representation. Comparing these features in different groups could skew the data. Therefore, future studies should also contemplate these related OSA conditions as exclusion criteria to avoid false discoveries.

Table 13. Contrast between control and obstructive sleep apnea (OSA) groups on a subset without differences either on age or on height on the male population.

Features	Control (apnea-hypopnea index, AHI<10), n=73	Obstructive sleep apnea, OSA (AHI>30), n=65	*P* value
	Median (SD)	Median (SD)	
Clinical variables			
Weight	85 (11.1)	99 (22.65)	.001[a]
Height	174 (7.26)	176 (6.61)	.24
Body mass index	27 (3.79)	32 (6.63)	.001
Age	43 (9.51)	47 (8.55)	.21
Cervical perimeter	41 (2.46)	44 (3.7)	.001[a]
Speech features			
Bandwidth, BW2/a/	116 (40.31)	148 (129.4)	.001[a]
BW3/e/	161 (98.23)	137 (94.6)	.03[a]

[a]There are significant differences between OSA groups at the .05 level.

Acknowledgments

The activities presented in this paper were funded by the Spanish Ministry of Economy and Competitiveness and the European Union (FEDER) as part of the TEC2015-68172-C2-2-P (Deep & Subspace Speech Learning, DSSL) project. The authors also thank Sonia Martinez Díaz for her effort in collecting the OSA database that is used in this study.

Conflicts of Interest

None declared.

References

1. Lam JC, Sharma SK, Lam B. Obstructive sleep apnoea: definitions, epidemiology & natural history. Indian J Med Res 2010 Feb;131:165-170 [FREE Full text] [Medline: 20308741]
2. [No authors listed]. AARC-APT (American Association of Respiratory Care-Association of Polysomnography Technologists) clinical practice guideline. Polysomnography. Respir Care 1995 Dec;40(12):1336-1343. [Medline: 10153260]
3. Schwab RJ, Pasirstein M, Pierson R, Mackley A, Hachadoorian R, Arens R, et al. Identification of upper airway anatomic risk factors for obstructive sleep apnea with volumetric magnetic resonance imaging. Am J Respir Crit Care Med 2003 Sep 1;168(5):522-530. [doi: 10.1164/rccm.200208-866OC] [Medline: 12746251]
4. Lee RW, Chan AS, Grunstein RR, Cistulli PA. Craniofacial phenotyping in obstructive sleep apnea--a novel quantitative photographic approach. Sleep 2009 Jan;32(1):37-45 [FREE Full text] [Medline: 19189777]
5. Lee RW, Petocz P, Prvan T, Chan AS, Grunstein RR, Cistulli PA. Prediction of obstructive sleep apnea with craniofacial photographic analysis. Sleep 2009 Jan;32(1):46-52 [FREE Full text] [Medline: 19189778]
6. Lowe AA, Fleetham JA, Adachi S, Ryan CF. Cephalometric and computed tomographic predictors of obstructive sleep apnea severity. Am J Orthod Dentofacial Orthop 1995 Jun;107(6):589-595. [Medline: 7771363]
7. Fox AW, Monoson PK, Morgan CD. Speech dysfunction of obstructive sleep apnea. A discriminant analysis of its descriptors. Chest 1989 Sep;96(3):589-595. [Medline: 2766817]
8. Monoson PK, Fox AW. Preliminary observation of speech disorder in obstructive and mixed sleep apnea. Chest 1987 Oct;92(4):670-675. [Medline: 3308347]
9. Espinoza-Cuadros F, Fernández-Pozo R, Toledano DT, Alcázar-Ramírez JD, López-Gonzalo E, Hernández-Gómez LA. Speech signal and facial image processing for obstructive sleep apnea assessment. Comput Math Methods Med 2015;2015:489761 [FREE Full text] [doi: 10.1155/2015/489761] [Medline: 26664493]
10. Goldshtein E, Tarasiuk A, Zigel Y. Automatic detection of obstructive sleep apnea using speech signals. IEEE Trans Biomed Eng 2011 May;58(5):1373-1382. [doi: 10.1109/TBME.2010.2100096] [Medline: 21172747]
11. Solé-Casals J, Munteanu C, Martín OC, Barbé F, Queipo C, Amilibia J, et al. Detection of severe obstructive sleep apnea through voice analysis. Appl Soft Comput 2014;23:346-354. [doi: 10.1016/j.asoc.2014.06.017]
12. Montero-Benavides A, Blanco-Murillo JL, Fernández-Pozo R, Espinoza-Cuadros F, Torre-Toledano D, Alcázar-Ramírez JD, et al. Formant frequencies and bandwidths in relation to clinical variables in an obstructive sleep apnea population. J Voice 2016 Jan;30(1):21-29. [doi: 10.1016/j.jvoice.2015.01.006] [Medline: 25795368]
13. Robb MP, Yates J, Morgan EJ. Vocal tract resonance characteristics of adults with obstructive sleep apnea. Acta Otolaryngol 1997 Sep;117(5):760-763. [Medline: 9349877]
14. Blanco-Murillo JL, Fernández-Pozo R, López-Gonzálo E, Hernández-Gómez LA. Exploring differences between phonetic classes in sleep apnoea syndrome patients using automatic speech processing techniques. The Phonetician 2008;97(1):35-54 [FREE Full text]
15. Boersma P, Weenink D. Praat. Amsterdam: University of Amsterdam Praat: doing phonetics by computer URL: http://www.praat.org/ [accessed 2017-10-26] [WebCite Cache ID 6uUtZAcbR]
16. Sjölander K. Speech.kth.se. Sweden: KTH Royal Institute of Technology The snack sound toolkit URL: http://www.speech.kth.se/snack/index.html [accessed 2017-10-26] [WebCite Cache ID 6uUtWte6S]
17. Liss JM, Weismer G, Rosenbek JC. Selected acoustic characteristics of speech production in very old males. J Gerontol 1990 Mar;45(2):P35-P45. [Medline: 2313046]
18. Cootes TF, Edwards GJ, Taylor CJ. Active appearance models. IEEE Trans Pattern Anal Mach Intell 2001 Jun;23(6):681-685. [doi: 10.1109/34.927467]
19. Cootes T. AAM tools. Manchester: The University of Manchester URL: http://personalpages.manchester.ac.uk/staff/timothy.f.cootes/software/aam_tools_doc/index.html [accessed 2017-10-26] [WebCite Cache ID 6uUtUDGvc]
20. Mortimore IL, Marshall I, Wraith PK, Sellar RJ, Douglas NJ. Neck and total body fat deposition in nonobese and obese patients with sleep apnea compared with that in control subjects. Am J Respir Crit Care Med 1998 Jan;157(1):280-283. [doi: 10.1164/ajrccm.157.1.9703018] [Medline: 9445310]
21. Lee RW, Sutherland K, Chan AS, Zeng B, Grunstein RR, Darendeliler MA, et al. Relationship between surface facial dimensions and upper airway structures in obstructive sleep apnea. Sleep 2010 Sep;33(9):1249-1254 [FREE Full text] [Medline: 20857873]
22. Fowler J, Cohen L, Jarvis P. Practical Statistics for Field Biology. New York: John Wiley & Sons; 1998.

23. Torre 3rd P, Barlow JA. Age-related changes in acoustic characteristics of adult speech. J Commun Disord 2009;42(5):324-333. [doi: 10.1016/j.jcomdis.2009.03.001] [Medline: 19394957]
24. Benjamin BJ. Speech production of normally aging adults. Semin Speech Lang 1997 May;18(2):135-141. [doi: 10.1055/s-2008-1064068] [Medline: 9195686]
25. Xue SA, Hao GJ. Changes in the human vocal tract due to aging and the acoustic correlates of speech production: a pilot study. J Speech Lang Hear Res 2003 Jun;46(3):689-701. [Medline: 14696995]
26. Rastatter MP, McGuire RA, Kalinowski J, Stuart A. Formant frequency characteristics of elderly speakers in contextual speech. Folia Phoniatr Logop 1997;49(1):1-8. [Medline: 9097490]
27. González J. Formant frequencies and body size of speaker: a weak relationship in adult humans. J Phon 2004 Apr;32(2):277-287. [doi: 10.1016/S0095-4470(03)00049-4]
28. Lee BJ, Ku B, Jang JS, Kim JY. A novel method for classifying body mass index on the basis of speech signals for future clinical applications: a pilot study. Evid Based Complement Alternat Med 2013;2013:150265 [FREE Full text] [doi: 10.1155/2013/150265] [Medline: 23573116]
29. Hamdan AL, Al Barazi R, Khneizer G, Turfe Z, Sinno S, Ashkar J, et al. Formant frequency in relation to body mass composition. J Voice 2013 Sep;27(5):567-571. [doi: 10.1016/j.jvoice.2012.09.005] [Medline: 23583206]
30. Rendall D, Kollias S, Ney C, Lloyd P. Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. J Acoust Soc Am 2005 Feb;117(2):944-955. [Medline: 15759713]
31. Anibor E, Eboh DE, Etetafia MO. A study of craniofacial parameters and total body height. Pelagia Res Libr Adv Appl Sci Res 2011;2(6):400-405 [FREE Full text]
32. Pelin C, Zağyapan R, Yazici C, Kürkçüoğlu A. Body height estimation from head and face dimensions: a different method. J Forensic Sci 2010 Sep;55(5):1326-1330. [doi: 10.1111/j.1556-4029.2010.01429.x] [Medline: 20487169]
33. Rexhepi AM, Brestovci B. Prediction of stature according to three head measurements. Int J Morphol 2015 Sep;33(3):1151-1155. [doi: 10.4067/S0717-95022015000300055]
34. Rajagopal KR, Abbrecht PH, Derderian SS, Pickett C, Hofeldt F, Tellis CJ, et al. Obstructive sleep apnea in hypothyroidism. Ann Intern Med 1984 Oct 1;101(4):491-494. [doi: 10.7326/0003-4819-101-4-491]
35. Mickelson SA, Senior BA, Rosenthal LD, Rock JP, Friduss ME. Obstructive sleep apnea syndrome and acromegaly. Otolaryngol Head Neck Surg 1994;111(1):25-30. [doi: 10.1177/019459989411100107]

Abbreviations

AAM: active appearance model
AHI: apnea-hypopnea index
BMI: body mass index
BW: bandwidth
F: formant
LPC: linear predictive coding
OSA: obstructive sleep apnea
PSG: polysomnography
TRG: tragion-ramus-stomion
UA: upper airway

Edited by C Dias; submitted 19.06.17; peer-reviewed by C Rabec, H Mishra, N Ocal; comments to author 25.07.17; revised version received 14.09.17; accepted 21.09.17; published 06.11.17.

Please cite as:
Tyan M, Espinoza-Cuadros F, Fernández Pozo R, Toledano D, Lopez Gonzalo E, Alcazar Ramirez JD, Hernandez Gomez LA
Obstructive Sleep Apnea in Women: Study of Speech and Craniofacial Characteristics
JMIR Mhealth Uhealth 2017;5(11):e169
URL: http://mhealth.jmir.org/2017/11/e169/
doi:10.2196/mhealth.8238
PMID:29109068

Original Paper

Uses of Mobile Device Digital Photography of Dermatologic Conditions in Primary Care

Jennifer L Pecina[1], MD; Kirk D Wyatt[2], MD; Nneka I Comfere[3], MD; Matthew E Bernard[1], MD; Frederick North[4], MD

[1]Department of Family Medicine, Mayo Clinic, Rochester, MN, United States

[2]Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, MN, United States

[3]Departments of Dermatology and Laboratory Medicine & Pathology, Mayo Clinic, Rochester, MN, United States

[4]Department of Primary Care Internal Medicine, Mayo Clinic, Rochester, MN, United States

Corresponding Author:
Jennifer L Pecina, MD
Department of Family Medicine, Mayo Clinic
200 1st Street SW
Rochester, MN
United States
Phone: 1 5072840423
Email: pecina.jennifer@mayo.edu

Abstract

Background: PhotoExam is a mobile app that incorporates digital photographs into the electronic health record (EHR) using iPhone operating system (iOS, Apple Inc)–based mobile devices.

Objective: The aim of this study was to describe usage patterns of PhotoExam in primary care and to assess clinician-level factors that influence the use of the PhotoExam app for teledermatology (TD) purposes.

Methods: Retrospective record review of primary care patients who had one or more photos taken with the PhotoExam app between February 16, 2015 to February 29, 2016 were reviewed for 30-day outcomes for rates of dermatology consult request, mode of dermatology consultation (curbside phone consult, eConsult, and in-person consult), specialty and training level of clinician using the app, performance of skin biopsy, and final pathological diagnosis (benign vs malignant).

Results: During the study period, there were 1139 photo sessions on 1059 unique patients. Of the 1139 sessions, 395 (34.68%) sessions documented dermatologist input in the EHR via dermatology curbside consultation, eConsult, and in-person dermatology consult. Clinicians utilized curbside phone consults preferentially over eConsults for TD. By clinician type, nurse practitioners (NPs) and physician assistants (PAs) were more likely to utilize the PhotoExam for TD as compared with physicians. By specialty type, pediatric clinicians were more likely to utilize the PhotoExam for TD as compared with family medicine and internal medicine clinicians. A total of 108 (9.5%) photo sessions had a biopsy performed of the photographed site. Of these, 46 biopsies (42.6%) were performed by a primary care clinician, and 27 (25.0%) biopsies were interpreted as a malignancy. Of the 27 biopsies that revealed malignant findings, 6 (22%) had a TD consultation before biopsy, and 10 (37%) of these biopsies were obtained by primary care clinicians.

Conclusions: Clinicians primarily used the PhotoExam for non-TD purposes. Nurse practitioners and PAs utilized the app for TD purposes more than physicians. Primary care clinicians requested curbside dermatology consults more frequently than dermatology eConsults.

(JMIR Mhealth Uhealth 2017;5(11):e165) doi:10.2196/mhealth.8257

KEYWORDS

telemedicine; teledermatology; mobile phone; mobile applications; primary health care; smartphone; remote consultation

Introduction

The visual nature of skin conditions in dermatology has promoted the application of medical photography in the specialty. Technological advances in digital cameras, computer memory, and processing power have contributed to an expansion of potential applications of digital medical photography in dermatology practice. Digital photography facilitates

communication with the health care team and patients and documentation for medical and academic purposes [1]. Multiple applications of digital photography in dermatology have emerged, including education, clinical archiving in patient records, surgical documentation pre- and postoperatively, follow-up of chronic conditions for progression or treatment response, diagnosis of melanocytic lesions, and teledermatology (TD) consultation [1].

TD is an established practice that enables remote access to dermatologic care using communications technology [2]. Three modes of TD care delivery are utilized: (1) store and forward (S&F) for transmitting digital images and clinical information to the dermatologist for consultation at a later time; (2) real-time video teleconferencing (VTC), where clinicians and patient interact live in videoconference; and (3) a hybrid model using a combination of the above. Store and forward represents the most commonly employed care delivery model for TD and may substitute for the dermatologic physical examination [3].

Uses of TD include consultation, triage, direct care for the diagnosis and management of skin disorders, and follow-up of chronic skin conditions [4]. Studies suggest that TD is cost-effective and associated with high levels of diagnostic accuracy, increased access to care, better clinical outcomes, and high levels of satisfaction reported by patients, referring clinicians, and dermatologists using this model [5-8]. TD via mobile phone apps has been targeted directly to consumers, as well as to clinicians [9]. Nelson et al demonstrated increased speed of dermatologic consultation and accessibility using S&F TD via a mobile app in an underserved primary care setting [10]. Nami et al found high levels of agreement between dermatologist in-person diagnoses and teledermatologist diagnoses when using asynchronous iPhone 4S (Apple Inc, Cupertino, CA, USA), skin images and clinical history from a mobile TD app (MugDerma e-derm-consult GmbH, Graz, Austria), [11]. A pilot study of resident physicians from emergency medicine, internal medicine, and dermatology that evaluated physician satisfaction with a mobile app for clinical photography found that the majority found the app useful and easy to use and desired to continue using the app [12]. Increasing access to mobile technologies with high quality cameras with or without electronic health record (EHR) integration, the rising burden of skin disease, uneven geographic distribution of dermatologists, prolonged wait times, and socioeconomic barriers to dermatologic care have together fueled the rise of TD consultative services across the United States [13]. Although privacy and confidentiality, image quality, and diagnostic confidence continue as barriers to widespread adoption [14-19], TD has great potential to minimize disparities in dermatologic care.

The purpose of this retrospective study was two-fold: (1) to describe usage patterns of a mobile iPhone operating system (iOS)–based app, PhotoExam, for clinical image capture of skin conditions in the primary care clinics of an integrated tertiary care system and (2) to assess clinician-level factors that influence the use of the PhotoExam app for TD purposes. Our primary outcome measures were 30-day post-S&F dermatology consultation request and mode of dermatology consultation (curbside consult, eConsult, and in-person consult). Our secondary outcome measures were specialty and training level of clinician using the app, performance of biopsy of skin condition of interest, and final pathological diagnosis (benign vs malignant).

Methods

Store and Forward Teledermatology Platform and Care Process

In 2015, Mayo Clinic launched PhotoExam, an internal app available to all clinicians that incorporates digital clinical photographs into the EHR using iOS-based mobile devices. In primary care practice, the promise with the app's release was greater access to dermatologic expertise, with the ability to either speak with an on-call dermatologist via telephone for curbside consultation or submit a dermatology eConsult. Curbside consultations occur via telephone in a live interactive consultation between the primary care clinician and the dermatologist either with or without the patient present. For a curbside consultation, the referring clinician calls a dedicated pager that is carried by a dermatologist who is on call to respond to curbside consultations. During a curbside consultation, the clinical documentation of the discussion and recommendations are performed by the requesting primary care clinician. An eConsult is an asynchronous electronic free text-based consultation between the requesting clinician and the dermatologist, where medical information including the PhotoExam images, diagnostic tests, and specific questions are made available electronically to the dermatology specialist to enable timely access to dermatologic expertise. eConsults are ordered by the referring provider, and completion is tracked using the same electronic ordering system as in-person referrals. eConsults result in a formal dermatology note in the patient's EHR.

PhotoExam is an app that can be downloaded by providers from an internal Mayo Clinic website to iPhone, iPad, or iPod touch devices when connected to the Mayo Clinic intranet. Once the clinician securely logs in to the patient record using the PhotoExam app, they are prompted to confirm that consent has been obtained for photography before accessing the photo capture screen. The provider is then prompted to enter the anatomical sites being photographed. Photos can then be taken in the app using the device's camera. Clinical images are transferred to institutional imaging systems via a secure network, stored, and made available for viewing in the EHR within minutes. PhotoExam is available to be run on any iOS 7.2 or higher device that is connected to the institutional network. As the devices are limited to this platform, photo resolution is assured to at least equal the criteria in current TD practice guidelines (8 megapixel or greater) [6]. Because no patient data remains on the mobile device, clinicians can use their own personally owned devices to take photos; however, all devices using the app must have a security profile installed. No training in medical photography is provided or required before using the app. A previous review of system-wide use of the app (ie, not limited to the primary care setting or specifically to TD) revealed that the quality of photos taken with the app was judged as generally good, with images receiving, on average, 91% of

possible points on the quality scoring rubric used for assessment [20].

Study Design and Data Collection

We performed a retrospective review of the EHR for patients who had clinical digital images captured by primary care clinicians in family medicine, internal medicine, and general pediatrics using the PhotoExam app from February 16, 2015 to February 29, 2016. Patients who did not have research authorization on file were excluded from review. This study was approved by the institutional review board of the Mayo Clinic. The Mayo Clinic Advanced Cohort Explorer is a clinical data repository with text search functionality that was used to text search the EHR for the keywords "dermatology," "dermatologic," "dermatologist," and "derm" and to search for pathology reports within 30 days after photo capture using the PhotoExam app. Medical records that contained the above search terms were then manually reviewed to assess for our primary outcomes. PhotoExam sessions represented our unit of study and were documented for each patient. A session was defined as all photos taken by a clinician of a single patient within a single calendar day. All data collection was performed by one coauthor (JLP).

Patient records were reviewed for the following outcomes within 30 days after photo capture: dermatology consultation, mode of dermatology consultation (curbside consult, eConsult, and/or in-person consult), performance of skin biopsy of the dermatologic condition captured with the PhotoExam app, medical specialty and training level of clinician using the app, and final pathologic diagnosis (malignant vs benign) of skin biopsy, if performed. For eConsults, we looked for statements from the responding dermatologist about the quality of the photographic images, noting particularly if the dermatologist made any comment on inadequacy of images for any reason. Descriptive statistics were used to summarize the data. Statistical analyses were performed using software, JMP Pro 12.0.0 (SAS Institute Inc). Excel 2010 (Microsoft) was used to generate a bubble chart of user types and specialties.

Results

During the study period, 1139 discrete PhotoExam sessions were captured on 1059 unique patients. Table 1 outlines the characteristics of the cohort.

Table 1. Features of PhotoExam sessions.

Features	n (% or range), N=1059 patients (1139 photo sessions)
Age (years), mean	44 (0-104)
Gender	
Female	564 (53.26%)
Race	
White	917 (86.59%)
Black	40 (3.78%)
Asian	44 (4.15%)
Other	36 (3.40%)
Chose not to disclose	12 (1.13%)
American Indian or Alaskan native	4 (<1%)
Native Hawaiian or Pacific Islander	1 (<1%)
Unknown	5 (<1%)
Photos per session, median	2 (1-18)
Number of body sites photographed per session, median	1 (1-6)
Photo sessions by training level of clinician	
Consultant physician	567 (49.78%)
Resident physician	155 (13.60%)
NP[a] and PA[b]	417 (36.61%)
Photo sessions by medical specialty	
Internal medicine	589 (51.71%)
Family medicine	282 (24.76%)
Pediatrics	268 (23.53%)

[a]NP: nurse practitioner.

[b]PA: physician assistant.

Of the 1139 sessions, 395 (34.68%) sessions documented dermatologist input in the EHR via dermatology curbside consultation, eConsult, and in-person dermatology consult. Figure 1 demonstrates a flowchart showing patterns of dermatology consultation after use of the app by the requesting primary care clinician.

The likelihood of requesting dermatology input on photos taken with the app varied significantly by both requesting clinician medical specialty and training level with NPs and PAs utilizing a higher proportion of photo sessions for TD consultation as compared with consultant and resident physicians (P=.03). Across the medical specialties in our cohort, pediatric clinicians utilized a higher percentage of photo sessions for dermatology consultation purposes as compared with family medicine and internal medicine clinicians (P=.003; Figure 2).

A total of 108 (9.5%) photo sessions had a biopsy performed of the photographed site, and only one biopsy was performed per session. Thus, 108 skin biopsies were performed; of these 46 biopsies (42.6%) were performed by a primary care clinician. Of the 108 photo sessions that had a biopsy of the site within 30 days of photo capture, 27 (25.0%) biopsies were interpreted as a malignancy (14 squamous cell carcinomas, 7 basal cell carcinomas, 3 melanomas, and 3 lymphomas). Of the 27 biopsies that revealed malignant findings, 6 (22%) had an eConsult or curbside consultation before biopsy, and 10 (37%) of these biopsies were obtained by primary care clinicians. Additionally, five dysplastic nevi were diagnosed on pathology, and four of these biopsies were obtained by primary care clinicians.

Review of the TD eConsults found that 16% (14/89) had statements from the responding dermatologist that suggested the images were not ideal for interpretation. In nine of these 14 statements, the comment from the dermatologist was that the photos were blurry or out of focus; however most commented that this was slight. For the other five, comments such as “difficult to ascertain” and “hard to tell” were used when describing the photos. An in-person dermatology visit was recommended by the dermatologist for nine out of the 14, whereas no in-person visit was recommended for the other five.

Figure 1. Flow chart of dermatology consultation rates and type of 1139 photo sessionsTD, Teledermatology.

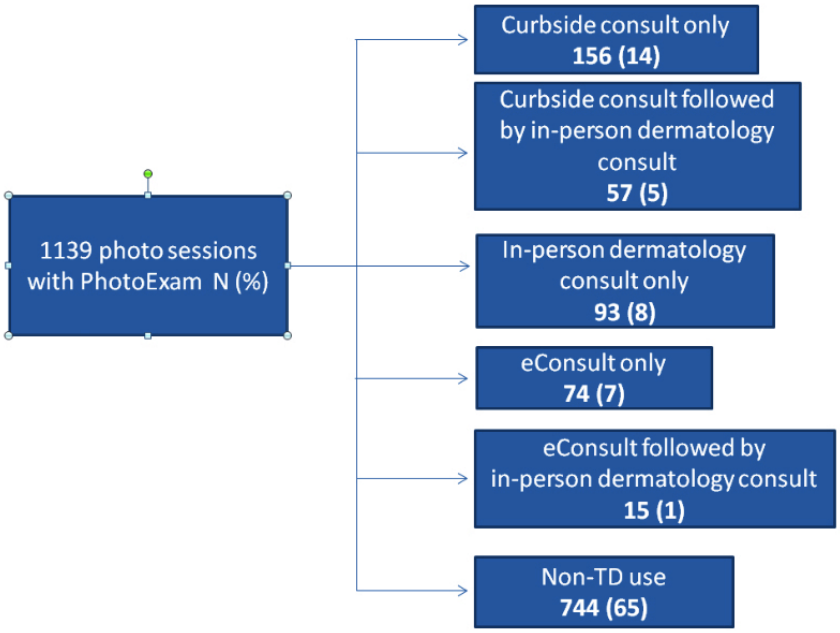


Figure 2. Bubble chart of percentage of photo sessions with dermatology input (via eConsult, curbside consult, and/or in person dermatology visit) by specialty and provider type. NP: nurse practitioner, PA: physician assistant, FM: family medicine, IM: internal medicine, PEDS: pediatrics.

Discussion

Principal Findings

This study describes our early experiences with a mobile-based PhotoExam app for clinical image capture in the primary care setting. Our primary findings highlight usage of the PhotoExam app by primary care clinicians for multiple purposes including TD. Notably, the predominant use (approximately two-thirds of sessions) of the clinical images captured using the PhotoExam app in our cohort was for purposes other than TD.

Whereas the PhotoExam tool is most accurately characterized as supporting an S&F TD model of care, given its situation in mobile devices, it also allows for near real-time consultation on dermatologic concerns at or close to the point of care. The PhotoExam TD model is thus a hybrid that combines the advantages of efficient still clinical image capture as in S&F TD and the near real-time dialogue inherent to VTC. The seamless transfer of clinical images acquired using the PhotoExam app into the shared EHR enables ease of access and timeliness of consultation by dermatology specialists. This allows the primary care clinician to dermatologist consultation to occur by curbside (telephone) consult, eConsult, or in-person visit. With this app, primary care clinicians are able to choose between multiple pathways to dermatology advice, depending on the needs of the patient and clinician. When utilizing the PhotoExam app for TD, requesting clinicians utilized the curbside dermatology consultation option more frequently than the eConsult pathway. We observed that when the PhotoExam was used for TD curbside and eConsults, three-quarters of the time there was no in-person dermatology visit within the following 30 days, suggesting that PhotoExam app use, in conjunction with TD consultation, averted the need for an in-person dermatology visit in the majority of instances where used, a finding supported by a recent review of telehealth literature [21].

The finding that approximately two-thirds of these photo sessions by primary care clinicians did not have any evidence for dermatology consultation within 30 days suggests that this type of mobile technology has value to primary care clinicians beyond its use in TD. The limited literature supports various applications for digital clinical photography in dermatologic care, including education of medical students and physicians-in-training, enhancement of clinical documentation in practice, archiving patient records, pre- and postoperative documentation for referring physicians, and diagnosis of melanoma [1]. During our manual chart review, we saw various non-TD examples of how primary care clinicians were using this technology in practice. The most common reason appeared to be for simply documenting exam findings for descriptive purposes. However, we also saw examples where the stated intention was to document the physical examination to follow a condition over time (such as an evolving cellulitis or to follow acne after initiation of treatment). We saw instances in which the primary care clinician commented in their note that they obtained the photograph to be able to obtain TD input if their initial treatment of the skin disorder did not lead to resolution. When photographs were obtained before in-person dermatology consultations, we saw comments that photos were taken to document the finding on the day of the primary care visit, in case exam findings changed before the in-person dermatology consultation. We also saw examples of the app being used for documentation in nondermatologic settings, including hand injuries (used in conjunction with orthopedic and plastic surgery curbside consults), as well as documentation of ocular infection exam findings. Finally, we saw instances where clinicians used the app to capture photographs taken by patients on their own devices and shared with the clinician. We can envision myriad other possibilities for this app including presurgical or preprocedural consultations [22] and documenting acquired deformities, congenital malformations, and trauma findings [23,24].

NPs and PAs were more likely to utilize the PhotoExam for TD as compared with physicians, whereas pediatric clinicians were more likely to utilize the PhotoExam for TD as compared with family medicine and internal medicine clinicians. Though of a different methodological design, a study comparing eConsult referral patterns to specialists between NPs and physicians in family medicine found that NPs directed a higher proportion of their eConsults to dermatology when compared with family medicine physicians [25]. Nurse practitioners and PAs in our study had a fairly consistent use of the PhotoExam for TD (37%-43% of sessions) across specialties, whereas there was a much broader range for rates of physicians using PhotoExam for TD across specialties with a range of 17% to 49% of sessions utilizing TD among physicians. The difference in rates of using PhotoExam for TD across specialties appears to be driven by the very low rate of family medicine residents using the PhotoExam for TD (17% of sessions) and the high rate of pediatric staff physicians using the PhotoExam for TD (49% of sessions). Our study design does not allow us to assess whether this difference is because of some provider types requesting more TD services than other provider types (thus leading to a higher percentage of PhotoExam sessions for TD) or if the difference is because of some provider types utilizing the PhotoExam more frequently for non-TD purposes (such as clinical documentation). If the latter is the case, this could lead to a higher overall use of PhotoExam with a lower rate of using the PhotoExam for TD without representing a difference in overall TD rates between specialties or provider types per patients seen in clinic.

We found that the majority of photos submitted by primary care clinicians involved only one anatomic site, which suggests that primary care clinicians are either utilizing this primarily for discrete skin lesions or conditions rather than diffuse conditions or clinicians are only taking single site photos of diffuse conditions. We found only a very small percentage (2%) of the sessions were of lesions diagnosed as malignant, though it is possible that some of these lesions may have been diagnosed as malignant at a date after our data collection period.

Review of the times that the app was used for a TD eConsult revealed that lack of a clear photo was not a major driver for in-person dermatology exam recommendations as this was seen only 10% (9/89) of the time. We could not retrospectively examine what the curbside dermatologist thought about the quality of the photos submitted. Documentation of these curbside

consultations was done by the primary care clinician rather than the responding dermatologist.

Limitations

Limitations of our study include the retrospective nature of our review that limited our ability to ensure that the extracted data and the actions of the clinician were concordant. For example, clinicians may have performed a curbside consultation with a dermatologist but not documented that conversation in the EHR. We did not collect data on primary care TD consults before the release of the PhotoExam app, and therefore, we were unable to assess whether there was a change in these rates with implementation of the PhotoExam app. Data extraction was performed by only one reviewer, though a standard process was used. Though our review of eConsults suggested that poor image quality was not a major driver for in-person dermatology referrals, our study design did not allow us to determine how often this was a factor for in-person dermatology referrals after curbside consults. Finally, our study population was primarily white, which may limit generalizability.

Conclusions

The PhotoExam app appears to be used as intended in primary care as a tool for TD. However, providers have also found other creative uses of the app, including augmenting textually constrained physical examination documentation, inputting into the medical record patient-taken photos (via a screenshot), documenting disease progression or regression, and consulting nondermatology specialists. Areas for future study include assessing health outcomes, app use in different medical settings, and clinician and patient satisfaction. We predict that the ubiquity of mobile devices with cameras and the availability of apps such as PhotoExam will accelerate the establishment of clinical photography as the standard of care for documentation in primary care and other areas of medicine.

Conflicts of Interest

None declared.

References

1. Ratner D, Thomas CO, Bickers D. The uses of digital photography in dermatology. J Am Acad Dermatol 1999 Nov;41(5 Pt 1):749-756. [Medline: 10534639]
2. Coates SJ, Kvedar J, Granstein RD. Teledermatology: from historical perspective to emerging techniques of the modern era: part II: emerging technologies in teledermatology, limitations and future directions. J Am Acad Dermatol 2015 Apr;72(4):577-86; quiz 587-88. [doi: 10.1016/j.jaad.2014.08.014] [Medline: 25773408]
3. Kvedar JC, Edwards RA, Menn ER, Mofid M, Gonzalez E, Dover J, et al. The substitution of digital images for dermatologic physical examination. Arch Dermatol 1997 Feb;133(2):161-167. [Medline: 9041828]
4. Pathipati AS, Lee L, Armstrong AW. Health-care delivery methods in teledermatology: consultative, triage and direct-care models. J Telemed Telecare 2011;17(4):214-216. [doi: 10.1258/jtt.2010.010002] [Medline: 21508078]
5. McKoy K, Antoniotti NM, Armstrong A, Bashshur R, Bernard J, Bernstein D, et al. Practice Guidelines for Teledermatology. Telemed J E Health 2016 Dec;22(12):981-990. [doi: 10.1089/tmj.2016.0137] [Medline: 27690203]
6. Datta SK, Warshaw EM, Edison KE, Kapur K, Thottapurathu L, Moritz TE, et al. Cost and Utility Analysis of a Store-and-Forward Teledermatology Referral System: A Randomized Clinical Trial. J Am Med Assoc Dermatol 2015 Dec 01;151(12):1323-1329. [doi: 10.1001/jamadermatol.2015.2362] [Medline: 26375589]
7. Snoswell C, Finnane A, Janda M, Soyer HP, Whitty JA. Cost-effectiveness of store-and-forward teledermatology: a systematic review. J Am Med Assoc Dermatol 2016 Jun 01;152(6):702-708. [doi: 10.1001/jamadermatol.2016.0525] [Medline: 27074289]
8. Mounessa JS, Chapman S, Braunberger T, Qin R, Lipoff JB, Dellavalle RP, et al. A systematic review of satisfaction with teledermatology. J Telemed Telecare 2017 Jan 01:1357633X17696587. [doi: 10.1177/1357633X17696587] [Medline: 28350281]
9. Moreno-Ramírez D, Argenziano G. Teledermatology and mobile applications in the management of patients with skin lesions. Acta Derm Venereol 2017 Jul 05;Suppl 218:31-35 [FREE Full text] [doi: 10.2340/00015555-2718] [Medline: 28676881]
10. Nelson CA, Takeshita J, Wanat KA, Bream KD, Holmes JH, Koenig HC, et al. Impact of store-and-forward (SAF) teledermatology on outpatient dermatologic care: a prospective study in an underserved urban primary care setting. J Am Acad Dermatol 2016 Mar;74(3):484-90.e1. [doi: 10.1016/j.jaad.2015.09.058] [Medline: 26679528]
11. Nami N, Massone C, Rubegni P, Cevenini G, Fimiani M, Hofmann-Wellenhof R. Concordance and time estimation of store-and-forward mobile teledermatology compared to classical face-to-face consultation. Acta Derm Venereol 2015 Jan;95(1):35-39 [FREE Full text] [doi: 10.2340/00015555-1876] [Medline: 24889827]
12. Landman A, Emani S, Carlile N, Rosenthal DI, Semakov S, Pallin DJ, et al. A mobile app for securely capturing and transferring clinical images to the electronic health record: description and preliminary usability study. JMIR Mhealth Uhealth 2015 Jan 02;3(1):e1 [FREE Full text] [doi: 10.2196/mhealth.3481] [Medline: 25565678]
13. Armstrong AW, Wu J, Kovarik CL, Goldyne ME, Oh DH, McKoy KC, et al. State of teledermatology programs in the United States. J Am Acad Dermatol 2012 Nov;67(5):939-944. [doi: 10.1016/j.jaad.2012.02.019] [Medline: 22459360]

14. Leger MC, Wu T, Haimovic A, Kaplan R, Sanchez M, Cohen D, et al. Patient perspectives on medical photography in dermatology. Dermatol Surg 2014 Sep;40(9):1028-1037. [doi: 10.1097/01.DSS.0000452632.22081.79] [Medline: 25099296]
15. Hsieh C, Yun D, Bhatia AC, Hsu JT, Ruiz de Luzuriaga AM. Patient perception on the usage of smartphones for medical photography and for reference in dermatology. Dermatol Surg 2015 Jan;41(1):149-154. [doi: 10.1097/DSS.0000000000000213] [Medline: 25533160]
16. Hacard F, Maruani A, Delaplace M, Caille A, Machet L, Lorette G, et al. Patients' acceptance of medical photography in a French adult and paediatric dermatology department: a questionnaire survey. Br J Dermatol 2013 Aug;169(2):298-305. [doi: 10.1111/bjd.12345] [Medline: 23551168]
17. Tomlinson J, Meads B, Myers A. Use of camera phones for hospital clinical communication: patient perspectives. ANZ J Surg 2014 Mar;84(3):105-106. [Medline: 24738157]
18. Sikka N, Carlin KN, Pines J, Pirri M, Strauss R, Rahimi F. The use of mobile phones for acute wound care: attitudes and opinions of emergency department patients. J Health Commun 2012;17(Suppl 1):37-42; quiz 42-3. [doi: 10.1080/10810730.2011.649161] [Medline: 22548597]
19. Kunde L, McMeniman E, Parker M. Clinical photography in dermatology: ethical and medico-legal considerations in the age of digital and smartphone technology. Australas J Dermatol 2013 Aug;54(3):192-197. [doi: 10.1111/ajd.12063] [Medline: 23713892]
20. Wyatt KD, Willaert BN, Pallagi PJ, Uribe RA, Yiannias JA, Hellmich TR. PhotoExam: adoption of an iOS-based clinical image capture application at Mayo Clinic. Int J Dermatol 2017 May 11 [Epub ahead of print]. [doi: 10.1111/ijd.13648] [Medline: 28497467]
21. Caffery LJ, Farjian M, Smith AC. Telehealth interventions for reducing waiting lists and waiting times for specialist outpatient services: a scoping review. J Telemed Telecare 2016 Dec;22(8):504-512. [doi: 10.1177/1357633X16670495] [Medline: 27686648]
22. Chan N, Charette J, Dumestre DO, Fraulin FO. Should 'smart phones' be used for patient photography? Plast Surg (Oakv) 2016;24(1):32-34 [FREE Full text] [Medline: 27054136]
23. Bloemen EM, Rosen T, Cline Schiroo JA, Clark S, Mulcare MR, Stern ME, et al. Photographing injuries in the acute care setting: development and evaluation of a standardized protocol for research, forensics, and clinical practice. Acad Emerg Med 2016 Dec;23(5):653-659 [FREE Full text] [doi: 10.1111/acem.12955] [Medline: 26932497]
24. Moumoulidis I, Mani N, Patel H, Leong P. A novel use of photo messaging in the assessment of nasal fractures. J Telemed Telecare 2007;13(8):387-390. [doi: 10.1258/135763307783064412] [Medline: 18078548]
25. Liddy C, Deri Armstrong C, McKellips F, Keely E. A comparison of referral patterns to a multispecialty eConsultation service between nurse practitioners and family physicians: the case for eConsult. J Am Assoc Nurse Pract 2016 Mar;28(3):144-150. [doi: 10.1002/2327-6924.12266] [Medline: 25965249]

Abbreviations

EHR: electronic health record
iOS: iPhone operating system
NP: nurse practitioner
PA: physician assistant
S&F: store and forward
TD: teledermatology
VTC: video teleconferencing

Edited by G Eysenbach; submitted 22.06.17; peer-reviewed by N Sikka, E Krupinski; comments to author 13.07.17; revised version received 30.08.17; accepted 10.09.17; published 08.11.17.

Please cite as:
Pecina JL, Wyatt KD, Comfere NI, Bernard ME, North F
Uses of Mobile Device Digital Photography of Dermatologic Conditions in Primary Care
JMIR Mhealth Uhealth 2017;5(11):e165
URL: http://mhealth.jmir.org/2017/11/e165/
doi:10.2196/mhealth.8257
PMID:29117934

Original Paper

Time Gain Needed for In-Ambulance Telemedicine: Cost-Utility Model

Alexis Valenzuela Espinoza[1], MSc; Stefanie Devos[1], PhD; Robbert-Jan van Hooff[2,3], MD, PhD; Maaike Fobelets[1], MSc; Alain Dupont[4], MD, PhD; Maarten Moens[5,6], MD, PhD; Ives Hubloue[7,8], MD, PhD; Door Lauwaert[7,8], RN; Pieter Cornu[4], PhD; Raf Brouns[2,9], MD, PhD; Koen Putman[1], PhD

[1]Interuniversity Center for Health Economics Research, Vrije Universiteit Brussel, Brussels, Belgium

[2]Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel, Brussels, Belgium

[3]Neurovascular Center, Department of Neurology, Zealand University Hospital, Roskilde, Denmark

[4]Research Group Clinical Pharmacology and Clinical Pharmacy, Vrije Universiteit Brussel, Brussels, Belgium

[5]Department of Neurosurgery, Universitair Ziekenhuis Brussel, Vrije Universiteit Brussel, Brussels, Belgium

[6]Center for Neurosciences, Vrije Universiteit Brussel, Brussels, Belgium

[7]Department of Emergency Medicine, Universitair Ziekenhuis Brussel, Vrije Universiteit Brussel, Brussels, Belgium

[8]Research Group on Emergency and Disaster Medicine, Vrije Universiteit Brussel, Brussels, Belgium

[9]Department of Neurology, ZorgSaam Hospital, Terneuzen, Netherlands

Corresponding Author:
Alexis Valenzuela Espinoza, MSc
Interuniversity Center for Health Economics Research
Vrije Universiteit Brussel
Laarbeeklaan 103
Brussels,
Belgium
Phone: 32 47 050 57 96
Email: alexis.valenzuela.espinoza@vub.be

Abstract

Background: Stroke is a very time-sensitive pathology, and many new solutions target the optimization of prehospital stroke care to improve the stroke management process. In-ambulance telemedicine, defined by live bidirectional audio-video between a patient and a neurologist in a moving ambulance and the automated transfer of vital parameters, is a promising new approach to speed up and improve the quality of acute stroke care. Currently, no evidence exists on the cost effectiveness of in-ambulance telemedicine.

Objective: We aim to develop a first cost effectiveness model for in-ambulance telemedicine and use this model to estimate the time savings needed before in-ambulance telemedicine becomes cost effective.

Methods: Current standard stroke care is compared with current standard stroke care supplemented with in-ambulance telemedicine using a cost-utility model measuring costs and quality-adjusted life-years (QALYs) from a health care perspective. We combine a decision tree with a Markov model. Data from the UZ Brussel Stroke Registry (2282 stroke patients) and linked hospital claims data at individual level are combined with literature data to populate the model. A 2-way sensitivity analysis varying both implementation costs and time gain is performed to map the different cost-effective combinations and identify the time gain needed for cost effectiveness and dominance. For several modeled time gains, the cost-effectiveness acceptability curve is calculated and mapped in 1 figure.

Results: Under the base-case scenario (implementation cost of US $159,425) and taking a lifetime horizon into account, in-ambulance telemedicine is a cost-effective strategy compared to standard stroke care alone starting from a time gain of 6 minutes. After 12 minutes, in-ambulance telemedicine becomes dominant, and this results in a mean decrease of costs by US –$30 (95% CI –$32 to –$29) per patient with 0.00456 (95% CI 0.00448 to 0.00463) QALYs on average gained per patient. In over 82% of all probabilistic simulations, in-ambulance telemedicine remains under the cost-effectiveness threshold of US $47,747.

Conclusions: Our model suggests that in-ambulance telemedicine can be cost effective starting from a time gain of 6 minutes and becomes a dominant strategy after approximately 15 minutes. This indicates that in-ambulance telemedicine has the potential to become a cost-effective intervention assuming time gains in clinical implementations are realized in the future.

(JMIR Mhealth Uhealth 2017;5(11):e175) doi:10.2196/mhealth.8288

KEYWORDS

telemedicine; prehospital; stroke; cost effectiveness

Introduction

Stroke is a very time-sensitive pathology, and many new solutions target the optimization of prehospital stroke care to improve the stroke management process [1]. One approach to speed up the stroke care process is the deployment of mobile stroke units (MSUs) that focus on the prehospital diagnosis and intravenous administration of recombinant tissue plasminogen activator (IVT) [2]. This is achieved by bringing the computed tomography (CT) scan to the patient, and time gains of 15 minutes between emergency call to IVT have been realized using this method [3]. In-ambulance telemedicine is another promising approach to reduce delays of the in-hospital stroke response by gathering and transferring relevant diagnostic information while the patient is underway to the hospital and therefore facilitating the clinical decision making on performing a CT scan and treatment initiation [4]. Recent progress in mobile connectivity enables virtually every ambulance to be equipped with telemedicine solutions, and several projects confirm the medical interest in this approach [4-8]. In-ambulance telemedicine allows head-to-toe examination of each patient through bidirectional audio-video communication between the ambulance and a remote teleconsultant and the secure transfer of medical data during emergency transportation of patients to a care facility. Pilot studies on stroke patients have shown that 24/7 in-ambulance telemedicine support is feasible, and stroke-specific information can be collected and communicated to the in-hospital team during emergency ambulance transportation [4,5]. The use of in-ambulance telemedicine is well accepted by patients and emergency personnel [5,9]. The combination of prehospital triage, early notification of the receiving in-hospital team, and communication of stroke-specific information by a remote stroke expert while the patient is being transported to the hospital has the potential to speed up the stroke diagnosis and treatment [4]. A time gain of 20 minutes has the potential to improve the probability of a favorable outcome after intravenous thrombolysis by 2.3% in a mixed stroke population [10] and is associated with reduced in-hospital mortality, lower risk of symptomatic intracranial hemorrhage, increased chance of independency at discharge, and increased probability to be discharged home [11]. Novel endovascular therapies are also highly time sensitive, and probabilities of favorable outcome increase relevantly when delays to treatment initiation decrease [12].

Currently, no evidence exists on the cost effectiveness of in-ambulance telemedicine. We aim to develop a model which predicts the potential costs and benefits associated with this new in-ambulance telemedicine approach. Consequently, this model allows identification of the minimum time gain that is needed before in-ambulance telemedicine becomes cost effective.

Methods

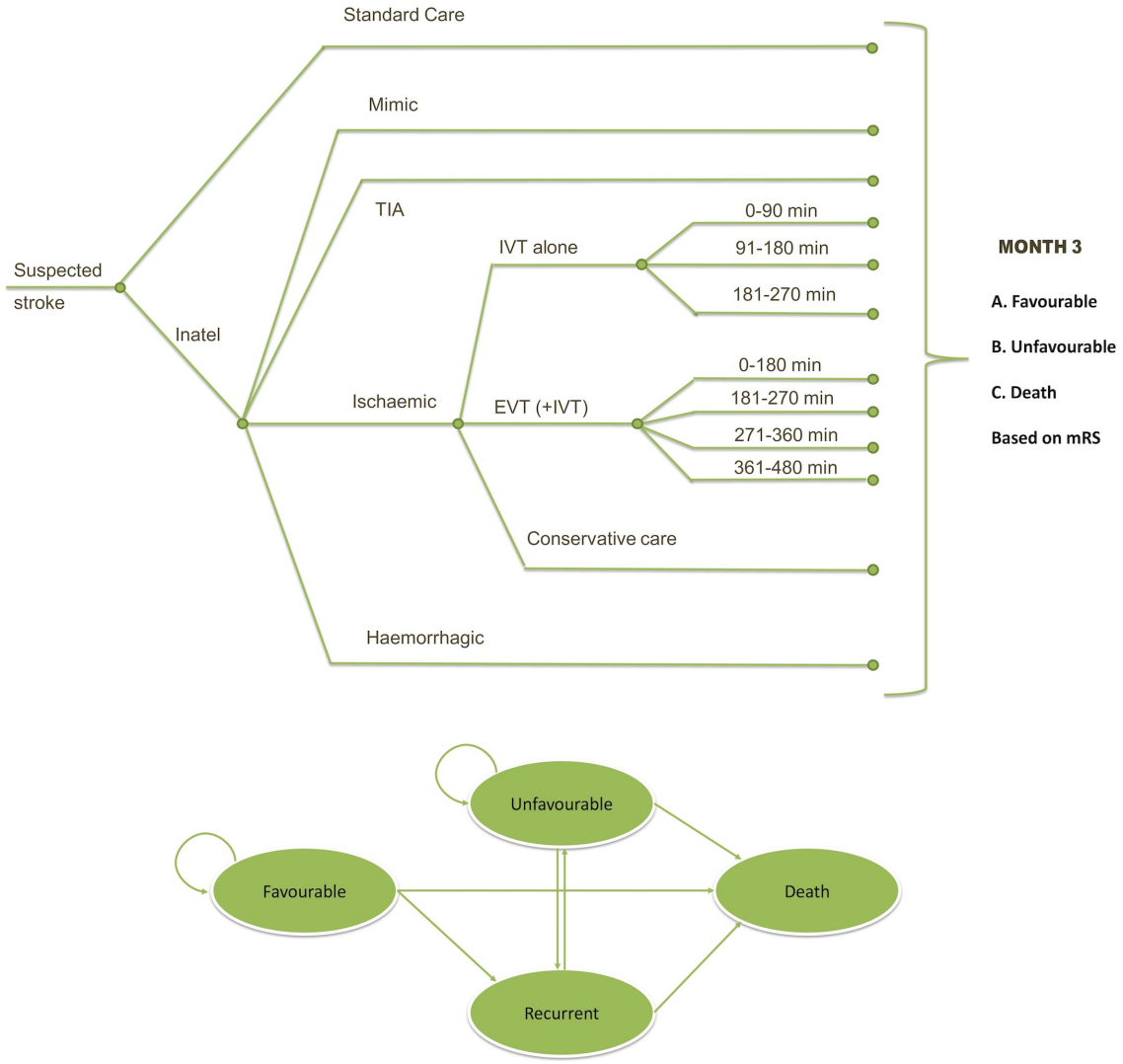
Model Description

Current standard stroke care is compared with current standard stroke care supplemented with in-ambulance telemedicine using a cost-utility model measuring costs and quality adjusted life-years (QALYs) from a health care perspective. We combine a decision tree model (3 months) with a Markov model using a lifetime horizon (Figure 1). One possible implementation of the intervention (in-ambulance telemedicine) was previously described [4]. In-ambulance telemedicine allows the automated transmission of vital parameters (heart rate, blood oxygen saturation, and systolic and diastolic blood pressure), glycemia, electronic patient identification, functional assessments, and prehospital notification of the in-hospital team. Teleconsultants are not required to remain in the hospital to ensure 24/7 coverage because telemedicine support can be provided from any location with access to the Internet. A report of the teleconsultation containing all available information is immediately sent to the in-hospital team. Teleconsultants and ambulance personnel are trained to adequately use the telemedicine system. The in-hospital team is taught how to securely access and interpret the teleconsultation report.

All patients with suspicion of acute stroke are included in the model and are divided into 4 main categories: (1) stroke mimic, (2) transient ischemic attack (TIA), (3) ischemic stroke, and (4) hemorrhagic stroke. Ischemic strokes are divided into 3 treatment groups: (1) intravenous administration of IVT alone, (2) intravenous administration of IVT in combination with endovascular treatment (EVT), or (3) conservative care. Patients in the in-ambulance telemedicine model can either receive in-ambulance telemedicine on top of standard stroke care or standard stroke care alone. This additional arm accounts for the missed opportunities related to the accuracy of dispatchers to recognize a stroke and the proportion of patients not transported by an ambulance equipped with telemedicine technology.

To effectively model the time-sensitive nature of IVT with and without EVT, the treatment effect per time interval is modeled. For IVT alone, we assume that patients are treated up to 4.5 hours after stroke onset [13]. In combination with EVT, a positive effect is observed until 8 hours after stroke onset [12].

Figure 1. Decision tree and Markov model for in-ambulance telemedicine for suspected stroke patients.



TIA = Transient Ischemic Attack; IVT = Intravenous Thrombolysis; EVT = Endovascular Therapy; mRS = modified Rankin Scale; Inatel = In-ambulance telemedicine

The comparator of standard stroke care is based on the performance indicators of Universitair Ziekenhuis Brussel (UZ Brussel), the university hospital of the Free University of Brussels (VUB), between 2011 and 2014. In this center, in-hospital prealerting of potential stroke patients and streamlined in-hospital workflows are part of standard medical practice [4].

Probabilities and Utilities

For each end node in the decision tree, the modified Rankin Scale (mRS) at month 3 is used to classify patients into 3 categories: (1) favorable outcome (mRS 0-2), (2) unfavorable outcome (mRS 3-5), and (3) death (mRS 6). Data from the UZ Brussel Stroke Registry (2282 suspected stroke patients admitted to the UZ Brussel Stroke Unit between February 2009 and February 2014) are combined with literature data to populate the model for standard care (Table 1 and Multimedia Appendix 1). Each mRS group is associated with utilities. The cycle length of the Markov model is 1 year with a lifetime horizon. Patients enter the Markov model in the favorable or unfavorable state and transition to their current state, to a death state, or to a recurrent stroke state. Transition after a recurrent stroke is limited to the unfavorable or death state.

Table 1. Parameters used to populate the standard care model.

Parameter	Base-case value (probability)	Source
Probabilities		
After Suspected Stroke		
Stroke mimic	0.22	Stroke Registry UZ Brussel[a]
TIA[b]	0.05	Sheppard et al [14]
Ischemic stroke	0.66	Stroke Registry UZ Brussel
Hemorrhagic stroke	0.07	Stroke Registry UZ Brussel
After ischemic stroke		
IVT[c]	0.15	Stroke Registry UZ Brussel
EVT[d]	0.05	Vanacker et al [15]
Conservative treatment	0.8	Stroke Registry UZ Brussel
After ischemic stroke and IVT		
0-90 min	0.12	OTT[e] distributions from Lees et al [16]
91-180 min	0.24	OTT distributions from Lees et al [16]
181-270 min	0.64	OTT distributions from Lees et al [16]
After ischemic stroke and EVT		
0-180 min	0.16	Campbell et al [17]
181-270 min	0.21	Campbell et al [17]
271-360 min	0.41	Campbell et al [17]
361-480 min	0.23	Campbell et al [17]
After stroke mimic		
Favorable	0.78	Stroke Registry UZ Brussel
Unfavorable	0.04	Stroke Registry UZ Brussel
Death	0.17	Stroke Registry UZ Brussel
After TIA		
Favorable	0.82	Stroke Registry UZ Brussel
Unfavorable	0.14	Stroke Registry UZ Brussel
Death	0.04	Stroke Registry UZ Brussel
After ischemic stroke and conservative treatment		
Favorable	0.48	Wardlaw et al [18]
Unfavorable	0.04	Wardlaw et al [18]
Death	0.12	Wardlaw et al [18]
After ischemic stroke and IVT		
0-90 min		
Favorable	0.7	Lees et al [16]
Unfavorable	0.18	Lees et al [16]
Death	0.12	Lees et al [16]
91-180 min		
Favorable	0.59	Lees et al [16]
Unfavorable	0.29	Lees et al [16]
Death	0.12	Lees et al [16]

Parameter	Base-case value (probability)	Source
181-270 min		
Favorable	0.55	Lees et al [16]
Unfavorable	0.33	Lees et al [16]
Death	0.12	Lees et al [16]
After ischemic stroke and EVT		
0-180 min		
Favorable	0.78	Multimedia Appendix 1, Fransen et al [12]
Unfavorable	0.1	Multimedia Appendix 1, Fransen et al [12]
Death	0.12	Multimedia Appendix 1, Fransen et al [12]
181-270 min		
Favorable	0.70	Multimedia Appendix 1, Fransen et al [12]
Unfavorable	0.18	Multimedia Appendix 1, Fransen et al [12]
Death	0.12	Multimedia Appendix 1, Fransen et al [12]
271-360 min		
Favorable	0.59	Multimedia Appendix 1, Fransen et al [12]
Unfavorable	0.29	Multimedia Appendix 1, Fransen et al [12]
Death	0.12	Multimedia Appendix 1, Fransen et al [12]
361-480 min		
Favorable	0.51	Multimedia Appendix 1, Fransen et al [12]
Unfavorable	0.29	Multimedia Appendix 1, Fransen et al [12]
Death	0.12	Multimedia Appendix 1, Fransen et al [12]
After hemorrhagic stroke		
Favorable	0.44	Anderson et al [19]
Unfavorable	0.44	Anderson et al [19]
Death	0.12	Anderson et al [19]
Utilities		
Utility in the favorable state (mRS[f] 0-2)	0.74	Dorman et al [20]
Utility in the unfavorable state (mRS 3-5)	0.38	Dorman et al [20]
Utility in the death state	0	
Utility in the recurrent state	0.34	Morris et al [21]
Markov transitions		
Probability recurrent stroke	0.05	Sandercock et al [22]
Increased mortality risk after recurrent stroke	0.25	Sandercock et al [22]
Multiplier for age-specific mortality among stroke patients	2.5	Sandercock et al [22]
Mortality after stroke		
70-74 years	0.05	Belgian mortality statistics corrected for age-specific mortality among stroke patients
75-79 years	0.08	Belgian mortality statistics corrected for age-specific mortality among stroke patients
80-84 years	0.14	Belgian mortality statistics corrected for age-specific mortality among stroke patients
85-89 years	0.26	Belgian mortality statistics corrected for age-specific mortality among stroke patients

Parameter	Base-case value (probability)	Source
90+ years	0.45	Belgian mortality statistics corrected for age-specific mortality among stroke patients
Other		
Average age of stroke patients	73	Thijs et al [23]
Discount rate for costs	0.03	KCE[g] [24]
Discount rate for utilities	0.015	KCE [24]

[a]UZ Brussel: Universitair Ziekenhuis Brussel.

[b]TIA: transient ischemic attack.

[c]IVT: intravenous administration of recombinant tissue plasminogen activator.

[d]EVT: endovascular treatment.

[e]OTT: onset to treatment time.

[f]mRS: modifed Rankin Scale.

[g]KCE: Belgian Health Care Knowledge Centre.

Costs

For the costs per treatment arm, we link hospital data and emergency claims data (including all payer costs) at the individual patient level (Multimedia Appendices 2 and 3). Claims for drugs, clinical biology, medical imaging, physicians' honoraria, other claims charged to patients (copayments), and health insurances were included. To calculate the total hospital cost, a fixed day price was added according to the year of admission. This fixed day price covers the financing of nonmedical hospital activities. For this study, the weighted average per diem prices (across Belgian hospitals) was used [24]. All costs are expressed in and discounted to 2014, and Euro and US $ equivalents are calculated using the average 2014 exchange rate (€1=US $1.329). We do not model any productivity loss, as the average age of our patient cohort is 73 years. The cost-effectiveness threshold is set at US $47,747 (€35,927), the gross domestic product per capita of Belgium in 2014.

Impact of In-Ambulance Telemedicine

The impact of in-ambulance telemedicine on top of standard care is modeled by assuming an average time gain ranging from 5 to 60 minutes (Table 2 and Multimedia Appendix 4). This influences the treatment of stroke patients in 2 fundamental ways. First, probabilities of a positive outcome after treatment with IVT and/or EVT increase, since more patients are shifted into an earlier time window. Second, more patients can be treated with IVT and/or EVT as more patients shift into the applicable time windows, 4.5 hours and 8 hours, respectively, after symptom onset. Costs of the intervention are modeled by adding a fee per teleconsultation and a fixed fee per ambulance in which a telemedicine device is installed (Table 2 and Multimedia Appendix 5). Training of all stakeholders and mobile connectivity costs are included in the telemedicine installation cost.

Based on previous in-ambulance telemedicine pilot studies [4,5], we assume that 150 patients can be treated with 1 ambulance on a yearly basis, resulting in 3 ambulances to be equipped with the telemedicine technology for a patient cohort of 1000 suspected stroke patients (390 patients receiving in-ambulance telemedicine). All other parameters are assumed equal to standard medical stroke care.

Model Output

Costs and QALYs are used to calculate the incremental cost-effectiveness ratio (ICER) after 3 months (decision tree only) and after a lifetime horizon (decision tree plus Markov model). Time gain after in-ambulance telemedicine is varied between 0 and 60 minutes, and cost for the implementation of in-ambulance telemedicine is varied between 50% and 400% of baseline cost in a 2-way sensitivity analysis, mapping the ICER for all combinations of both variables. Based on this analysis, we select the time gain for which in-ambulance telemedicine becomes cost effective. For this time gain, we perform a 1-way sensitivity analysis, varying all input parameters between 70% and 130% of their deterministic value and ranking the parameters according to the highest interval between calculated outcome parameters (both cost and QALYs are calculated). A probabilistic sensitivity analysis is applied using Monte Carlo simulations with 5000 bootstraps to account for the uncertainty around the input parameters and assess the robustness of the model. The cost-effectiveness acceptability curve is constructed for 5, 10, 15, and 30 minutes of time gain (1000 bootstraps). The health economic model was built and runs in Excel (Microsoft Office Professional Plus 2013, Microsoft Corp), and the UZ Brussel Stroke Registry was analyzed using Stata MP 13.0 for Windows (StataCorp LLC). The study was approved by the UZ Brussel ethical committee, and the model was validated by MF.

Table 2. Adapted parameters for in-ambulance telemedicine under 12 minutes time gain on average per patient and additional costs.

Parameter	Base-case value	Source/assumption
Probabilities		
After suspected stroke		
Standard care	0.61	Multimedia Appendix 4
In-ambulance telemedicine	0.39	Multimedia Appendix 4
After ischemic stroke		
IVT[a]	0.19	Multimedia Appendix 4
EVT[b]	0.07	Multimedia Appendix 4
Conservative treatment	0.73	Multimedia Appendix 4
After ischemic stroke and IVT		
0-90 min	0.15	Multimedia Appendix 4
91-180 min	0.29	Multimedia Appendix 4
181-270 min	0.56	Multimedia Appendix 4
After ischemic stroke and EVT		
0-180 min	0.19	Multimedia Appendix 4
181-270 min	0.23	Multimedia Appendix 4
271-360 min	0.38	Multimedia Appendix 4
361-480 min	0.19	Multimedia Appendix 4
Costs, US $ (€)		
Cost per teleconsultation	142.89 (107.52)	Multimedia Appendix 5
Cost of installation of 1 telemedicine device	29,011 (26,000)	Offer from Zebra Academy
Estimated total cost for in-ambulance telemedicine for 390 treated patients in 1 year	159,425 (119,959)	Multimedia Appendix 5
Number of patients that can be treated with 1 device in 1 year	150	Activation rates of the PreSSUB-I[c] trial [4]

[a]IVT: intravenous administration of recombinant tissue plasminogen activator.

[b]EVT: endovascular treatment.

[c]PreSSUB-I: Prehospital Study at the Universitair Ziekenhuis Brussel I.

Results

Base-Case

Under the base-case scenario (implementation cost of US $159,425) and taking a lifetime horizon into account, in-ambulance telemedicine is a cost-effective strategy compared to standard stroke care alone, starting from a time gain of 6 minutes (Figure 2).

After 12 minutes, in-ambulance telemedicine becomes a dominant strategy over standard best medical practice (Table 3). In a cohort of 1000 patients, 4.9 QALYs are gained (0.005 QALY/patient) and US $4040 (€3040) in long-term costs are avoided (–$4/patient). The savings of earlier stroke treatment outweigh the cost for implementation of in-ambulance telemedicine (cost equals US $159,425 [€119,959] for 390 patients receiving in-ambulance telemedicine) and higher utilization rates of specific stroke treatments (IVT and EVT). After 3 months, unfavorable outcome is avoided in 2.42 additional patients, resulting in long-term savings for society.

Not taking into account these long-term savings, in-ambulance telemedicine yields an ICER of US $201,557/QALY (€151,660/QALY) after 3 months. This incremental cost per saved QALY is explained by the cumulative cost of the intervention and the costs associated with more IVT and EVT after implementation of in-ambulance telemedicine.

One-Way Sensitivity Analysis

One-way sensitivity analysis at 12 minutes time gain reveals that parameters involving outcome of ischemic stroke have the largest impact on calculated costs and QALYs (Multimedia Appendices 6-9). For likelihood of unfavorable outcome after conservative care, the incremental cost/patient varies between US –$252 (–€190) (130%) and US $244 (€184) (70%); for favorable outcome after conservative care, the incremental QALY/patient varies between 0.013 (70%) and –0.003 (130%). This is not surprising given the model's rationale (shift from conservative care to IVT) and the time-sensitive nature of IVT. We note, however, that these parameters are taken from the analysis of pooled randomized controlled trials (RCT) [16,18], and the time-sensitive nature of IVT has been confirmed in

larger clinical populations [11,25]. The proportion of patients receiving standard care (vs in-ambulance telemedicine) influences the outcome parameters of the model, indicating that regions with more ambulances equipped with telemedicine, with a higher ability of ambulance dispatchers to recognize a stroke, and with a higher proportion of patients being transported by ambulance will benefit more from in-ambulance telemedicine.

Two-Way Sensitivity Analysis

Under 4 times baseline costs (>US $500,000 implementation costs), in-ambulance telemedicine becomes cost effective after 19 minutes and is dominant after 39 minutes (Figure 2). For lower cost implementations (<US $70,000), in-ambulance telemedicine can be cost effective after 3 minutes and dominant after 7 minutes of achieved time gain.

Probabilistic Results

The cost-effectiveness acceptability curves (Figure 3) reveal that under probabilistic analysis more than 90% of simulations are cost effective at the threshold of US $47,747, starting from 15 minutes time gain. This number drops below 80% under a scenario of 10 minutes time gain.

Figure 2. Two-way sensitivity analysis for in-ambulance telemedicine compared to standard care. Implementation costs are varied between 0.5 and 4 times the base case cost and time-gain is varied between 0 and 60 minutes.

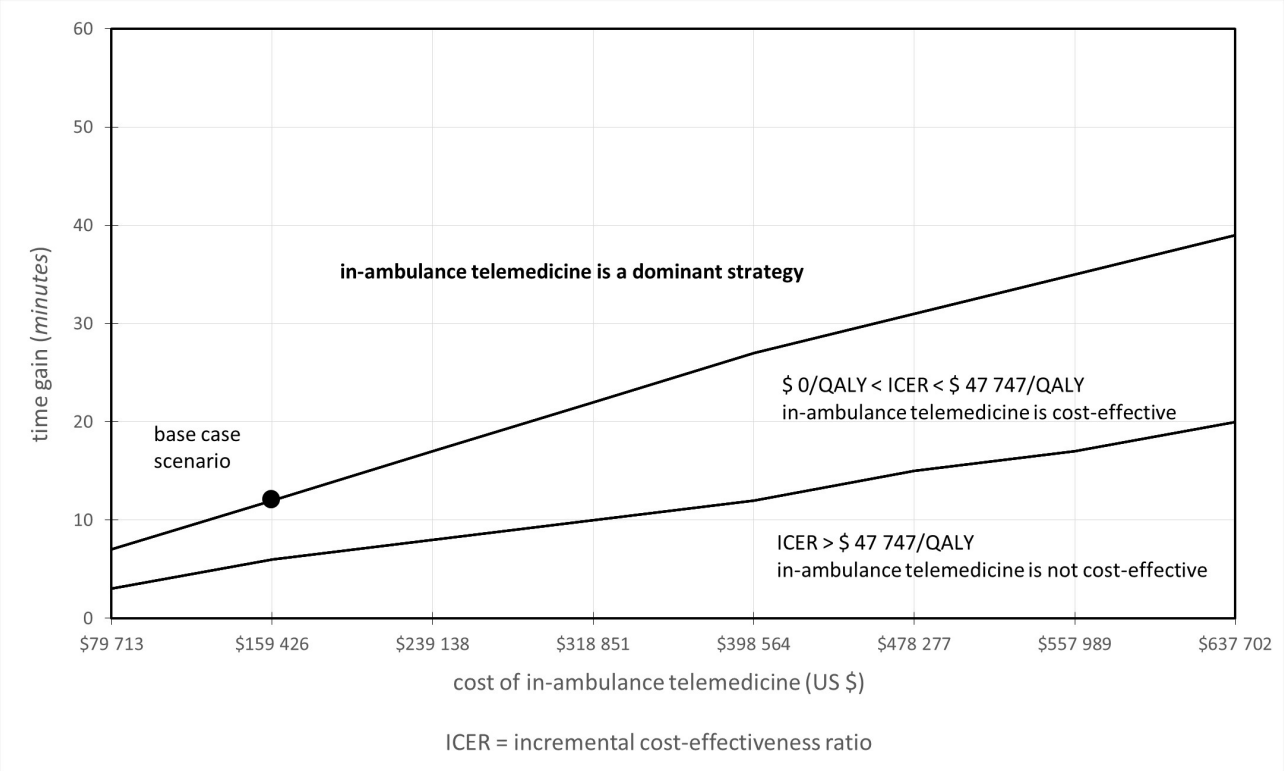


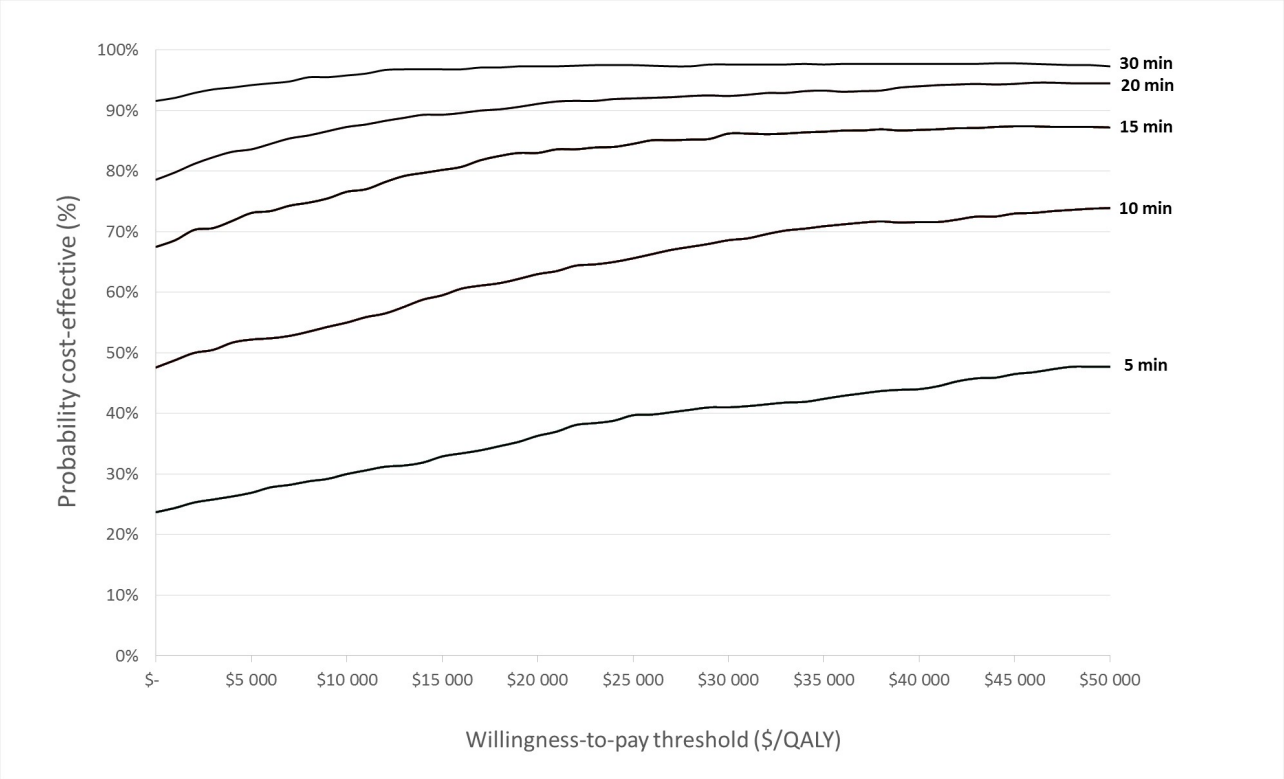
Table 3. Deterministic costs and quality-adjusted life-years after 3 months and after a lifetime horizon under 12 minutes time gain.

Cohort of 1000 patients	3 months		Lifetime horizon	
	Total costs ($)	Outcome (QALY[a])	Total costs ($)	Outcome (QALY)
Standard care	21,530,867	537.6	92,068,697	3649.8
In-ambulance telemedicine	21,706,449	538.4	92,064,657	3654.8
Difference	175,582	0.9	–4040	4.9
ICER[b], $/QALY	201,557		–817	

[a]QALY: quality-adjusted life-year.

[b]ICER: incremental cost-effectiveness ratio.

Figure 3. The cost-effectiveness acceptability curve for in-ambulance telemedicine compared to standard care is calculated for 1000 bootstraps per time interval (5, 10, 15, 20, and 30 minutes).



Discussion

Principal Findings

We report on the first ever comprehensive in-ambulance telemedicine cost-utility model combining a decision tree with a Markov model, using detailed cost information per treatment arm. In-ambulance telemedicine is dominant from a health care payer perspective starting from 12 minutes time gain.

Two previous publications describe health economic aspects of MSUs for improvement of prehospital stroke care [26,27]. Both models use the same source for calculation of improved outcome following faster IVT [16]. We applied the same methodology in our model by calculating the absolute risk difference based on the reported numbers needed to treat.

Limitations

No impact of in-ambulance telemedicine on mortality was modeled, even though evidence exists that earlier stroke treatment reduces in-hospital mortality [11]. We chose not to model mortality because sufficiently specific information on mortality is not available in RCTs and because it may be unlikely that our population of 2200 patients would be sufficiently powered to model a possible impact on mortality.

The impact of in-ambulance telemedicine can vary greatly from 1 hospital or region to another depending on the standard quality and speed of care. Streamlining in-hospital workflows will always be a crucial part of a successful in-ambulance telemedicine implementation and can influence costs and potential benefit. For example, realized time gains could be more modest if prehospital notification through mobile phones or tablets is already part of current practice.

The major limitation of the presented model lies in the absence of information from RCTs evaluating the effects of in-ambulance telemedicine on costs and patient outcome. This drawback was addressed by implementing only solid criteria originating from RCTs for the outcome parameters. However, combining data from multiple trials is not without risk, and the results of this model should be interpreted carefully.

Other Considerations

An advantage of in-ambulance telemedicine is the limited amount of additional resources needed from the hospital. If state-of-the art stroke care is available, no additional staff may be needed and existing ambulances can readily be equipped with the technology. However, depending on the catchment area and the number of stroke patients supported via in-ambulance telemedicine, additional teleconsultants on call may be needed. Although this could increase the organizational cost of in-ambulance telemedicine, more patients would be treated with in-ambulance telemedicine, which would further decrease the cost per patient and improve outcome. Training is required but was included in the cost of the telemedicine implementation.

Total implementation costs will decrease when in-ambulance telemedicine technology becomes more widely available. Lower cost alternatives such as tablet-based approaches are currently being investigated for remote stroke severity assessment in driving ambulances [6,7].

Widespread implementation of in-ambulance telemedicine will not only depend on its (cost-)effectiveness but also on the creation of the required legal framework. In Belgium, currently consultations are only officially recognized if face-to-face

contact between the patient and treating physician occurs. Other issues include clear regulations for reimbursement and liability.

Our analysis only takes benefits of the expected time gain into account. Other benefits of in-ambulance telemedicine include a lower risk of stroke misdiagnosis and consequently missed opportunities for treatment with IVT or EVT and triage of patients to inadequate facilities [28]. Estimates of missed stroke diagnosis by emergency personnel range from 22% to 47%, indicating the potential for in-ambulance telemedicine to curtail the risk of misdiagnosis [29,30]. Reducing missed opportunities was not modeled here to avoid double counting when combined with faster treatment effects.

We excluded the implementation of rapid blood pressure lowering for hemorrhagic stroke patients in our model, even though indications exist on the benefit of this approach [19]. In-ambulance telemedicine could increase the proportion of patients receiving rapid blood pressure lowering, resulting in an underestimation of the potential cost effectiveness.

Further, expert prehospital care may help avoid secondary brain damage, as the teleconsultant can support the ambulance personnel in obtaining and maintaining homeostasis during ambulance transportation through optimal application of the standard operating procedures for airway protection, blood oxygen saturation, arterial blood pressure, heart rate, cardiac arrhythmia, decreased level of consciousness, dysglycemia, and other supportive measures (eg, antiemetics, analgesics).

The combination of in-ambulance telemedicine with MSUs is another interesting approach that could further decrease costs of the MSU as it would avoid sending highly trained physicians into the field for each individual patient. This approach is feasible, and preliminary analysis has shown that median time savings of 23 minutes between alarm-to-CT times can be attained when compared to standard care [31].

Recent clinical trials showing impressive benefits from EVT [32-34] herald a new era in acute ischemic stroke care. This highly effective treatment is resulting in a paradigm shift toward optimization of prehospital stroke diagnosis, identification of suitable candidates for IVT and EVT, and patient triage to appropriate centers [35]. A care model that avoids secondary transportations of stroke patients from primary stroke centers to comprehensive stroke centers for EVT is expected to result in better patient outcome for at least 0.2% of the patient cohort [10]. We did not take these effects of in-ambulance telestroke into account as their supportive evidence currently is insufficient to allow robust modeling. This probably results in an underestimation of the benefits yielded by in-ambulance telemedicine.

In-ambulance telemedicine has the potential to improve the organization of care for other medical emergencies, further strengthening the cost-effectiveness potential of this technology. The use of tablet computers as a support system for general emergency medical services and better patient triage have shown a decrease in transportation times by ambulances, showing the possibilities for further innovation in emergency care organization and delivery using telemedicine and mobile health solutions [36].

We believe that this positive health economic evaluation can inspire decision makers in hospitals and governments to actively pursue the implementation of and further research on in-ambulance telemedicine.

Conclusions

In-ambulance telestroke is highly cost effective from a health care perspective, resulting in more QALYs and less costs starting from a realized time gain of 12 minutes. The model is not directly based on results from RCTs on the effects of in-ambulance telemedicine, and trials to further crystalize these effects in various care models are needed. Support from governments and hospitals to facilitate implementation in clinical practice is indispensable and can be justified by the dominant cost effectiveness of in-ambulance telemedicine under several scenarios both in terms of implementation costs and time gain.

Conflicts of Interest

AD, PC, IH, DL, KP, SD, and MF report no conflicts of interest. AVE has received funding from the Brussels Institute for Research and Innovation (INNOVIRIS). RJvH has received consulting honoraria from Boehringer-Ingelheim. RB is a senior clinical investigator of the Fonds Wetenschappelijk Onderzoek. He has received funding from the Strategic Research Project Growth Fund and the Industrial Research Fund of the Vrije Universiteit Brussel, the Wetenschappelijk Fonds Willy Gepts of the UZ Brussel, INNOVIRIS, the Future Internet Public-Private Partnership program of the European Commission, and the Caring Entrepreneurship Fund of the King Baudouin Foundation. He serves on the editorial board of Clinical Neurology and Neurosurgery and of the Journal of Translational Internal Medicine, and he has received consultancy or speaker honoraria from Pfizer, Medtronic, Shire Human Genetics Therapies, Sanofi-Aventis, Boehringer-Ingelheim, Daiichi Sankyo, Amgen, and Bayer. RB is cofounder of Zebra Academy. AVE, RJvH, and RB are holders of a patent on in-ambulance telestroke. MM has received consultancy or speaker honoraria from Medtronic and Pfizer.

Multimedia Appendix 1

Details on all parameters used to populate the standard-care model.

[PDF File (Adobe PDF File), 49KB - mhealth_v5i11e175_app1.pdf]

Multimedia Appendix 2

Details on the costs used in the model.

[PDF File (Adobe PDF File), 43KB - mhealth_v5i11e175_app2.pdf]

Multimedia Appendix 3

Details on 1-year poststroke costs.

[PDF File (Adobe PDF File), 50KB - mhealth_v5i11e175_app3.pdf]

Multimedia Appendix 4

Adapted probabilities for in-ambulance telemedicine.

[PDF File (Adobe PDF File), 33KB - mhealth_v5i11e175_app4.pdf]

Multimedia Appendix 5

Costs related to in-ambulance telemedicine.

[PDF File (Adobe PDF File), 38KB - mhealth_v5i11e175_app5.pdf]

Multimedia Appendix 6

Tornado input probabilities—incremental cost per patient.

[PDF File (Adobe PDF File), 255KB - mhealth_v5i11e175_app6.pdf]

Multimedia Appendix 7

Tornado input probabilities—incremental quality-adjusted life-years per patient.

[PDF File (Adobe PDF File), 152KB - mhealth_v5i11e175_app7.pdf]

Multimedia Appendix 8

Tornado input costs, utilities, and other parameters—incremental cost per patient.

[PDF File (Adobe PDF File), 167KB - mhealth_v5i11e175_app8.pdf]

Multimedia Appendix 9

Tornado input costs, utilities, and other parameters—incremental quality-adjusted life-years per patient.

[PDF File (Adobe PDF File), 46KB - mhealth_v5i11e175_app9.pdf]

References

1. Yperzeele L, Van Hooff R, De Smedt A, Valenzuela EA, Van de Casseye R, Hubloue I, et al. Prehospital stroke care: limitations of current interventions and focus on new developments. Cerebrovasc Dis 2014;38(1):1-9 [FREE Full text] [doi: 10.1159/000363617] [Medline: 25116305]
2. Fassbender K, Grotta JC, Walter S, Grunwald IQ, Ragoschke-Schumm A, Saver JL. Mobile stroke units for prehospital thrombolysis, triage, and beyond: benefits and challenges. Lancet Neurol 2017 Mar;16(3):227-237. [doi: 10.1016/S1474-4422(17)30008-X] [Medline: 28229894]
3. Ebinger M, Winter B, Wendt M, Weber JE, Waldschmidt C, Rozanski M, et al. Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial. JAMA 2014;311(16):1622-1631. [doi: 10.1001/jama.2014.2850] [Medline: 24756512]
4. Valenzuela EA, Van Hooff R, De Smedt A, Moens M, Yperzeele L, Nieboer K, PreSSUB-Consortium. Development and pilot testing of 24/7 in-ambulance telemedicine for acute stroke: prehospital stroke study at the Universitair Ziekenhuis Brussel-Project. Cerebrovasc Dis 2016;42(1-2):15-22 [FREE Full text] [doi: 10.1159/000444175] [Medline: 26950076]
5. Yperzeele L, Van Hooff R, De Smedt A, Valenzuela EA, Van Dyck R, Van de Casseye R, et al. Feasibility of AmbulanCe-Based Telemedicine (FACT) study: safety, feasibility and reliability of third generation in-ambulance telemedicine. PLoS One 2014 Oct;9(10):e110043 [FREE Full text] [doi: 10.1371/journal.pone.0110043] [Medline: 25343246]

6. Chapman SSN, Govindarajan P, Padrick MM, Lippman JM, McMurry TL, Resler BL, et al. A low-cost, tablet-based option for prehospital neurologic assessment: the iTREAT study. Neurology 2016 Jul 05;87(1):19-26 [FREE Full text] [doi: 10.1212/WNL.0000000000002799] [Medline: 27281534]
7. Barrett KM, Pizzi MA, Kesari V, TerKonda SP, Mauricio EA, Silvers SM, et al. Ambulance-based assessment of NIH Stroke Scale with telemedicine: A feasibility pilot study. J Telemed Telecare 2017 May;23(4):476-483. [doi: 10.1177/1357633X16648490] [Medline: 27177870]
8. Felzen M, Brokmann JC, Beckers SK, Czaplik M, Hirsch F, Tamm M, et al. Improved technical performance of a multifunctional prehospital telemedicine system between the research phase and the routine use phase: an observational study. J Telemed Telecare 2017 Apr;23(3):402-409. [doi: 10.1177/1357633X16644115] [Medline: 27080747]
9. Valenzuela EA, De Smedt A, Guldolf K, Vandervorst F, Van Hooff R, Fernandez TH, et al. Opinions and beliefs about telemedicine for emergency treatment during ambulance transportation and for chronic care at home. Interact J Med Res 2016 Mar 30;5(1):e9 [FREE Full text] [doi: 10.2196/ijmr.5015] [Medline: 27029999]
10. Lorenz MW, Lauer A, Foerch C. Quantifying the benefit of prehospital rapid treatment in acute stroke: benchmark for future innovative clinical trials. Stroke 2015 Nov;46(11):3168-3176 [FREE Full text] [doi: 10.1161/STROKEAHA.115.010445] [Medline: 26443833]
11. Saver JL, Fonarow GC, Smith EE, Reeves MJ, Grau-Sepulveda MV, Pan W, et al. Time to treatment with intravenous tissue plasminogen activator and outcome from acute ischemic stroke. JAMA 2013 Jun 19;309(23):2480-2488. [doi: 10.1001/jama.2013.6959] [Medline: 23780461]
12. Fransen PSS, Berkhemer OA, Lingsma HF, Beumer D, van den Berg LA, Yoo AJ, Multicenter Randomized Clinical Trial of Endovascular Treatment of Acute Ischemic Stroke in the Netherlands Investigators. Time to reperfusion and treatment effect for acute ischemic stroke: a randomized clinical trial. JAMA Neurol 2016 Feb;73(2):190-196. [doi: 10.1001/jamaneurol.2015.3886] [Medline: 26716735]
13. Jauch EC, Saver JL, Adams HP, Bruno A, Connors JJB, Demaerschalk BM, American Heart Association Stroke Council, Council on Cardiovascular Nursing, Council on Peripheral Vascular Disease, Council on Clinical Cardiology. Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2013 Mar;44(3):870-947 [FREE Full text] [doi: 10.1161/STR.0b013e318284056a] [Medline: 23370205]
14. Sheppard JP, Lindenmeyer A, Mellor RM, Greenfield S, Mant J, Quinn T, CLAHRC BBC investigators. Prevalence and predictors of hospital prealerting in acute stroke: a mixed methods study. Emerg Med J 2016 Jul;33(7):482-488 [FREE Full text] [doi: 10.1136/emermed-2014-204392] [Medline: 26949969]
15. Vanacker P, Lambrou D, Eskandari A, Mosimann PJ, Maghraoui A, Michel P. Eligibility and predictors for acute revascularization procedures in a stroke center. Stroke 2016 Jul;47(7):1844-1849 [FREE Full text] [doi: 10.1161/STROKEAHA.115.012577] [Medline: 27301945]
16. Lees KR, Bluhmki E, Brott TG, Toni D, Grotta JC, Albers GW, ECASS, ATLANTIS, NINDSEPITHET rt-PA Study Group, et al. Time to treatment with intravenous alteplase and outcome in stroke: an updated pooled analysis of ECASS, ATLANTIS, NINDS, and EPITHET trials. Lancet 2010 May 15;375(9727):1695-1703. [doi: 10.1016/S0140-6736(10)60491-6] [Medline: 20472172]
17. Campbell BCV, Hill MD, Rubiera M, Menon BK, Demchuk A, Donnan GA, et al. Safety and efficacy of solitaire stent thrombectomy: individual patient data meta-analysis of randomized trials. Stroke 2016 Mar;47(3):798-806 [FREE Full text] [doi: 10.1161/STROKEAHA.115.012360] [Medline: 26888532]
18. Wardlaw JM, Murray V, Berge E, del Zoppo G, Sandercock P, Lindley RL, et al. Recombinant tissue plasminogen activator for acute ischaemic stroke: an updated systematic review and meta-analysis. Lancet 2012 Jun 23;379(9834):2364-2372 [FREE Full text] [doi: 10.1016/S0140-6736(12)60738-7] [Medline: 22632907]
19. Anderson CS, Heeley E, Huang Y, Wang J, Stapf C, Delcourt C, INTERACT2 Investigators. Rapid blood-pressure lowering in patients with acute intracerebral hemorrhage. N Engl J Med 2013 Jun 20;368(25):2355-2365. [doi: 10.1056/NEJMoa1214609] [Medline: 23713578]
20. Dorman PJ, Waddell F, Slattery J, Dennis M, Sandercock P. Is the EuroQol a valid measure of health-related quality of life after stroke? Stroke 1997 Oct;28(10):1876-1882 [FREE Full text] [Medline: 9341688]
21. Morris S, Hunter R, Davie C, Thompson A, Walker H, Thomson N. Cost-Effectiveness Analysis of the London Stroke Service. London: University College; 2011.
22. Sandercock P, Berge E, Dennis M, Forbes J, Hand P, Kwan J, et al. Cost-effectiveness of thrombolysis with recombinant tissue plasminogen activator for acute ischemic stroke assessed by a model based on UK NHS costs. Stroke 2004 Jun;35(6):1490-1497 [FREE Full text] [doi: 10.1161/01.STR.0000126871.98801.6E] [Medline: 15105519]
23. Thijs V, Dewilde S, Putman K, Pince H. Cost of hospitalization for cerebrovascular disorders in Belgium. Acta Neurol Belg 2011 Jun;111(2):104-110. [Medline: 21748928]
24. Cleemput I, Neyt M, Van de Sande S, Thiry N. Belgische richtlijnen voor economische evaluaties en budget impact analyses: tweede editie. Health Technology Assessment. KCE report 183A. Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2012. URL: https://kce.fgov.be/sites/default/files/atoms/files/KCE_183A_economische_evaluaties_tweede_editie_0.pdf [accessed 2017-11-06] [WebCite Cache ID 6umqeRUcP]

25. Meretoja A, Strbian D, Mustanoja S, Tatlisumak T, Lindsberg PJ, Kaste M. Reducing in-hospital delay to 20 minutes in stroke thrombolysis. Neurology 2012 Jul 24;79(4):306-313. [doi: 10.1212/WNL.0b013e31825d6011] [Medline: 22622858]
26. Dietrich M, Walter S, Ragoschke-Schumm A, Helwig S, Levine S, Balucani C, et al. Is prehospital treatment of acute stroke too expensive? An economic evaluation based on the first trial. Cerebrovasc Dis 2014;38(6):457-463. [doi: 10.1159/000371427] [Medline: 25531507]
27. Gyrd-Hansen D, Olsen KR, Bollweg K, Kronborg C, Ebinger M, Audebert HJ. Cost-effectiveness estimate of prehospital thrombolysis: results of the PHANTOM-S study. Neurology 2015 Mar 17;84(11):1090-1097. [doi: 10.1212/WNL.0000000000001366] [Medline: 25672925]
28. Audebert HJ, Schwamm L. Telestroke: scientific results. Cerebrovasc Dis 2009;27 Suppl 4:15-20. [doi: 10.1159/000213054] [Medline: 19546537]
29. Arch AE, Weisman DC, Coca S, Nystrom KV, Wira CR, Schindler JL. Missed ischemic stroke diagnosis in the emergency department by emergency medicine and neurology services. Stroke 2016 Mar;47(3):668-673 [FREE Full text] [doi: 10.1161/STROKEAHA.115.010613] [Medline: 26846858]
30. Caceres JA, Adil MM, Jadhav V, Chaudhry SA, Pawar S, Rodriguez GJ, et al. Diagnosis of stroke by emergency medical dispatchers and its impact on the prehospital care of patients. J Stroke Cerebrovasc Dis 2013 Nov;22(8):e610-e614. [doi: 10.1016/j.jstrokecerebrovasdis.2013.07.039] [Medline: 24075587]
31. Taqui A, Cerejo R, Itrat A, Briggs FBS, Reimer AP, Winners S, Cleveland Pre-Hospital Acute Stroke Treatment (PHAST) Group. Reduction in time to treatment in prehospital telemedicine evaluation and thrombolysis. Neurology 2017 Apr 04;88(14):1305-1312. [doi: 10.1212/WNL.0000000000003786] [Medline: 28275084]
32. Berkhemer OA, Fransen PSS, Beumer D, van den Berg LA, Lingsma HF, Yoo AJ, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. N Engl J Med 2015 Jan 01;372(1):11-20. [doi: 10.1056/NEJMoa1411587] [Medline: 25517348]
33. Saver JL, Goyal M, Bonafe A, Diener H, Levy EI, Pereira VM, et al. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. N Engl J Med 2015 Jun 11;372(24):2285-2295. [doi: 10.1056/NEJMoa1415061] [Medline: 25882376]
34. Jovin TG, Chamorro A, Cobo E, Molina CA, Rovira A, Serena J, et al. Thrombectomy within 8 hours after symptom onset in ischemic stroke. N Engl J Med 2015 Jun 11;372(24):2296-2306. [doi: 10.1056/NEJMoa1503780] [Medline: 25882510]
35. Mohamad NF, Hastrup S, Rasmussen M, Andersen MS, Johnsen SP, Andersen G, et al. Bypassing primary stroke centre reduces delay and improves outcomes for patients with large vessel occlusion. European Stroke Journal 2016 Jun;1(2):85-92. [doi: 10.1177/2396987316647857]
36. Yamada KC, Inoue S, Sakamoto Y. An effective support system of emergency medical services with tablet computers. JMIR Mhealth Uhealth 2015 Feb 27;3(1):e23 [FREE Full text] [doi: 10.2196/mhealth.3293] [Medline: 25803096]

Abbreviations

CT: computed tomography
EVT: endovascular treatment
ICER: incremental cost-effectiveness ratio
INNOVIRIS: Brussels Institute for Research and Innovation
IVT: intravenous administration of recombinant tissue plasminogen activator
KCE: Belgian Health Care Knowledge Centre
mRS: modified Rankin Scale
MSU: mobile stroke unit
OTT: onset to treatment time
QALY: quality-adjusted life-year
TIA: transient ischemic attack
UZ Brussel: Universitair Ziekenhuis Brussel
VUB: Free University of Brussels

Edited by G Eysenbach; submitted 26.06.17; peer-reviewed by K Bache, K Fassbender, S Walter; comments to author 03.08.17; revised version received 27.09.17; accepted 03.10.17; published 24.11.17.

Please cite as:
Valenzuela Espinoza A, Devos S, van Hooff RJ, Fobelets M, Dupont A, Moens M, Hubloue I, Lauwaert D, Cornu P, Brouns R, Putman K
Time Gain Needed for In-Ambulance Telemedicine: Cost-Utility Model
JMIR Mhealth Uhealth 2017;5(11):e175
URL: http://mhealth.jmir.org/2017/11/e175/
doi:10.2196/mhealth.8288
PMID:29175808

Original Paper

Recruitment and Ongoing Engagement in a UK Smartphone Study Examining the Association Between Weather and Pain: Cohort Study

Katie L Druce[1], PhD; John McBeth[1,2], PhD; Sabine N van der Veer[1], PhD; David A Selby[3], BSc (Hons); Bertie Vidgen[4], MA; Konstantinos Georgatzis[5], PhD; Bruce Hellman[6], MBA; Rashmi Lakshminarayana[6], MBA, MPH, MD; Afiqul Chowdhury[1]; David M Schultz[7], PhD; Caroline Sanders[8], PhD; Jamie C Sergeant[1,2], PhD; William G Dixon[1,2], MRCP, PhD

[1]Arthritis Research UK Centre for Epidemiology, University of Manchester, Manchester, United Kingdom

[2]NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, Manchester, United Kingdom

[3]Department of Statistics, University of Warwick, Coventry, United Kingdom

[4]Oxford Internet Institute, University of Oxford, Oxford, United Kingdom

[5]School of Informatics, University of Edinburgh, Edinburgh, United Kingdom

[6]uMotif, London, United Kingdom

[7]Centre for Atmospheric Science, School of Earth and Environmental Sciences, University of Manchester, Manchester, United Kingdom

[8]Medical Sociology, Division of Population Health, Health Services Research and Primary Care, University of Manchester, Manchester, United Kingdom

Corresponding Author:
Katie L Druce, PhD
Arthritis Research UK Centre for Epidemiology
University of Manchester
Stopford Building
Oxford Road
Manchester, M13 9PT
United Kingdom
Phone: 44 161 275 1604
Email: katie.druce@manchester.ac.uk

Abstract

Background: The huge increase in smartphone use heralds an enormous opportunity for epidemiology research, but there is limited evidence regarding long-term engagement and attrition in mobile health (mHealth) studies.

Objective: The objective of this study was to examine how representative the Cloudy with a Chance of Pain study population is of wider chronic-pain populations and to explore patterns of engagement among participants during the first 6 months of the study.

Methods: Participants in the United Kingdom who had chronic pain (≥3 months) and enrolled between January 20, 2016 and January 29, 2016 were eligible if they were aged ≥17 years and used the study app to report any of 10 pain-related symptoms during the study period. Participant characteristics were compared with data from the Health Survey for England (HSE) 2011. Distinct clusters of engagement over time were determined using first-order hidden Markov models, and participant characteristics were compared between the clusters.

Results: Compared with the data from the HSE, our sample comprised a higher proportion of women (80.51%, 5129/6370 vs 55.61%, 4782/8599) and fewer persons at the extremes of age (16-34 and 75+). Four clusters of engagement were identified: high (13.60%, 865/6370), moderate (21.76%, 1384/6370), low (39.35%, 2503/6370), and tourists (25.44%, 1618/6370), between which median days of data entry ranged from 1 (interquartile range; IQR: 1-1; tourist) to 149 (124-163; high). Those in the high-engagement cluster were typically older, whereas those in the tourist cluster were mostly male. Few other differences distinguished the clusters.

Conclusions: Cloudy with a Chance of Pain demonstrates a rapid and successful recruitment of a large, representative, and engaged sample of people with chronic pain and provides strong evidence to suggest that smartphones could provide a viable alternative to traditional data collection methods.

(JMIR Mhealth Uhealth 2017;5(11):e168) doi:10.2196/mhealth.8162

KEYWORDS

epidemiology; mHealth; chronic pain; methods

Introduction

In the United Kingdom, 70% of adults own a smartphone, over half of whom use apps [1]. This growth in smartphone use within the general population heralds an enormous opportunity for epidemiology and population-health research [2-4], allowing data collection to be integrated into people's lives. Smartphone apps for health monitoring can potentially deliver frequent and regular self-reported symptoms, whereas sensors on smartphones can aid collection of new data types, including position, movement, and environmental exposures [5].

Despite high expectations about mobile health (or mHealth) [4] studies and initial evidence that mHealth studies can recruit at scale [5], limited evidence exists on representativeness of populations who participate in digital health studies and patterns of engagement over time [6-8]. This is particularly pertinent, given the known existence of both primary and secondary digital divides, in which younger adults from higher socioeconomic backgrounds are not only more likely to have access to a smartphone device but will also utilize them differently from older adults [1,9]. Thus, though younger adults are more likely to download apps and play games on their devices, older users primarily view their smartphone as a means of communication [1].

Although smartphones appear to offer a more rapid and mobile method of data collection without compromising completion rates obtained by traditional methods [10,11], relatively little detailed information is available regarding participant recruitment and retention, or engagement, in smartphone studies, particularly when compared with other traditional methods [12] or Web-based studies [13]. Engagement has previously been defined in ways which fail to account for the potentially variable patterns of use through time, including continuity of data entry [5,14-16], and this nonuniformity in definitions makes it difficult to draw conclusions regarding the viability of mHealth studies for longitudinal research.

Cloudy with a Chance of Pain is a UK smartphone-based, prospective cohort study investigating the link between the weather and pain in people with chronic pain. Specifically, Cloudy with a Chance of Pain seeks to investigate whether self-reported pain severity is associated with weather variables and whether the observed relationships differ between specific patient groups. Earlier research on this topic has been inconclusive [17], despite more than two-thirds of patients with musculoskeletal pain believing that there is an association between the weather and pain [18,19]. The numerous methodological challenges that have traditionally contributed to this ambiguity include small sample sizes, a lack of temporally rich data, and poor availability of data pertaining to geographical and meteorological variability. However, smartphone apps have the capacity to overcome these challenges, if they can recruit and continue to engage a representative study population.

The two aims of this paper were to examine how representative the Cloudy with a Chance of Pain study population is of wider chronic-pain populations and to explore patterns of engagement among participants during the first 6 months of the study.

Methods

From January 20, 2016 to January 20, 2017, Cloudy with a Chance of Pain aimed to recruit over 1000 UK residents aged 17 or over who owned an Android or iPhone operating system (iOS; Apple Inc) smartphone, and who experienced pain for at least the preceding 3 months. The study was advertised through national and regional television, radio and newspaper media, social media, and via charity and patient partner organizations (Multimedia Appendix 1). Further information for interested participants was available on the study website [20].

To enroll in the study, participants downloaded the uMotif app [21] on their smartphone from the Apple App Store or Google Play Store. After completion of digital consent, the app enabled participants to report their symptoms daily for 6 months, or longer if willing. In the background, the smartphone's Global Positioning System (GPS) reported hourly location, allowing linkage to local weather data from the Met Office (the UK's national weather service) and investigation of the association between weather and pain. More details on the app and data collection are provided below.

Participants included in this analysis were those recruited between January 20, 2016 and February 29, 2016, with patterns of engagement examined through to July 20, 2016, 6 months from the study launch date. Participants provided a year of birth through the consent process to confirm that they were 17 years of age or older. Not everyone who downloaded the app used it, so eligibility was further restricted to those who had reported their symptoms at least once between enrollment and July 20, 2016.

Ethical approval was obtained in December 2015 from the University of Manchester Research Ethics Committee 4 (ref: ethics/15522).

Data

Baseline Data

The baseline questionnaire collected demographic data: sex, year of birth, and first half of participant's postcode. Participants

reported the site of pain (eg, head, face, knee) and were able to report pain at multiple sites or having pain all over the body. Participants were asked to record whether they had been diagnosed (by a doctor) with rheumatoid arthritis, ankylosing spondylitis or spondyloarthropathy, gout or other calcium-crystal arthritis (eg, pseudogout), arthritis (type not specified), fibromyalgia or chronic widespread pain, chronic headache, or neuropathic pain. A free-text entry box was provided for any diagnoses not otherwise listed. Due to a coding error, diagnoses of osteoarthritis (OA) were not collected for the first 9 weeks of data collection, after which it was included within the above list. A push notification was sent out on March 24, 2016 asking existing participants to indicate whether or not they had the condition. Responses were received from 1157 of 8267 (13.99%) of participants recruited by March 24, 2016. For this reason, prevalence rates of OA are not provided in this paper.

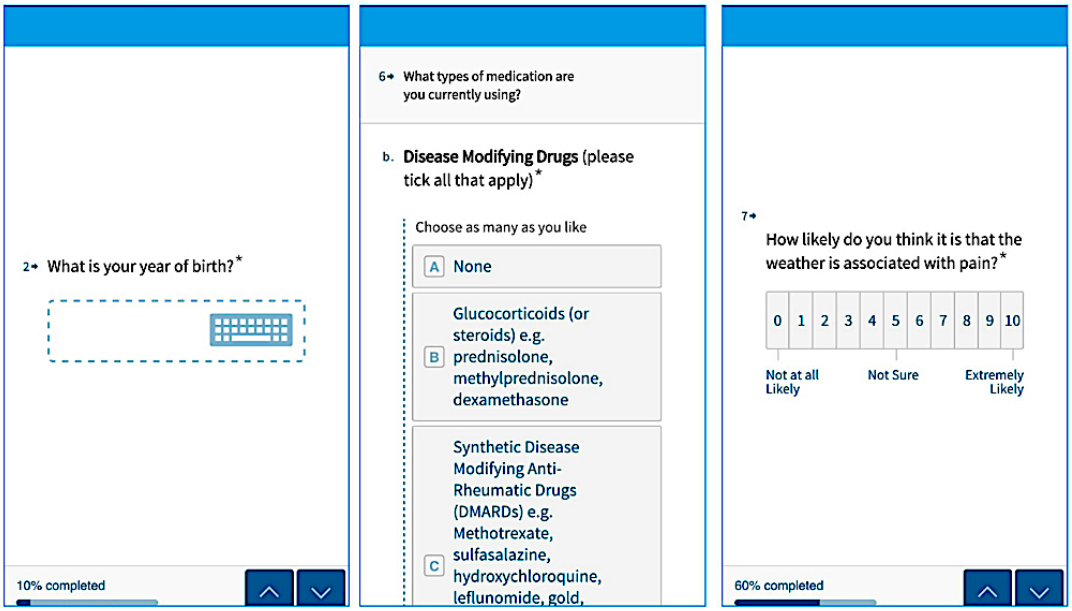
Participants reported their use of paracetamol, nonsteroidal anti-inflammatory drugs (NSAIDs), simple analgesics, weak opiates, strong opiates, and drugs for neuropathic pain. Participants reported their use of glucocorticoids (steroids), synthetic disease modifying antirheumatic drugs (DMARDs), and biologic DMARDS. Participants could also report the use of other medications. If "other" was selected, a free-text entry box was provided.

Participants reported how likely they thought it was that the weather was associated with pain using a 0 to 10 numerical rating scale (NRS), where 0 indicated not at all likely and 10 indicated extremely likely. Participants were also asked which weather conditions they most felt were associated with pain, selecting from damp/rain, cold, heat, change in barometric pressure, change in temperature, and other (free-text box provided to specify belief). Examples of data-entry screens are shown in Figure 1.

Daily Symptom Domains

Following completion of the baseline questionnaire, participants were asked to report 10 symptoms every day using the uMotif app (Figure 2), prompted by a daily notification at 6:24 p.m. Each symptom was scored in five ordinal categories (eg, pain was scored as no pain, mild, moderate, severe, or very severe). The symptoms were pain severity, fatigue, morning stiffness, the impact of pain on activities, sleep quality, time spent outside, feeling tired on waking, physical activity, mood, and well-being. A study motif was considered complete when all 10 variables were reported at a single time point. The app was codesigned with a patient and public involvement group and refined after a feasibility study of 20 participants with rheumatoid arthritis [22].

Figure 1. Screenshot of example baseline data collection.



Analysis

Representativeness of Participants

To explore the representativeness of participants recruited to this study, we compared the age and sex distribution of participants with that of a sample of persons with chronic pain (≥3 months) from the Health Survey for England (2011) [23]. The Health Survey for England is a large-scale annual survey that has been conducted since 1994 and recruits a stratified random probability sample of private households within England. Full description of the methods of data collection are available elsewhere [24].

Engagement

We sought to define common patterns of engagement (ie, data entry), using a three-step process.

Following recruitment, individuals were labeled as engaged if they reported any of the ten symptoms on a given day. A first-order hidden Markov model [25,26] was then used to estimate the levels of engagement of participants, using the depmixS4 R package (I Visser, Netherlands)[27] (Multimedia Appendix 2). The model assumed three latent engagement states: high, low, and disengaged. The model was initialized assuming every participant started highly engaged. Furthermore, the model assumed that disengagement was an "absorbing state," so that

participants entering this state could not reengage with the study. Finally, clusters were defined according to different probabilities of transitioning between high engagement, low engagement, and disengagement during the study. The optimal number of clusters between 2 and 10 was identified visually using the "elbow method" [28]. The elbow method involves plotting the curve of log-likelihood against number of clusters, such that the location of a bend ("elbow") in the plot is considered to identify the best number of clusters. The clusters were generated by a "blind" algorithmic process. Therefore, to assign names to the clusters, the engagement patterns of a random selection of users within each cluster were inspected.

Comparisons were then made between the clusters regarding duration of study engagement, defined as (1) the median number of days "in study" (defined as the number of days from first to last symptoms report) and (2) the median number of days of data entry (defined as a day when any symptoms were reported). Data completion was compared between the clusters, defined as (1) the total number of segments reported, (2) the total number of complete motifs, (3) the proportion of days in the study (days between enrollment and July 20, 2016) on which complete motifs were reported (days of data entry/total days in study), and (4) the proportion of days of data entry on which complete motifs were reported. Baseline data were then compared between the clusters, with data presented as median and interquartile range (IQR), or proportion and 95% CI where appropriate. Due to the initial configuration of the app, data regarding the mobile-phone platform used by participants are not available for all participants, and we are unable to compare or draw conclusions about how app use differs between Apple and Android platforms.

Figure 2. Screenshot of motif for daily symptom collection.

Results

Of 7972 participants enrolled in the study between January 20, 2016 and February 29, 2016, 6370 (79.90%) were eligible for the analysis in this paper (Table 1). Reasons for failing eligibility included no baseline data (n=802), age indeterminate (n=308) and age <17 (n=3). A further 489 participants had downloaded the app but never reported symptoms. Those who installed the app but did not prospectively record symptoms did not differ from those who recorded symptoms based on age (median 51; IQR 41-61 vs 49; IQR 41-59) or strength of belief in the association between the weather and pain (median 7; IQR 5-9 vs 7; IQR 6-9). However, a larger proportion were male (30.3%, 95% CI 26.2-34.4 vs 19.5, 18.5-20.5).

Eligible participants were 80.51% (5129/6370) female, with a mean age of 49 years. The majority of those included in the analysis reported pain at more than one site (73.39%, 4675/6370). A further 16.62% (1059/6370) reported pain "all over" and 9.49% (605/6370) reported pain at a single site. The most common diagnosis was arthritis (40.29% type unspecified [2567/6370], 19.12% rheumatoid arthritis [1218/6370]), followed by fibromyalgia/chronic widespread pain (23.75% , 1513/6370) and "other pain diagnosis" (22.64%, 1442/6370). Beliefs about the existence of a relationship between the weather and pain were strong, with a median belief score of 7 (IQR: 6-9). Participants most commonly believed that pain was affected by the damp/rain (74.43% , 4741/6370) and the cold (68.67%, 4374/6370) but least commonly believed that hot weather affected pain (14.76%, 940/6370).

Table 1. Baseline characteristics of participants eligible for analysis.

Characteristics	All eligible participants, n (% or SD or IQR[a]) (n=6370)
Demographics	
Female	5129 (80.52%)
Mean age in years	49.2 (12.9)
Pain condition	
Site of pain	
Single	605 (9.50%)
Multisite	4675 (73. 39%)
All over pain	1059 (16.62%)
Missing	31 (0.49%)
Diagnosis of conditions	
Rheumatoid arthritis	1218 (19.12%)
Ankylosing spondylitis/spondyloarthropathy	576 (9.04%)
Gout	231 (3.63%)
Arthritis (type not specified)	2567 (40.29%)
Fibromyalgia/chronic widespread pain	1513 (23.79%)
Chronic headache	462 (7.25%)
Neuropathic	821 (12.89%)
Other	1442 (22.64%)
Medications used at baseline	
Analgesics	
None	619 (9.72%)
Paracetamol	3154 (49.51%)
Nonsteroidal anti-inflammatory drugs	3694 (57.99%)
Simple analgesics	1937 (30.41%)
Weak opiates	1902 (29.86%)
Strong opiates	782 (12.28%)
Neuropathic pain medication	1297 (20.36%)
Other pain medications	717 (11.26%)
Disease modifying treatment	
None	4407 (69.18%)
Steroids	480 (7.54%)
Synthetic DMARDs[b]	1282 (20.13%)
Biologic DMARDs	560 (8.79%)
Other DMARDs	406 (6.37%)
Beliefs	
Median strength of belief in the association between weather and pain	7 (6-9)
Weather conditions that participants think most affect their pain	
Damp or rain	4741 (74.43%)
Cold	4374 (68.67%)
Hot	940 (14.76%)
Changes in barometric pressure	1945 (30.53%)

Characteristics	All eligible participants, n (% or SD or IQR[a]) (n=6370)
Changes in temperature	1967 (30.88%)

[a]IQR: interquartile range.

[b]DMARDs: disease-modifying antirheumatic drugs.

Comparison With Other Chronic Pain Populations

Compared with data from the Health Survey for England (2011) [23], a greater proportion of participants in this study were women (80.52%, 5129/6370 compared with 55.61%, 4782/8599 expected). The age bands 35 to 64 years were over-represented in this study (73.11%, 4657/6370 compared with 51.18%, 4401/8599 expected; Table 2), with fewer participants in the extremes of age (<35: 14.65%, 933/6370 compared with 24.7%, 2126/8599 expected; ≥75: 1.19%, 76/6370 compared with 11.13%, 957/8599 expected).

Identifying Clusters of Engagement

Following inspection of the log-likelihood plot (Figure 3), a four-cluster solution was retained. The clusters (Figure 4) were allocated names based on the best description of their engagement patterns: high engagement (14%, 865/6370; red), moderate engagement (22%, 1384/6370; purple), low engagement (39%, 2503/6370; green), and tourists (25%, 1618/6370; teal).

The proportion of days on which data were entered and rates of data completion varied substantially between clusters (Table 3).The median days "in study" ranged from 175 days (IQR: 152-177) in the high-engagement cluster to 1 day (IQR: 1-1) in the tourist cluster. Participants in the moderate-engagement cluster stayed in the study 10 times longer than those in the low-engagement cluster (88 days, 42-163 vs 8 days, 4-16).

Those in the high-engagement cluster provided data on most days throughout follow-up (Figure 4). The high-engagement cluster reported complete motifs on 89.13% (106,360/119,332) of the days that they provided data, and the moderate-engagement cluster provided complete motifs on 87.5% (67,704/77,368) of all data-entry days. Rates of completion were slightly lower in the other clusters, with the low-engagement cluster and tourists recording complete motifs on 82.88% (13,415/16,186) and 64.85% (1947/2848) of the days on which any data were reported, respectively (Table 3).

Table 2. Comparison of the sex and age distribution of persons with chronic pain from the Health Survey for England (2011) and participants recruited to Cloudy with a Chance of Pain.

Population	Sex		Age (in bands), years					
	Male, n (%)	Female, n (%)	16-34, n (%)	35-44, n (%)	45-54, n (%)	55-64, n (%)	65-74, n (%)	75+, n (%)
Health Survey for England (2011)	3817 (44.39)	4782 (55.61)	2126 (24.72)	1512 (17.58)	1490 (17.33)	1399 (16.27)	1115 (12.97)	957 (11.13)
Cloudy with a Chance of Pain	1231 (19.48)	5129 (80.52)	933 (14.65)	1280 (20.09)	1840 (28.89)	1537 (24.13)	704 (11.05)	76 (1.19)

Figure 3. Plot of the log-likelihood of different numbers of clusters in hidden Markov sequences; the elbow indicates the optimal number of clusters which should be accepted.

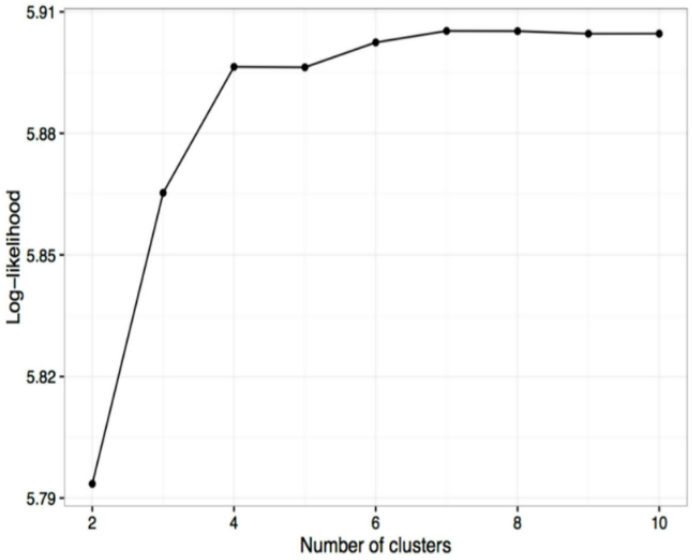


Table 3. Data provided by 6370 Cloudy with a Chance of Pain participants clustered by levels of engagement.

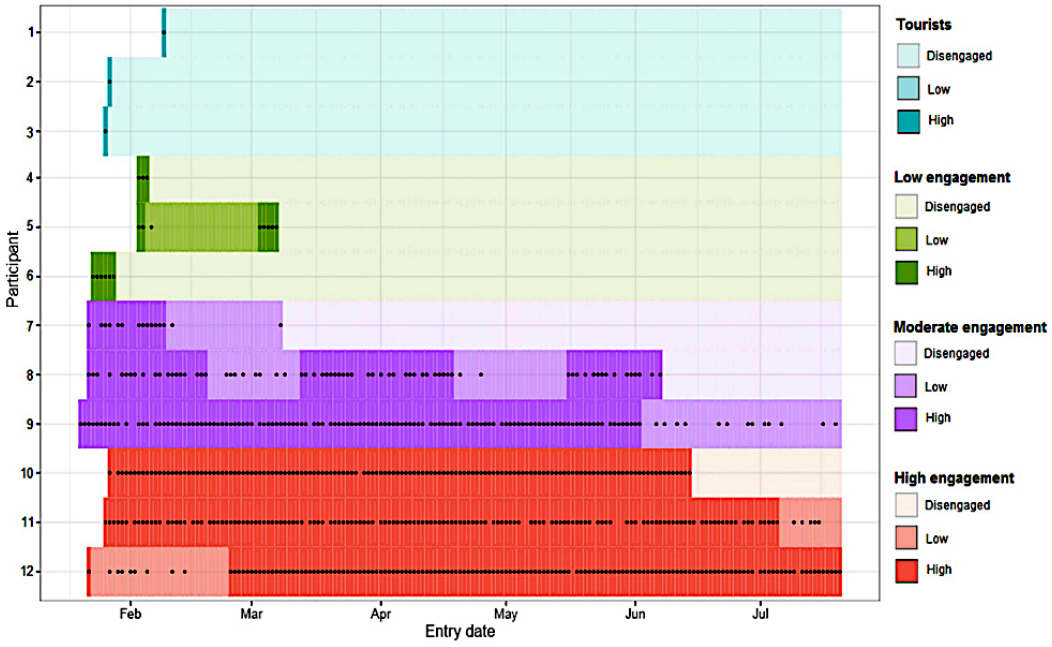
Data	High	Moderate	Low	Tourist
Participants in cluster, n (%)	865 (13.6)	1384 (21.7)	2503 (39.3)	1618 (25.4)
Total number of segments reported	1,233,685	799,872	171,545	26,344
Total number of complete motifs	106,360	67,704	13,415	1847
Total number of days in study	151,187	240,841	435,678	279,755
Total number of days of data entry	119,332	77,368	16,186	2848
Median number of days in study[a]	175 (152-177)	88 (42-163)	8 (4-16)	1 (1-1)
Median number of days of data entry[b] (IQR[c])	149 (124-163)	44 (27-80.5)	4 (3-9)	1 (1-1)
Proportion (%) of days in study on which complete motifs were reported	(70.16)	(28.11)	(3.08)	(0.66)
Proportion (%) of days of data entry on which complete motifs were reported	(89.13)	(87.51)	(82.88)	(64.85)

[a]Days between first and final symptom report.

[b]Data entry: any symptom reported.

[c]IQR: interquartile range.

Figure 4. Examples of participants from clusters; High engagement (red), Moderate engagement (purple), Low engagement (green), Tourists (teal).



Between-Cluster Differences

Higher engagement was associated with increased age, with a difference of more than 5 years between the median age of those who were in the low-engagement (47, IQR: 39-57), or tourist clusters (49, IQR: 40-58), and those who were in the high-engagement cluster (median 56 years, IQR: 47-63). A substantially lower proportion of those in the tourist cluster were women (76.27%, 1234/1618; 95% CI 74.2-78.3) than any other cluster (high engagement: 82.31%, 712/865; 95% CI 79.6-84.7; moderate engagement: 84.10%, 1164/1384; 95% CI 82.1-85.9; low engagement: 80.66%, 2019/2503; 95% CI 19.1-82.2).

There were no differences between clusters with respect to the site of pain or in the prevalence of rheumatic disease diagnoses (eg, rheumatoid arthritis, fibromyalgia). The proportion of people in the tourist cluster (17.74%, 287/1618; 95% CI 15.88-19.60) who reported "other" pain conditions was also lower than in the high-engagement (23.70%, 205/865; 95% CI 20.87-26.55), moderate-engagement (24.49%, 339/1384; 95% CI 22.22-26.76), and low-engagement (24.41%, 611/2503; 95% CI 22.73-26.09) groups.

Table 4. Characteristics of the 6370 Cloudy with a Chance of Pain participants clustered by levels of engagement.

Data	High (n=865)	Moderate (n=1384)		Low (n=2503)		Tourist (n=1618)	
	n (% or IQR[a])	95% CI	n (% or IQR)	95% CI	n (% or IQR)	95% CI	n (% or IQR)	95% CI
Demographics								
Female	712 (82.31%)	79.77-84.85	1164 (84.10%)	82.17-86.03	2019 (80.66%)	79.11-82.21	1234 (76.27%)	74.20-78.34
Median age in years	56 (47-63)		50 (41-59)		47 (39-57)		49 (40-58)	
Pain condition								
Site of pain								
Single	71 (8.21%)	6.38-10.04	121 (8.74%)	7.25-10.23	226 (9.03%)	7.91-10.15	187 (11.56%)	10.00-13.12
Multisite	668 (77.23%)	74.44-80.02	1026 (74.13%)	71.82-76.44	1809 (72.27%)	70.52-74.02	1172 (72.44%)	70.26-74.62
All over pain	121 (13.99%)	11.68-16.30	230 (16.62%)	14.66-18.58	462 (18.46%)	16.94-19.98	246 (15.20%)	13.45-16.95
Missing	5 (0.58%)	0.07-1.09	7 (0.51%)	0.13-0.89	6 (0.24%)	0.05-0.43	13 (0.80%)	0.37-1.23
Diagnosis of conditions								
Rheumatoid arthritis	176 (20.35%)	17.67-23.03	271 (19.58%)	17.49-21.67	473 (18.90%)	17.37-20.43	298 (18.42%)	16.53-20.31
Ankylosing spondylitis/spondyloarthropathy	70 (8.09%)	6.27-9.91	130 (9.39%)	7.785-10.93	230 (9.19%)	8.06-10.32	146 (9.02%)	7.62-10.42
Gout	29 (3.35%)	2.15-4.55	55 (3.97%)	2.94-5.00	80 (3.20%)	2.51-3.89	67 (4.14%)	3.17-5.11
Arthritis (unspecified)	394 (45.55%)	42.23-48.87	556 (40.17%)	37.59-42.75	958 (38.27%)	36.37-40.17	659 (40.73%)	38.34-43.12
Fibromyalgia/chronic widespread pain	188 (21.73%)	18.98-24.48	336 (24.28%)	22.02-26.54	634 (25.33%)	23.63-27.03	355 (21.94%)	19.92-23.96
Chronic headache	43 (4.97%)	3.52-6.42	100 (7.23%)	5.87-8.59	209 (8.35%)	7.27-9.43	110 (6.80%)	5.57-8.03
Neuropathic	112 (12.95%)	10.71-15.19	182 (13.15%)	11.37-14.93	337 (13.46%)	12.12-14.80	190 (11.74%)	10.17-13.31
Other	205 (23.70%)	20.87-26.53	339 (24.49%)	22.22-26.76	611 (24.41%)	22.73-26.09	287 (17.74%)	15.88-19.60
Medications used at baseline								
Analgesics								
None	81 (9.36%)	7.42-11.30	115 (8.31%)	6.86-9.76	236 (9.43%)	8.29-10.57	187 (11.56%)	10.00-13.12
Paracetamol	454 (52.49%)	49.16-55.82	707 (51.08%)	48.45-53.71	1241 (49.58%)	47.62-51.54	752 (46.48%)	44.05-48.91
Nonsteroidal anti-inflammatory drugs	498 (57.57%)	54.28-60.86	833 (60.19%)	57.61-62.77	1470 (58.73%)	56.80-60.66	893 (55.19%)	52.77-57.61
Simple analgesics	254 (29.36%)	26.33-32.39	406 (29.34%)	26.94-31.74	773 (30.88%)	29.07-32.69	504 (31.15%)	28.89-33.41
Weak opiates	253 (29.25%)	26.22-32.28	426 (30.78%)	28.35-33.21	773 (30.88%)	29.07-32.69	450 (27.81%)	25.63-29.99
Strong opiates	82 (9.48%)	7.53-11.43	154 (11.13%)	9.47-12.79	356 (14.22%)	12.85-15.59	190 (11.74%)	10.17-13.31
Neuropathic pain medication	167 (19.31%)	16.68-21.94	278 (20.09%)	17.98-22.20	538 (21.49%)	19.88-23.10	314 (19.41%)	17.48-21.34

Data	High (n=865)	Moderate (n=1384)		Low (n=2503)		Tourist (n=1618)	
	n (% or IQR[a])	95% CI	n (% or IQR)	95% CI	n (% or IQR)	95% CI	n (% or IQR)	95% CI
Other pain medications	106 (12.25%)	10.07-14.43	188 (13.58%)	11.78-15.38	270 (10.79%)	9.57-12.01	153 (9.46%)	8.03-10.89
Steroids	57 (6.59%)	4.94-8.24	96 (6.94%)	5.60-8.28	202 (8.07%)	7.00-9.14	125 (7.73%)	6.43-9.03
DMARDs[b]								
None	591 (68.32%)	65.22-71.42	974 (70.38%)	67.97-72.79	1690 (67.52%)	65.69-69.35	1152 (71.20%)	68.99-73.41
Synthetic DMARDs	192 (22.20%)	19.43-24.97	292 (21.10%)	18.95-23.25	529 (21.13%)	19.53-22.73	269 (16.63%)	14.82-18.44
Biologic DMARDs	80 (9.25%)	7.32-11.18	121 (8.74%)	7.25-10.23	226 (9.03%)	7.91-10.15	133 (8.22%)	6.88-9.56
Other DMARDs	58 (6.71%)	5.04-8.38	73 (5.27%)	4.09-6.45	156 (6.23%)	5.28-7.18	119 (7.35%)	6.08-8.62
Beliefs								
Median strength of belief in the association between weather and pain	7 (6-9)		7 (6-9)		7 (6-9)		7 (5-9)	
Weather condition(s) that participants think most affect their pain								
Damp or rain	647 (74.80%)	71.91-77.69	1030 (74.42%)	72.12-76.72	1883 (75.23%)	73.54-76.92	1181 (72.99%)	70.83-75.15
Cold	539 (62.31%)	59.08-65.54	931 (67.27%)	64.80-70.4	1799 (71.87%)	70.11-73.63	1105 (68.29%)	66.02-70.56
Hot	117 (13.53%)	11.25-15.81	210 (15.17%)	13.28-17.06	383 (15.30%)	13.89-16.71	230 (14.22%)	12.52-15.92
Changes in barometric pressure	307 (35.49%)	32.30-38.68	455 (32.88%)	30.41-35.35	714 (28.53%)	26.76-30.30	469 (28.99%)	26.78-31.20
Changes in temperature	238 (27.51%)	24.53-30.49	422 (30.49%)	28.06-32.92	797 (31.84%)	30.01-33.67	510 (31.52%)	29.26-33.78

[a]IQR: interquartile range.

[b]DMARDs: disease-modifying antirheumatic drugs.

No differences were observed between the clusters regarding the use of analgesics and steroids. Only the use of synthetic DMARDs differed substantially between the clusters, with less of those in the tourist cluster (16.63%, 269/1618; 95% CI 14.82-18.44) reporting taking the medication than those in the other engagement clusters (high engagement: 22.20%, 192/865; 95% CI 19.43-24.97; moderate engagement: 21.10%, 292/1384; 95% CI 18.95-23.25; low engagement: 21.13%, 529/2503; 95% CI 19.53-22.73). Comparable proportions were using biologic or other DMARDs.

There were no differences in the strength of belief that the weather affected pain, but fewer of those in the high-engagement cluster believed the cold affected their pain (62.31%, 539/865; 95% CI 59.08-65.54) when compared with those in the low-engagement and tourist clusters (71.87%,1799/2503; 95% CI 70.11-73.63 and 68.29%, 1105/1618; 95% CI 66.02-70.56, respectively). Conversely, more of those who were highly engaged (35.49%, 307/865; 95% CI 32.30-38.68) believed that changes in barometric pressure were associated with pain that those in the low-engagement and tourist clusters (28.35%, 714/2503; 95% CI 26.76-30.30 and 28.99%, 469/1618; 95% CI 26.78-31.20, respectively). There were no observed differences in the proportion of participants who believed their pain is associated with damp or rain, heat, or changes in temperature (Table 4).

Discussion

Principal Findings

Cloudy with a Chance of Pain is the first mHealth study to demonstrate successful and rapid mass recruitment of a largely representative sample of highly engaged participants. Among our sample, patterns of ongoing engagement showed that around 1 in 7 participants provided data on most days in the first 6 months, completing full data entry on 89% of those days.

A major strength of Cloudy with a Chance of Pain is the rapid mass recruitment of eligible participants. Our study benefitted from wide promotion by the UK national media at the time of

the study launch, which emphasizes the power of national media to promote. Indeed, as a result of coverage including, among others, the BBC2 television show *Trust Me I'm a Doctor* on January 20, 2016 and BBC *Breakfast* on January 26, 2016, 90% of participants enrolled in the study by July 20 were recruited within 1 month of the study launch.

Furthermore, ongoing engagement within Cloudy with a Chance of Pain was high. More than 30% of participants were in the high-engagement or moderate-engagement cluster, entering data on at least half of days throughout the 6 months. In comparison, fewer than 25% of participants in Apple's ResearchKit studies were active by 10 weeks [29], with similar proportions active in a physical-activity study by 42 days [14]. In one of the largest mHealth studies reported to date (mPower study of people with Parkinson disease and healthy controls), less than 10% of enrolled participants completed 5 or more days within the first 6 months of the study [5]. One in 7 participants were in the high-engagement cluster and provided data on most days throughout the 6 months; we are not aware of other mHealth studies that have reported such high levels of ongoing engagement to date.

Previous analyses have used arbitrary definitions that fail to capture the patterns of use through time and may ignore the importance of continuity of data entry [5,14-16]. In contrast, this analysis attempted to account fully for data complexity and made no a priori decisions to define engagement. Thus, this study has improved understanding of the extent to which participants remain engaged over time and provides a promising method for future engagement studies.

Our recruitment strategy enrolled a sample which comprised an under-representation of males and persons at the extremes of age (<35 years and ≥75 years) than would have been expected from the general population data of the Health Survey for England (2011) [23]. Although women are more likely to respond to more traditional population surveys [30-33], we recruited a much higher proportion of women than would have been expected using traditional recruitment methods. A possible explanation is that women recruited to this study more commonly viewed the television programs or may have perceived the potential for additional benefits for participation. However, we also note that self-selection likely accounts for the observed differences in our population. For example, not only are women known to use social media [34] and health apps more than men [35], but they also use digital content differently [36,37].

Therefore, future mHealth studies may benefit from the use of supplementary and targeted recruitment strategies used by other digital health interventions [38] in which it would be possible to oversample men, such as the use of health professionals, friends, and families, or work-based campaigns, as well as outreach programs designed to access hard-to-reach groups. Although similar methods could also promote the recruitment of younger adults, other opportunities to promote participation among this group include the use of social networks, community components, and app gamification [39-41].

Nevertheless, the internal validity of the study results (ie, the relationship between the weather and pain within our sample) is unlikely to be influenced by the excess of women entering the study, as there is no reason to suspect the relationship differs by sex. Analysis of the relationship between the weather and pain, and whether the relationship differs between the sexes, is underway and will be reported separately.

The impact on external validity (ie, the generalizability) is unclear, as people with a particular belief may have been more inclined to participate (which may, in turn, differ by sex). That said, our findings about beliefs align with prior research that suggests that as many as 92% of patients with arthritis believe in an association between weather and pain [42].

The reasons for the unprecedented rates of engagement observed in this study are worth exploring, particularly as we sought to collect a large amount of daily data, and this burden on the participants might well have been expected to result in a higher loss to follow-up through time and over such a long period. This study found that older participants were more likely to remain engaged in the study. One possible explanation for this is that older persons are less likely to use smartphone apps [1] and therefore may be less likely to experience "app fatigue" than younger participants. They may also feel a greater responsibility to complete the ongoing data entry once registered or have more time to give to the study. Furthermore, functionalities such as geolocation consume battery power, which may have a greater impact on younger persons, who use their smartphones for a greater number of varied tasks, than on older persons [1]. We note, however, that reasons for declining engagement are likely numerous.

Earlier studies have sought to examine possible mechanisms of engagement, including the complexity of tasks [3], the time of day data are entered [43], and various functionality features such as reminders, interactivity, tailored content, and delivery of feedback [14,15]. In a feasibility study [22], we reported that key motivators for ongoing engagement were the simple graphical user interface, automated reminders for data entry, a desire to contribute to answering an understandable and engaging research question, and visualization of data. However, limited information was available in this larger study to delineate the motivators of engagement in this population. We also acknowledge that the study did not capture education and income, which would have enabled this study to investigate the potential impact of the digital divide on recruitment and engagement.

Conclusions

In summary, Cloudy with a Chance of Pain demonstrates a rapid and successful recruitment of a large and engaged sample of people with chronic pain. Although there may be selection bias toward older females in our study, younger men are also less likely to participate in studies using traditional data-collection methods. Thus, our study provides strong evidence to suggest that smartphones could provide a viable alternative to traditional data collection methods, particularly for collecting daily data over long periods.

Acknowledgments

This work was supported by Arthritis Research UK (grant number 21225); the Medical Research Council (MRC)'s Confidence in Concept Scheme (Grant number MC_PC_13070); the Arthritis Research UK Centre for Epidemiology and the Farr Institute @Health eResearch Centre (HeRC) (Grant number MR/K006665/1); MRC Clinician Scientist Award (Grant number G0902272 to W.G.D); and the National Institute for Health Research (NIHR) Manchester Musculoskeletal Biomedical Research Unit (for support of JS). This report includes independent research supported by the National Institute for Health Research Biomedical Research Unit Funding Scheme. The views expressed in this publication are those of the authors and not necessarily those of the National Health Service (NHS), the National Institute for Health Research, or the Department of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Charity and patient partner organisations who facilitated participant recruitment.

[PDF File (Adobe PDF File), 12KB - mhealth_v5i11e168_app1.pdf]

Multimedia Appendix 2

Hidden Markov model.

[PDF File (Adobe PDF File), 23KB - mhealth_v5i11e168_app2.pdf]

References

1. Ofcom. Ofcom: adults media use and attitudes report 2016 URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0026/80828/2016-adults-media-use-and-attitudes.pdf [WebCite Cache ID 6r2gB1EES]
2. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. PLoS Comput Biol 2012;8(7):e1002616 [FREE Full text] [doi: 10.1371/journal.pcbi.1002616] [Medline: 22844241]
3. Check Hayden E. Mobile-phone health apps deliver data bounty. Nature 2016 Mar 24;531(7595):422-423. [doi: 10.1038/531422a] [Medline: 27008946]
4. Who.int. 2011. mHealth: new horizons for health through mobile technologies URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [WebCite Cache ID 6r2pn463w]
5. Bot BB, 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;3:160011. [Medline: 26938265]
6. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
7. Sanders C, Rogers A, Bowen R, Bower P, Hirani S, Cartwright M, et al. Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study. BMC Health Serv Res 2012;12:220 [FREE Full text] [doi: 10.1186/1472-6963-12-220] [Medline: 22834978]
8. Rixon L, Hirani S, Cartwright M, Beynon M, Selva A, Sanders C, et al. What influences withdrawal because of rejection of telehealth - the whole system demonstrator evaluation. J Assist Technol 2013;7(4):219-227 [FREE Full text] [doi: 10.1108/JAT-06-2013-0017]
9. Latulippe K, Hamel C, Giroux D. Social health inequalities and eHealth: a literature review with qualitative synthesis of theoretical and empirical studies. J Med Internet Res 2017 Apr 27;19(4):e136 [FREE Full text] [doi: 10.2196/jmir.6731] [Medline: 28450271]
10. Buskirk TD, Andrus CH. Making mobile browser surveys smarter: results from a randomized experiment comparing online surveys completed via computer or smartphone. Field Methods 2014 Apr 14;26(4):322-342. [doi: 10.1177/1525822X14526146]
11. Dale O, Hagen KB. Despite technical problems personal digital assistants outperform pen and paper when collecting patient diary data. J Clin Epidemiol 2007 Jan;60(1):8-17. [doi: 10.1016/j.jclinepi.2006.04.005] [Medline: 17161749]
12. Marcano Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. Cochrane Database Syst Rev 2015 Jul 27(7):MR000042. [doi: 10.1002/14651858.MR000042.pub2] [Medline: 26212714]
13. Lane TS, Armin J, Gordon JS. Online recruitment methods for web-based and mobile health studies: a review of the literature. J Med Internet Res 2015;17(7):e183 [FREE Full text] [doi: 10.2196/jmir.4359] [Medline: 26202991]
14. Guertler D, Vandelanotte C, Kirwan M, Duncan MJ. Engagement and nonusage attrition with a free physical activity promotion program: the case of 10,000 steps Australia. J Med Internet Res 2015;17(7):e176 [FREE Full text] [doi: 10.2196/jmir.4339] [Medline: 26180040]

15. 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]
16. Zan S, Agboola S, Moore SA, Parks KA, Kvedar JC, Jethwani K. Patient engagement with a mobile web-based telemonitoring system for heart failure self-management: a pilot study. JMIR Mhealth Uhealth 2015;3(2):e33 [FREE Full text] [doi: 10.2196/mhealth.3789] [Medline: 25842282]
17. Smedslund G, Hagen KB. Does rain really cause pain? A systematic review of the associations between weather factors and severity of pain in people with rheumatoid arthritis. Eur J Pain 2011 Jan;15(1):5-10. [doi: 10.1016/j.ejpain.2010.05.003] [Medline: 20570193]
18. Jamison RN, Anderson KO, Slater MA. Weather changes and pain: perceived influence of local climate on pain complaint in chronic pain patients. Pain 1995 May;61(2):309-315. [Medline: 7659442]
19. Ng J, Scott D, Taneja A, Gow P, Gosai A. Weather changes and pain in rheumatology patients. APLAR J Rheumatol 2004;7(3):204-206. [doi: 10.1111/j.1479-8077.2004.00099]
20. Cloudywithachanceofpain. URL: https://www.cloudywithachanceofpain.com/ [WebCite Cache ID 6r2pz9H0X]
21. Umotif. Next-generation patient data capture platform URL: http://umotif.com/ [accessed 2017-06-07] [WebCite Cache ID 6r2q9a3B8]
22. Reade S, Spencer K, Sergeant JC, Sperrin M, Schultz DM, Ainsworth J, et al. Cloudy with a Chance of Pain: engagement and subsequent attrition of daily data entry in a smartphone pilot study tracking weather, disease severity, and physical activity in patients with rheumatoid arthritis. JMIR Mhealth Uhealth 2017 Mar 24;5(3):e37 [FREE Full text] [doi: 10.2196/mhealth.6496] [Medline: 28341616]
23. Bridges S. Digital.nhs. Health survey for England - 2011. Chronic pain URL: http://digital.nhs.uk/media/22200/Health-Survey-for-England-2011-Chapter-9-Chronic-pain/Any/HSE2011-Ch9-Chronic-Pain [accessed 2017-06-07] [WebCite Cache ID 6r2g20hQQ]
24. Boniface S, Bridges S, Craig R, Fuller E, Hancock R, Henderson C, et al. Digital.nhs. 2011. Health Survey for England 2011. Methods and documentation URL: http://digital.nhs.uk/media/22203/Health-Survey-for-England-2011-Methods-and-documentation/Any/HSE2011-Methods-and-docs [accessed 2017-06-07] [WebCite Cache ID 6r2g20hQQ]
25. Rabiner L. Ece.ucsb.edu. 1989. A tutorial on hidden Markov models and selected applications in speech recognition URL: http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/tutorial%20on%20hmm%20and%20applications.pdf [accessed 2017-10-19] [WebCite Cache ID 6uLRC1IPr]
26. Barber D. Discrete-state Markov models. In: Bayesian Reasoning and Machine Learning. Cambridge, UK: Cambridge University Press; 2012:492-508.
27. Visser I, Speekenbrink M. R-forge.r-project. 2010. depmixS4: An R-package for hidden Markov models URL: https://r-forge.r-project.org/scm/viewvc.php/*checkout*/papers/jss/V1/depmixS4.pdf?root=depmix [accessed 2017-06-07] [WebCite Cache ID 6r2psHywd]
28. Alpaydin E. Clustering. In: Introduction to Machine Learning. Cambridge, Massachusetts: MIT Press; 2004.
29. 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]
30. Flüß E, Bond CM, Jones GT, Macfarlane GJ. The effect of an internet option and single-sided printing format to increase the response rate to a population-based study: a randomized controlled trial. BMC Med Res Methodol 2014 Sep 09;14:104 [FREE Full text] [doi: 10.1186/1471-2288-14-104] [Medline: 25204467]
31. Ayorinde AA, Bhattacharya S, Druce KL, Jones GT, Macfarlane GJ. Chronic pelvic pain in women of reproductive and post-reproductive age: a population-based study. Eur J Pain 2017 Mar;21(3):445-455. [doi: 10.1002/ejp.938] [Medline: 27634190]
32. Korkeila K, Suominen S, Ahvenainen J, Ojanlatva A, Rautava P, Helenius H, et al. Non-response and related factors in a nation-wide health survey. Eur J Epidemiol 2001;17(11):991-999. [Medline: 12380710]
33. Torrance N, Smith BH, Bennett MI, Lee AJ. The epidemiology of chronic pain of predominantly neuropathic origin. Results from a general population survey. J Pain 2006 Apr;7(4):281-289. [doi: 10.1016/j.jpain.2005.11.008] [Medline: 16618472]
34. Vermeren I. Brandwatch. Men vs. women: who is more active on social media? URL: https://www.brandwatch.com/blog/men-vs-women-active-social-media/ [accessed 2017-06-07] [WebCite Cache ID 6r2gCo5xq]
35. Statista. 2013. Usage penetration of medical, health, and fitness-related apps in Great Britain in 2013, by demographic group URL: https://www.statista.com/statistics/286875/fitness-health-and-medical-app-usage-in-great-britain-by-demographic-group/ [accessed 2017-06-07] [WebCite Cache ID 6r2g5jRzo]
36. Seale C, Charteris-Black J, MacFarlane A, McPherson A. Interviews and internet forums: a comparison of two sources of qualitative data. Qual Health Res 2010 May;20(5):595-606. [doi: 10.1177/1049732309354094] [Medline: 20008955]
37. Seale C, Ziebland S, Charteris-Black J. Gender, cancer experience and internet use: a comparative keyword analysis of interviews and online cancer support groups. Soc Sci Med 2006 May;62(10):2577-2590. [doi: 10.1016/j.socscimed.2005.11.016] [Medline: 16361029]

38. O'Connor S, Hanlon P, O'Donnell C, Garcia S, Glanville J, Mair F. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016;16(1):120. [Medline: 27630020]
39. Thornton L, Batterham PJ, Fassnacht DB, Kay-Lambkin F, Calear AL, Hunt S. Recruiting for health, medical or psychosocial research using Facebook: systematic review. Internet Interv 2016 May;4:72-81. [doi: 10.1016/j.invent.2016.02.001]
40. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
41. Herschman J, Kasenberg T, Levy D, Ruth N, Taberner C, Kaufman M, et al. Development of a smartphone app for adolescents with lupus: a collaborative meeting-based methodology inclusive of a wide range of stakeholders. Rev Panam Salud Publica 2014 Jun;35(5-6):471-476 [FREE Full text] [Medline: 25211579]
42. Aikman H. The association between arthritis and the weather. Int J Biometeorol 1997 Jun;40(4):192-199. [Medline: 9225595]
43. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016;18(5):e97. [Medline: 27185295]

Abbreviations

DMARDs: disease-modifying antirheumatic drugs
GPS: global positioning system
HSE: Health Survey for England
iOS: iPhone operating system
IQR: interquartile range
mHealth: mobile health
NRS: numerical rating scale
NSAIDs: nonsteroidal anti-inflammatory drugs
OA: osteoarthritis

Edited by C Dias; submitted 07.06.17; peer-reviewed by A Elmessiry, M Numans, M Keller; comments to author 13.07.17; revised version received 18.08.17; accepted 27.08.17; published 01.11.17.

Please cite as:
Druce KL, McBeth J, van der Veer SN, Selby DA, Vidgen B, Georgatzis K, Hellman B, Lakshminarayana R, Chowdhury A, Schultz DM, Sanders C, Sergeant JC, Dixon WG
Recruitment and Ongoing Engagement in a UK Smartphone Study Examining the Association Between Weather and Pain: Cohort Study
JMIR Mhealth Uhealth 2017;5(11):e168
URL: http://mhealth.jmir.org/2017/11/e168/
doi:10.2196/mhealth.8162
PMID:29092810

Original Paper

Diabetes Data Management System to Improve Glycemic Control in People With Type 1 Diabetes: Prospective Cohort Study

Concetta Irace[1], MD, PhD; Matthias Axel Schweitzer[2], MD; Cesare Tripolino[3], MD; Faustina Barbara Scavelli[3], MD, PhD; Agostino Gnasso[3], MD

[1]Metabolic Diseases Unit, Department of Health Science, University of Catanzaro, Catanzaro, Italy

[2]Roche Diagnostics GmbH, Mannheim, Germany

[3]Metabolic Diseases Unit, Department of Clinical and Experimental Medicine, University of Catanzaro, Catanzaro, Italy

Corresponding Author:
Agostino Gnasso, MD
Metabolic Diseases Unit
Department of Clinical and Experimental Medicine
University of Catanzaro
University Campus
Catanzaro, 88100
Italy
Phone: 39 09613697039
Fax: 39 09613647250
Email: gnasso@unicz.it

Abstract

Background: Smartphone and Web technology can improve the health care process, especially in chronic diseases.

Objective: The aim of this study was to investigate whether the use of blood glucose (BG) data management system, which enables connection to smartphones, the Web, the cloud, and downloading, can improve glycemic control in subjects with type 1 diabetes mellitus (T1DM).

Methods: This study was a prospective, single-arm, cohort feasibility study with 6 months of duration. T1DM subjects enrolled had experience in self-monitoring blood glucose, but were download data naïve. Fasting BG and glycated hemoglobin (HbA_{1c}) were collected at the enrollment and at follow-up. Subjects were divided into Downloader (DL) and No-downloader (NDL).

Results: A total of 63 subjects were analyzed, of which 30 were classified as DL and 33 as NDL. At the end of the study, DL had significantly lower HbA_{1c}, mean daily glucose, standard deviation, percentage of BG values above target, and pre- and postprandial (lunch and dinner) values compared with NDL (all $P<.05$). The percentage of BG values within treatment target was significantly higher in DL compared with NDL (47% [SD 9] vs 37% [SD 13]; $P=.001$).

Conclusions: The findings suggest that, in T1DM, downloading of BG from data management system, which enables connection to smartphones, the Web, and the cloud, might be a valuable contributor to improved glycemic control.

(JMIR Mhealth Uhealth 2017;5(11):e170) doi:10.2196/mhealth.8532

KEYWORDS

diabetes mellitus; blood glucose self-monitoring; smartphone; internet

Introduction

Background

Optimizing insulin therapy and achieving good metabolic control is still a challenge in the management of type 1 diabetes mellitus (T1DM). Indeed, subjects with T1DM experience higher glycemic variability than those with type 2 diabetes mellitus (T2DM), and this variability is associated with higher risk of hypoglycemia and worse metabolic control [1]. Self-monitoring blood glucose (SMBG) provides real-time information to patients, allowing adjustment of therapy and also prevention of hypoglycemia in everyday life and during specific conditions such as physical activity, stress, and illness. It also allows sufficient interaction between patients and the health care team to analyze blood glucose (BG) data and to evaluate glycemic trends, glycemic variability, and the risk of hypoglycemia and hyperglycemia [2-4]. So far, many studies have demonstrated

the efficacy of SMBG in improving decision making, obtaining better glycemic control, and facilitating a more timely and aggressive change of diabetes therapy, as well as in starting insulin therapy both in type 1 and type 2 diabetes [5-7]. These studies have stressed the need of availability of sufficient BG data, to involve caregivers and patients in the management of the disease, and to share information to achieve a good and stable metabolic control. The idea arising from these trials is that health care providers (HCPs), caregivers, and patients should be in close collaboration for optimal diabetes therapy and outcome. An overall prerequisite is the use of BG data or information. BG information can be gathered through new technologies such as the Internet-enabled BG meter connected to computer systems, mobile phones, and the Web.

The management of diabetes requires some basic steps and rules such as knowing target BG values, gathering glucose data, interpreting glycemic patterns, and taking therapeutic action [8,9].

However, once BG data have been collected by the patient, downloading of the data, using them, and sharing them with HCPs is still limited and sometimes overcomplicated. In clinical practice, BG data in many cases are still shown during the scheduled office visits only or maybe forwarded in advance by fax or mail or via social media. New technologies such as smartphone apps and Web-enabled systems are more and more commonly used to connect all people involved in the management and monitoring of diabetes therapy. Recently, a new connected, Web- and cloud-based system, the Accu-Chek Connect diabetes management system (DMS), has been developed, with an aim to improve collection and management of BG data [10]. The system consists of 3 elements, the BG meter with Bluetooth low energy connectivity, smartphone apps, and a respective Web portal, all of which are wirelessly connected. Patients and caregivers check and tag BG with the meter and download the data into the app and the portal. A dedicated software provides analysis of the glucose data and generates different reports to visualize the information and pattern. Once glucose data are downloaded, HCPs can access the data in the cloud after logging onto the system with their personal account. Analyzed BG data, glycemic trends, hypo- and hyperglycemic events, glucose variability, and mean BG values can be detected. Ease of use and time efficiency of the Accu-Chek Connect DMS have been demonstrated by HCPs, patients with diabetes, and caregivers to obtain information, interpret data, and make therapy decisions [10]. HCPs assessed the system easily and quickly to identify glycemic pattern and to take therapeutic decisions. Identification and therapy decision making was also done in a shorter time compared with traditional BG logbook approaches.

Objective

On the basis of this knowledge and those findings, we have designed a prospective, single-arm, cohort feasibility study in subjects with T1DM using a connected BG data management system and put the primary objective and the focus of the analysis of this study on the relationship between BG download or no-download and the impact of frequency of BG download on glycemic control and therapy success. The hypothesis of the study was that BG downloading has a positive effect on diabetes therapy success and glycemic control, which would go along with findings from previous clinical studies but would also be true and maybe even advanced using a connected BG data management system such as the Accu-Chek Connect DMS.

Methods

Subjects and Study Design

A prospective, single-arm, cohort feasibility study was conducted including consecutive adult subjects with T1DM, who visited our hospital from January to June 2015 and met the inclusion and exclusion criteria. Inclusion criteria were as follows: age ≥18 years, diagnosis of T1DM, glycated hemoglobin (HbA_{1c}) <10%, recent admission to our clinic (≤6 months), ability to perform SMBG and carbohydrates counting or alternative way to adjust bolus insulin before meals, current use of BG meters connecting with the Accu-Chek Connect Smartphone app and the Accu-Chek Connect Web portal, and knowledge on how to download BG data into the system, but so far download naïve.

Exclusion criteria were as follows: current use of continuous glucose monitoring; pregnancy; alcohol consumption exceeding 20 grams per day; clinical conditions requiring intensive insulin treatment such as infection, surgery, and acute vascular event; concomitant corticosteroid treatment; and hyper- or hypothyroidism.

The protocol of the study was submitted to the local ethical committee, and the research was conducted according to local legal requirements and good clinical practice. Eligible patients were included in the study after signing informed consent. Data were collected at baseline and after 6 months (follow-up visit at the end of the study).

Participants were asked to perform self-monitoring following previous habits and download data into the system at least once during the 6-month period. Patients were not encouraged to examine and interpret reports following data downloading. Physicians analyzed downloaded data and suggested therapeutic changes when appropriate and according to clinical guidelines.

Office visits were scheduled as suggested by the conventional standard of diabetes care. If necessary, adjustment of ongoing therapy was communicated to the patient after reviewing the download of SMBG.

At 6-month follow-up visit (end of the study), primary analysis was made from two groups: those who downloaded the data from the BG meter into the system (Downloader, DL) at least one time from enrollment to the follow-up visit and those who did not download the data (No-downloader, NDL). Secondary analyses were made specifically on the impact of frequency of downloading on glycemic control and the impact of downloading in subjects on insulin pump therapy (continuous subcutaneous insulin injection, CSII) as a special subgroup.

On the basis of the educational level, subjects were defined as “student,” “graduated,” and “not-graduated.”

Accu-Chek Connect DMS and Glycemic Outcomes

Accu-Check Connect DMS (Roche, Indianapolis, IN) consists of 3 elements: the BG meter, the smartphone app, and the Web portal. The meter allows measurement of BG values, storing the values in the meter and connecting and transmitting wirelessly the values to the smartphone app. BG data can be sent from the smartphone app to the Web portal or can also directly be downloaded from the BG meter into the portal. The smartphone app can generate messages and automatic reports, for example, a 3-day glycemic trend derived from Structured SMBG. SMBG data, once downloaded, are stored in the cloud and are accessible for the HCP after signing in to the Web portal system. The Web portal is also able to generate automatic messages reporting that the patient file has been updated. The underlying software outputs different analyses and BG data visualization, illustrating glycemic pattern and variables, as well as a traditional logbook design.

Biochemical Variables

Fasting blood glucose (FBG) and HbA_{1c} were measured as recommended by the national guidelines. FBG was measured by the glucose-hexokinase method (Roche, Base, Switzerland); HbA_{1c} was measured with a high-performance liquid chromatographer standardized and aligned to the United Kingdom Prospective Diabetes Study (UKPDS) and the Diabetes Chronic Complications Trial (DCCT) (Menarini, Florence, Italy). For this study, we collected data at baseline and after 6 months.

Statistical Analyses

Statistical analyses were performed using PASW 18.0 for Windows (SPSS, Quarry Bay, HK). Variables not normally distributed were as follows: total daily insulin, percentage of values below the target and within the target, absolute and percentage difference of HbA_{1c} between follow-up and baseline visit, and mean postprandial (lunch and dinner) glucose. A 2-step rank transformation was performed to normalize these variables before applying parametric tests. The *t* test for unpaired data was used to compare means between two groups, and the chi-square test was used to compare percentages. The *t* test for paired data was used to compare variables measured at baseline and follow-up visit within each group.

Results

A total of 63 subjects were enrolled in the study, 44% (28/63) males, aged between 18 and 60 years. Moreover, 52% of participants (33/63) were treated with CSII with insulin pumps, and 48% (30/63) with multiple daily insulin injection. The prevalence of students was 41% (26/63), graduated 16% (10/63), and not-graduated 43% (27/63) (graduated and not-graduated are reported as nonstudent). Among nonstudent participants 31% (11/37), were unemployed, 40% (15/37) employed, and 29% (11/37) independent professional.

On the basis of the download of BG data, subjects were divided into two groups: DL (48%, 30/63) and NDL (52%, 33/63). Among DL, 63% (19/30) subjects downloaded the BG data one time from the baseline to follow-up visit, 34% (10/30) downloaded 2 times, and 3% (1/30) downloaded 3 times. All data downloaded into the system were reviewed by the physicians, and if necessary, therapy was changed accordingly following the current practice.

Characteristics of subjects included in the study at the time of the enrollment, grouped as DL and NDL, are reported in Table 1.

Age, disease duration, and prevalence of male sex were comparable between the groups. The percentage of subjects on CSII therapy was higher in DL compared with NDL, even if the difference was not statistically significant. The prevalence of unemployed, employed, and independent professional in the two groups, respectively, was as follows: DL 19% (6/30), 62% (18/30), and 19% (6/30); NDL 42% (14/33), 37% (12/33), and 21% (7/33), (*P*=.04).

Table 2 shows biochemical and clinical characteristics of DL and NDL collected at baseline and at the follow-up visit.

At baseline, no statistically significant difference between DL and NDL was observed with regard to glycemic control and BG data. At follow-up, DL had significantly lower FBG and HbA_{1c} compared with NDL. HbA_{1c} significantly decreased in DL at follow-up but remained unchanged in NDL.

Table 1. Characteristics of subjects enrolled in the study and grouped as No-downloader (NDL) and Downloader (DL).

Variable	No-downloader (N=33)	Downloader (N=30)
Males, n (%)	14 (42)	14 (46)
Age in years, mean (SD)	29.4 (13.4)	28.8 (12.2)
Disease duration in years, mean (SD)	14.3 (8.4)	17.5 (10.3)
CSII[a], n (%)	15 (45)	18 (60)
Student, n (%)	14 (43)	12 (40)
Graduated, n (%)	7 (21)	3 (10)
Not-graduated, n (%)	12 (36)	50

[a]CSII: continuous subcutaneous insulin injection.

Table 2. Fasting blood glucose, glycated hemoglobin, body weight, and total daily insulin at baseline and follow-up visit in subjects enrolled in the study and divided as No-downloader and Downloader.

Variables	Baseline		Follow-up	
	No-downloader (N=33) mean (SD)	Downloader (N=30) mean (SD)	No-downloader (N=33) mean (SD)	Downloader (N=30) mean (SD)
FBG[a], mg/dL	160 (58)	150 (43)	166 (51)	143 (36)[b]
HbA_{1c}[c], %	7.85 (0.62)	7.51 (0.71)	7.95 (0.74)	7.38 (0.66)[b,d]
Body weight, kg	68.9 (13.6)	69.2 (14.7)	69.4 (13.9)	70.6 (14.4)
TDI[e], unit/day	46.3 (18.3)	47.3 (16.8)	41.9 (15.0)	40.2 (14.8)
Unit per kg body weight	0.67	0.68	0.60	0.57

[a]FBG: fasting blood glucose.

[b]*P*=.003 versus NDL (unpaired *t* test).

[c]HbA_{1c}: glycated hemoglobin.

[d]*P*=.045 versus baseline.

[e]TDI: total daily insulin.

Figure 1. Absolute change in glycated hemoglobin (HbA_{1c}) between baseline and follow-up visit in No-downloader (NDL) and Downloader (DL). Values are expressed as mean (SD). *P=.03.

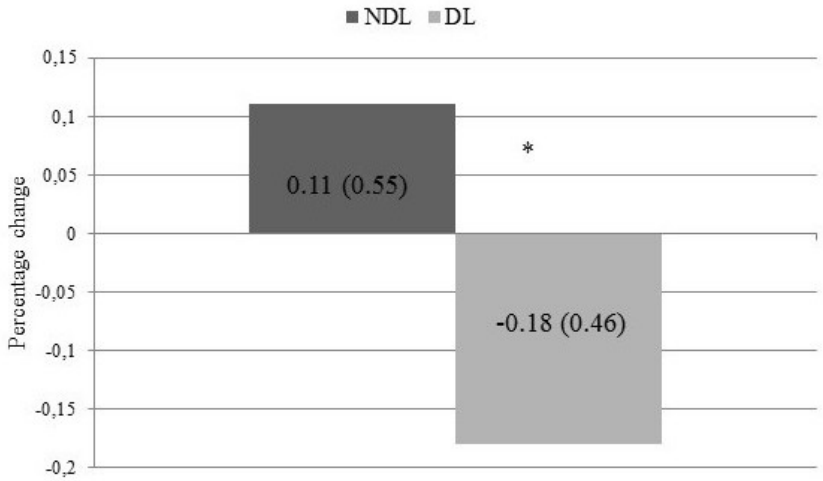


Figure 1 shows the absolute change of HbA_{1c} between baseline and follow-up visit in DL and NDL.

Glycemic variables generated by the BG data management system at follow-up visit as mean of the last 4 weeks are displayed in Figures 2-4.

Figure 2 shows the mean daily glucose and standard deviation, and both were significantly lower in DL compared with NDL. Premeal and postmeal (lunch and dinner) glucose were significantly different between the two groups, whereas prebreakfast glucose was comparable (Figure 3). Figure 4 illustrates the percentage of values within, above, and below target.

At follow-up visit, the percentage of values within the target was significantly higher, and the percentage of values above the target was significantly lower in DL. The prevalence of values below the target was comparable between the two groups. The mean number and (SD) of BG testing per day was comparable between DL and NDL, 4.2 (1.5) versus 3.7 (1.4), *P*=.21.

The same analyses, comparing DL versus NDL, were also performed in patients on CSII therapy, of which 17 were DL and 13 NDL. Mean HbA_{1c} and (SD) at baseline and follow-up visit were 7.6 (0.8) versus 7.4 (0.8) % in DL, and 8.2 (0.4) versus 8.1 (0.6) % in NDL, respectively. Mean absolute difference was −0.26 (0.55) in DL and −0.10 (0.51) in NDL (*P*=.39).

To verify whether the frequency of downloads performed during the study period would influence HbA_{1c} and therapy success, patients were divided into two groups: those who downloaded BG one time (19/30) and those who downloaded ≥2 times (11/30). Mean (SD) HbA_{1c} at baseline and follow-up were, respectively, 7.3 (0.6) versus 7.1 (0.3) % in subjects who downloaded one time, and 7.8 (0.7) versus 7.7 (0.5) % in subjects who downloaded ≥2 times. Mean absolute difference was −0.26 (0.53) in those who downloaded one time and −0.10 (0.29) in those who downloaded ≥2 times (*P*=.22).

Figure 2. Mean (SD) daily glucose and mean standard deviation of No-downloader (NDL) and Downloader (DL) at follow-up visit. *P=.001.

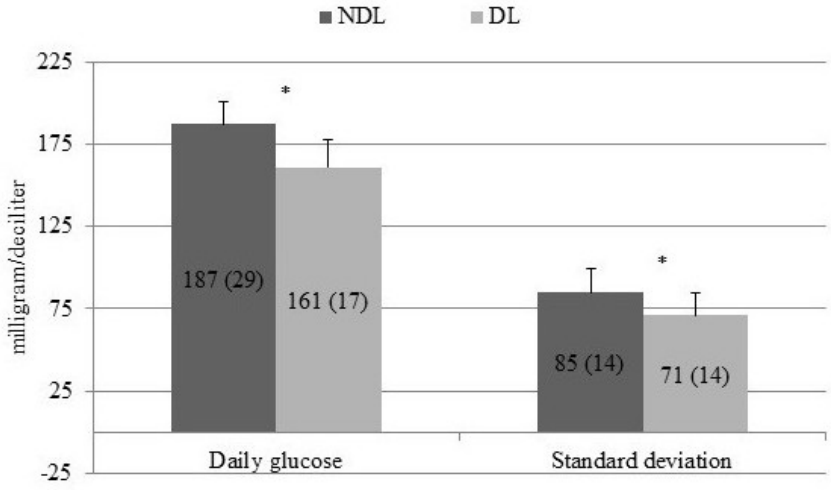
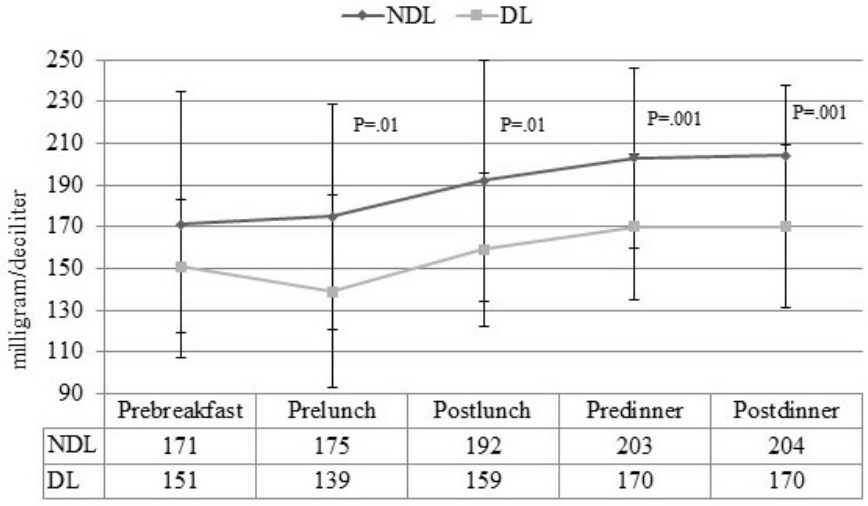
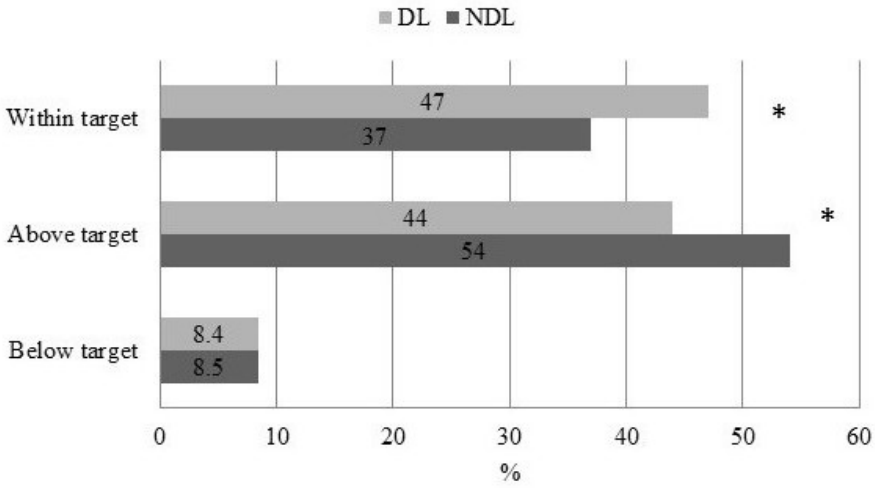


Figure 3. Self-monitoring blood glucose of No-downloader (NDL) and Downloader (DL) at follow-up visit.



	Prebreakfast	Prelunch	Postlunch	Predinner	Postdinner
NDL	171	175	192	203	204
DL	151	139	159	170	170

Figure 4. Percentage of values within, above, and below target of No-downloader (NDL) and Downloader (DL) at follow-up visit. *P=.01.



Discussion

Principal Findings

This study has demonstrated that a new BG data management system, the Accu-Chek Connect DMS, designed to collect, analyze, download, and share BG data, may offer benefits to improve the management of the disease. Indeed, the main finding of this prospective cohort study is the significant difference in HbA_{1c} from baseline to follow-up visit between patients who used the BG information better, DL compared with NDL, with an overall difference of approximately 0.3%. The change might seem modest but is worthy of consideration.

Other BG parameters measured or analyzed in this study, that is, mean daily glucose, standard deviation, and other BG values showed similar advances in DL compared with NDL, all consistent with the difference in HbA_{1c}.

T1DM and T2DM require optimal glycemic control to prevent acute and chronic complications [11,12]. Over the years, technology has supported more and more diabetes monitoring and management. BG meters and software enabling downloading of BG data are tools potentially helping to overcome the limits and drawbacks of manual BG recording to make the data more intelligible and usable [13-15]. Clinical studies with SMBG performed in a structured manner (structured testing) confirmed improvement of glycemic control and therapy success. These findings were based on better quality of BG information, better use of BG information available, and BG information sharing with HCP. Therapy change and decision making were more frequent and aggressive when structured testing was approached. Today's evidence supports that diabetes therapy success is largely dependent on the amount of glucose information available (frequency), the quality of information (eg, structured testing and analyses), and the use of BG data (download and data sharing) in daily diabetes management.

New technology might support better analyses of BG data and better downloading, availability, and use of BG data, thereby improving decision-making and glycemic outcomes [16].

Today, there is much debate about which BG parameters generated by BG data management systems can be more useful for the assessment of glycemic control and the management of the disease. In this regard, some considerations may be useful. Mean value might not be representative of glycemic variability because BG data are generally not normally distributed and the mean value is affected by the number of observations, single outlier, or aberrant values [17,18]. However, based on our results and previous studies, we suggest that BG parameters generated by the Accu-Chek Connect DMS from BG data might assist medical decision making in addition to traditional BG analysis and HbA_{1c}.

Furthermore, another reason that might contribute to making the BG data management systems able to positively impact glycemic control is the ability of those systems to store and facilitate sharing of the data. Once the patient with diabetes has downloaded the data into the cloud, an automatic message is generated, and in turn, the HCP is able to update subject files, further analyze BG, and, if necessary, contact the patient and give suggestion and advice about therapy or even adopt or escalate medication. We have not reported the results on how and when the therapy of subjects has been modified because that was not the aim of our study. However, sharing information through the BG data management system might promote the contact between patients and HCPs and allow use of glucose data for therapy advice or optimization whenever it is needed. The recent paper by Chow et al has reported that the frequency of online communication with HCP, along with an adequate number of tests per day (twice or more), is associated with a lower HbA_{1c} in T2DM patients on oral medication. In other words, the number of tests per day matters, but if glycemic values are frequently communicated via Internet to the physician, the efficacy of the numbers in terms of HbA_{1c} reduction is greater [19].

We have designed our study as a prospective study and included subjects with T1DM who benefit from close relationship with their HCPs, continuous support in insulin dose adjustment, high level of disease knowledge, and motivation of constant self-management. In this specific scenario, the remote data management might contribute and support those needs and enable close collaboration between patients with diabetes and HCPs. The ability of collecting, downloading, and sharing BG data can improve timely availability and the use of BG data and enhance the achievement of a better glycemic control. The comparison of DL versus NDL stands for availability and use of BG data, and improvement or difference in HbA_{1c} as found in the study was largely expected. The improvement in glycemic control was a consistent finding in the overall study population and in the CSII subgroup even if the result lacked significance in that subpopulation, probably because of the low number of subjects. In addition to the download of the data, the improvement of HbA_{1c} was probably because of the review and interpretation of the data. Indeed, in our study, all data were reviewed by a physician and, if necessary, the therapy was changed.

We did not find any difference between DL and NDL in the percentage of hypoglycemic events, defined as capillary glucose lower than 70 mg/dL. Hypoglycemia is a very common event in T1DM and the pathogenesis is very complex. It has been estimated that each individual experiences about 2 episodes of symptomatic hypoglycemia per week in real life [20]. As far as our results are concerned, we might argue that a more frequent downloading, or an active interpretation of the results by the patient, should offer the opportunity to significantly decrease the number of hypoglycemic events. Similarly, we did not find any significant difference of HbA_{1c} between subjects who downloaded once and those who downloaded more than once during the observation period. Due to the small sample size and the fact that patients essentially downloaded only 1 or 2 times in the 6 months, we could not assess whether more frequent downloading or reporting would have impacted the HbA_{1c} among DLs.

The number of subjects included in this study did also allow analyzing the impact of the frequency of downloading on glycemic control and therapy success. Our study should be

considered as a preliminary study, and based on the results, further information about number of data obtained during the day and downloaded, structured testing, use of additional technologies, and motivation of the subjects might influence the results.

Consistent to the findings that BG downloading improved glycemic control and therapy success in the overall study population and in the subgroup treated with CSII, more frequent downloading resulted in a similar effect, even if the results (due to small number of subjects) were not significant. This prospective cohort study adds evidence that downloading BG data from a BG data management system, which stands for availability and use of BG data, has positive effects on glycemic control and diabetes therapy success.

Finally, we would like to hypothesize that BG data management systems might be effective in reducing the number of ambulatory care visits. It has been estimated that approximately 50% of subjects with T1DM and T2DM make 4 or more visits in 1 year, and in approximately one-third of all visits, no change of therapy is suggested or new drugs are added to titrate therapy [21].

Future Work

However, even if the observation in this study is encouraging, additional data from larger randomized controlled studies are needed to better identify the clinical setting and the patients who can benefit from the change of care delivered by new technology, including different software available.

In chronic diseases in general, and especially in T1DM, which often occurs at a young age, the need for close collaboration between patients and HCPs, constant checks, and therapeutic adjustments often have a major impact on the quality of life. Technology cannot fully replace personal interaction between patients and physicians but can help to find new ways of delivering care and contribute to therapy success and daily diabetes management.

Conclusions

In conclusion, BG data management system, which allows collecting, analyzing, downloading, and sharing BG data, offers the opportunity to improve communication between patients and HCPs and connects patients with all stakeholders needed or wanted. The finding of the study supports the importance of BG data download for good glycemic control and diabetes therapy success, as downloading stands for availability and use of BG information. Other functionalities resulting from a connected BG meter to smartphone apps, Web portals, and the cloud also remotely analyzed BG might help in addition to take better therapeutic decisions, potentially decrease the number of office-based visits, and adopt the way diabetes care is delivered. Larger, controlled clinical studies are needed to fully endorse reported findings and support the use of new technologies further.

Conflicts of Interest

MAS was Senior Vice President Medical Affairs Diabetes Care, Roche Diagnostic.

References

1. Kilpatrick ES, Rigby AS, Goode K, Atkin SL. Relating mean blood glucose and glucose variability to the risk of multiple episodes of hypoglycaemia in type 1 diabetes. Diabetologia 2007 Dec;50(12):2553-2561. [doi: 10.1007/s00125-007-0820-z] [Medline: 17882397]
2. Handelsman Y, Bloomgarden ZT, Grunberger G, Umpierrez G, Zimmerman RS, Bailey TS, et al. American association of clinical endocrinologists and american college of endocrinology - clinical practice guidelines for developing a diabetes mellitus comprehensive care plan - 2015. Endocr Pract 2015 Apr;21(Suppl 1):1-87 [FREE Full text] [doi: 10.4158/EP15672.GL] [Medline: 25869408]
3. Powers MA, Bardsley J, Cypress M, Duker P, Funnell MM, Fischl AH, et al. Diabetes self-management education and support in type 2 diabetes: a joint position statement of the American Diabetes Association, the American Association of Diabetes Educators, and the Academy of Nutrition and Dietetics. Diabetes Educ 2015 Aug;41(4):417-430. [doi: 10.1177/0145721715588904] [Medline: 26047627]
4. American Diabetes Association. Standards of medical care in diabetes-2015 abridged for primary care providers. Clin Diabetes 2015 Apr;33(2):97-111. [doi: 10.2337/diaclin.33.2.97] [Medline: 25897193]
5. Mannucci E, Antenore A, Giorgino F, Scavini M. Effects of structured versus unstructured self monitoring of blood glucose on glucose control in patients with non-insulin treated type 2 diabetes: a meta-analysis of randomized controlled trials. J Diabetes Sci Technol 2017 Jul 01 (forthcoming). [doi: 10.1177/1932296817719290] [Medline: 28697625]
6. Young LA, Buse JB, Weaver MA, Vu MB, Mitchell CM, Blakeney T, Monitor Trial Group. Glucose self-monitoring in non-insulin treated patients with type 2 diabetes in primary care settings: a randomized trial. JAMA Intern Med 2017 Jul 01;177(7):920-929. [doi: 10.1001/jamainternmed.2017.1233] [Medline: 28600913]
7. American Diabetes Association. (6) Glycemic targets. Diabetes Care 2015 Jan;38 Suppl:S33-S40. [doi: 10.2337/dc15-S009] [Medline: 25537705]
8. Parkin CG, Davidson JA. Value of self-monitoring blood glucose pattern analysis in improving diabetes outcomes. J Diabetes Sci Technol 2009 May 01;3(3):500-508 [FREE Full text] [doi: 10.1177/193229680900300314] [Medline: 20144288]

9. Or CK, Tao D. Does the use of consumer health information technology improve outcomes in the patient self-management of diabetes? A meta-analysis and narrative review of randomized controlled trials. Int J Med Inform 2014 May;83(5):320-329. [doi: 10.1016/j.ijmedinf.2014.01.009] [Medline: 24534118]
10. Hinnen DA, Buskirk A, Lyden M, Amstutz L, Hunter T, Parkin CG, et al. Use of diabetes data management software reports by health care providers, patients with diabetes, and caregivers improves accuracy and efficiency of data analysis and interpretation compared with traditional logbook data: first results of the Accu-Chek Connect Reports Utility and Efficiency Study (ACCRUES). J Diabetes Sci Technol 2015 Mar;9(2):293-301 [FREE Full text] [doi: 10.1177/1932296814557188] [Medline: 25367012]
11. Diabetes Control and Complications Trial Research Group, Nathan DM, Genuth S, Lachin J, Cleary P, Crofford O, et al. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 1993 Dec 30;329(14):977-986. [doi: 10.1056/NEJM199309303291401] [Medline: 8366922]
12. Paneni F, Lüscher TF. Cardiovascular protection in the treatment of type 2 diabetes: a review of clinical trial results across drug classes. Am J Cardiol 2017 Jul 01;120(1S):S17-S27 [FREE Full text] [doi: 10.1016/j.amjcard.2017.05.015] [Medline: 28606340]
13. Colorafi K. Connected health: a review of the literature. Mhealth 2016;2:13 [FREE Full text] [doi: 10.21037/mhealth.2016.03.09] [Medline: 28293591]
14. Hunt CW. Technology and diabetes self-management: an integrative review. World J Diabetes 2015 Mar 15;6(2):225-233 [FREE Full text] [doi: 10.4239/wjd.v6.i2.225] [Medline: 25789104]
15. Carallo C, Scavelli FB, Cipolla M, Merante V, Medaglia V, Irace C, et al. Management of type2 diabetes mellitus through telemedicine. PLoS One 2015;10(5):e0126858 [FREE Full text] [doi: 10.1371/journal.pone.0126858] [Medline: 25974092]
16. Given JE, O'Kane MJ, Bunting BP, Coates VE. Comparing patient-generated blood glucose diary records with meter memory in diabetes: a systematic review. Diabet Med 2013 Aug;30(8):901-913. [doi: 10.1111/dme.12130] [Medline: 23324062]
17. Rodbard D. Optimizing display, analysis, interpretation and utility of self-monitoring of blood glucose (SMBG) data for management of patients with diabetes. J Diabetes Sci Technol 2007 Jan;1(1):62-71 [FREE Full text] [doi: 10.1177/193229680700100111] [Medline: 19888382]
18. Austin MM, Haas L, Johnson T, Parkin CG, Parkin CL, Spollett G, et al. Self-monitoring of blood glucose: benefits and utilization. Diabetes Educ 2006;32(6):835-6, 844. [doi: 10.1177/0145721706295873] [Medline: 17102152]
19. Chow N, Shearer D, Aydin PJ, Pottinger B, Pawlowska M, White A, et al. Blood glucose self-monitoring and internet diabetes management on A1C outcomes in patients with type 2 diabetes. BMJ Open Diabetes Res Care 2016;4(1):e000134 [FREE Full text] [doi: 10.1136/bmjdrc-2015-000134] [Medline: 27158516]
20. McCrimmon RJ, Sherwin RS. Hypoglycemia in type 1 diabetes. Diabetes 2010 Oct;59(10):2333-2339 [FREE Full text] [doi: 10.2337/db10-0103] [Medline: 20876723]
21. Ashman JJ, Talwalkar A, Taylor SA. CDC. 2014 Jul. Age differences in visits to office-based physicians by patients with diabetes: United States, 2010 URL: https://www.cdc.gov/nchs/products/databriefs/db161.htm [accessed 2017-11-01] [WebCite Cache ID 6udtRmLQZ]

Abbreviations

BG: blood glucose
CSII: continuous subcutaneous insulin injection
DCCT: Diabetes Chronic Complications Trial
DL: Downloader
DMS: diabetes management system
FBG: fasting blood glucose
HCPs: health care providers
NDL: No-downloader
SMBG: self-monitoring blood glucose
T1DM: type 1 diabetes mellitus
T2DM: type 2 diabetes mellitus
UKPDS: United Kingdom Prospective Diabetes Study

Edited by G Eysenbach; submitted 22.07.17; peer-reviewed by M Pawlowska, K Colorafi; comments to author 23.08.17; accepted 04.09.17; published 21.11.17.

Please cite as:
Irace C, Schweitzer MA, Tripolino C, Scavelli FB, Gnasso A
Diabetes Data Management System to Improve Glycemic Control in People With Type 1 Diabetes: Prospective Cohort Study
JMIR Mhealth Uhealth 2017;5(11):e170
URL: http://mhealth.jmir.org/2017/11/e170/
doi:10.2196/mhealth.8532
PMID:29162560

Original Paper

One Drop | Mobile on iPhone and Apple Watch: An Evaluation of HbA1c Improvement Associated With Tracking Self-Care

Chandra Y Osborn[1], PhD, MPH; Joost R van Ginkel[2], PhD; David G Marrero[3], PhD; David Rodbard[4], MD; Brian Huddleston[5], JD; Jeff Dachis[1], MA

[1]Informed Data Systems Inc, New York, NY, United States
[2]Leiden University, Leiden, Netherlands
[3]The University of Arizona Health Sciences, Tucson, AZ, United States
[4]Biomedical Informatics Consultants LLC, Potomac, MD, United States
[5]Informed Data Systems Inc, Austin, TX, United States

Corresponding Author:
Chandra Y Osborn, PhD, MPH
Informed Data Systems Inc
85 Delancey St, Ste 71
New York, NY, 10002
United States
Phone: 1 8604242858
Email: chandra@onedrop.today

Abstract

Background: The One Drop | Mobile app supports manual and passive (via HealthKit and One Drop's glucose meter) tracking of self-care and glycated hemoglobin A_{1c} (HbA_{1c}).

Objective: We assessed the HbA_{1c} change of a sample of people with type 1 diabetes (T1D) or type 2 diabetes (T2D) using the One Drop | Mobile app on iPhone and Apple Watch, and tested relationships between self-care tracking with the app and HbA_{1c} change.

Methods: In June 2017, we identified people with diabetes using the One Drop | Mobile app on iPhone and Apple Watch who entered two HbA_{1c} measurements in the app 60 to 365 days apart. We assessed the relationship between using the app and HbA_{1c} change.

Results: Users had T1D (n=65) or T2D (n=191), were 22.7% (58/219) female, with diabetes for a mean 8.34 (SD 8.79) years, and tracked a mean 2176.35 (SD 3430.23) self-care activities between HbA_{1c} entries. There was a significant 1.36% or 14.9 mmol/mol HbA_{1c} reduction (F=62.60, P<.001) from the first (8.72%, 71.8 mmol/mol) to second HbA_{1c} (7.36%, 56.9 mmol/mol) measurement. Tracking carbohydrates was independently associated with greater HbA_{1c} improvement (all P<.01).

Conclusions: Using One Drop | Mobile on iPhone and Apple Watch may favorably impact glycemic control.

(JMIR Mhealth Uhealth 2017;5(11):e179) doi:10.2196/mhealth.8781

KEYWORDS

type 1 diabetes; type 2 diabetes; mobile health; mobile phone; smartwatch; glycated hemoglobin A1c; HbA1c; glycemic control; self-care behavior

Introduction

The digital diabetes ecosystem is booming [1,2], with more than 1500 mobile apps supporting diabetes management [3], yet very few diabetes apps have been studied. For the few that have, they significantly reduce glycated hemoglobin A_{1c} (HbA_{1c}) by an average 0.49% [4].

The HbA_{1c} measurement is the amount of hemoglobin in the blood with glucose attached to it. People are diagnosed with diabetes when their HbA_{1c} level is 6.5% or greater. An HbA_{1c} of 7.0% or greater puts people with diabetes at risk of developing macrovascular and microvascular complications, whereas a HbA_{1c} less than 7.0% or reducing HbA_{1c} by 1.0% prevents complications [5,6]. Diabetes self-care (eg, eating

fewer carbohydrate grams, being more active, taking medications) improves HbA_{1c} levels.

Diabetes apps offer tracking of self-care and can educate and motivate people to better care for their health [1]. Together, the widely used diabetes apps rate highly in terms of functionality, aesthetics, and engagement [7]. Devices, sensors, wearables, and watches that passively collect data may bolster engagement. Passive data collection makes a more useful and less burdensome diabetes app [1,8]. Very few apps, however, offer manual and passive data collection from a mobile phone and a smartwatch, and no study to our knowledge has explored the health benefit of this type of digital solution.

The One Drop | Mobile app offers manual data entry, but also passive data collection via Apple's HealthKit, Apple Watch, and the Bluetooth-enabled One Drop | Chrome glucose meter. We hypothesized that there would be a pre-post HbA_{1c} change among people with diabetes using the One Drop | Mobile app on an iPhone and Apple Watch. We also hypothesized self-care tracking with the app would be associated with HbA_{1c} change.

Methods

One Drop | Mobile: A Mobile Phone and Smartwatch App

The One Drop | Mobile app is free and available on iOS, WatchOS, and Android operating systems. One Drop users manually and passively (via HealthKit for iPhone and Apple Watch, Google Fit for Android mobile phones, and the Bluetooth-enabled One Drop | Chrome blood glucose meter) store and track blood glucose readings, medication doses, physical activity, and carbohydrates consumed. A built-in food library expedites carbohydrate tracking. A medication scheduler reminds users when a dose is due, and tracks doses upon confirmation. Statistics of tracked data are viewable on iPhone, Android, and Apple Watch.

Watch app users can enter data directly from their Watch, and view statistics of their data and monitor goal progress on the Watch face. They can get push notifications on their Watch, including medication reminders and motivational messages prompting and reinforcing self-care.

On the mobile phone app, users can view in-depth statistics of their data and track HbA_{1c} test results and body weight. An in-app "Newsfeed" delivers health tips, articles, infographics, and more. A "Community" section facilitates learning from, supporting, and receiving support from other users. The iPhone app has a "Notifications" inbox with data-driven insights, achievements, reminders, and support accumulated from other users.

Procedures

On June 6, 2017, we identified people with type 1 (T1D) or type 2 diabetes (T2D) using the One Drop | Mobile app on an iPhone and Apple Watch who had manually entered at least two HbA_{1c} values in the app with HbA_{1c} test dates 60 to 365 days apart. We did not recruit participants. Instead, we analyzed data collected from real users who elected to use the One Drop | Mobile app on their mobile phone and smartwatch devices.

Users enter and store self-care and health data in the One Drop | Mobile app. All data exist in a secure server in the cloud. We characterized users with app-entered demographics (eg, gender, diabetes type). We tested their HbA_{1c} change (ie, self-reported HbA_{1c} collected in the app). We also tested if tracking self-care with the app (ie, the number of times food, activity, blood glucose, and medications were stored in the app between HbA_{1c} measurements) was associated with HbA_{1c} change.

All users agree to an end-user license agreement (EULA). In this agreement, it states that, as a user, you "grant One Drop a perpetual, transferrable, sublicensable, worldwide, nonexclusive, royalty-free license to reproduce, distribute, use, modify, remove, publish, transmit, publicly perform, publicly display, or create derivative works of Your User Content for any purpose without compensation to you, including for the purpose of promoting One Drop and the App, including after your account is cancelled or otherwise terminated." It also states that, "One Drop...may track and report your activity inside of the App, including for analytics purposes." The full EULA is available in the app and online.

Measures

User Characteristics

Gender, diabetes type, and year of diagnosis are self-reported in the app. The difference between year of diagnosis and year of One Drop account creation determined years of diagnosed diabetes. Passively collected time zone data determined user location. User location was dichotomized as United States versus non-United States in analyses because few users outside the United States had entered two HbA_{1c} measurements required for inclusion.

Insulin Status

We reviewed medication names tracked and scheduled in the app to determine if a user was taking insulin or not.

Self-Care

We summed self-care data tracked between two HbA_{1c} entries (60-365 days apart), generating counts of blood glucose, food (carbohydrates), medications, activity, and the overall number of self-care entries tracked in the app during that time.

Glycemic Control

Test results and test dates of HbA_{1c} were self-reported in the app. Self-reported recall of a HbA_{1c} test is highly sensitive (99%) to medical records and claims data documenting an actual HbA_{1c} test [9]. A self-reported HbA_{1c} result is sensitive (79%) to a lab HbA_{1c} test result [10]. Further, we used mean blood glucose measured before the second HbA_{1c} test date to exclude invalid HbA_{1c} measurements and, subsequently, validate self-reported HbA_{1c} at that time point (see Analyses section).

We used HbA_{1c} test dates to calculate the number of days between HbA_{1c} entries. We divided 365 days by 12 months to get 30.42 (days per) month. We divided the number of days

between HbA_{1c} entries by 30.42 (days per) month to get the number of months between HbA_{1c} measurements.

Study Oversight

One Drop, Informed Data Systems Inc (IDS) received an exemption for institutional review board approval and a waiver of informed consent from Solutions IRB, an independent ethics review company (Little Rock, AR and Yarnell, AZ) to study all de-identified data owned by One Drop IDS. All One Drop | Mobile app users must actively agree to a EULA detailing data ownership and use.

Analyses

All analyses were performed using SPSS version 23 (IBM Corp). Summary statistics characterized the sample. Mann-Whitney *U* tests were used for diabetes type differences with continuous variables, and chi-square tests for differences with dichotomous variables. One user with T1D selected "other" for gender. Because "other" gender was infrequently selected, we removed the "other" gender subgroup prior to testing diabetes type differences on gender.

To exclude invalid self-reported HbA_{1c} data, we used the formula HbA_{1c}=(90-day mean blood glucose + 77.3)/35.6 [11] to compare self-reported HbA_{1c} to 90-day mean blood glucose, and excluded users with a greater or less than 2.0% difference (n=44 were excluded). Spearman rho correlations verified the relationship between self-reported HbA_{1c} and mean blood glucose consistent with prior research [12].

Two variables had missing data: gender (37/256, 14.4%) and duration of diagnosed diabetes (47/256, 18.3%). Multiple imputation corrected for missing data on these variables [13]. We used predictive mean matching [14,15] to impute 100 datasets.

Three mixed-effects repeated measures models tested mean HbA_{1c} differences. The first unadjusted model tested the effects of time, diabetes type, and the interaction of time by diabetes type. The second model tested these effects adjusted for a priori covariates: gender, location, years of diagnosed diabetes, and months between HbA_{1c} measurements. We restricted the third model to users with T2D and tested the time effect only adjusted for a priori covariates and insulin status.

Finally, four multiple regression models tested relationships between self-care tracking with the app and HbA_{1c} change. The first unadjusted model assessed the relationships between the amount of tracking by self-care type and HbA_{1c} change. The second model introduced diabetes type. The third model added a priori covariates. The fourth model included users with T2D only, a priori covariates, and insulin status.

Results

Users (N=256) had T1D (n=65) or T2D (n=191), and were 22.7% (58/219) female, diagnosed with diabetes for a mean 8.34 (SD 8.79) years, and tracked a mean 2176.35 (SD 3430.23) self-care activities in the app between HbA_{1c} entries. Across each of four self-care types, the Shapiro-Wilk test statistic ranged from 0.22 to 0.86 (all P<.001), signifying a non-normal distribution. We dichotomized each self-care variable to tracked versus not tracked to satisfy assumptions of statistical tests.

Table 1 presents median and interquartile ranges, n (%), or mean and standard deviation with P values for diabetes type differences on observed variables before multiple imputation. Compared to users with T2D, users with T1D had diabetes for more years and entered more self-care data in the app between HbA_{1c} measurements, particularly blood glucose readings. Self-reported HbA_{1c} and 90-day mean blood glucose were strongly correlated (ρ=.75, P<.001), even when stratified by diabetes type (T1D: ρ=.84, P<.001; T2D: ρ=.72, P<.001). This is consistent with previous cohort studies reporting correlations varying from .71 to .86 [12].

In unadjusted and adjusted models, there was a significant 1.36% (14.9 mmol/mol) HbA_{1c} reduction (unadjusted and adjusted F=62.60, P<.001) during a median 4.06 (IQR 2.82) months (unadjusted: 8.26% [66.8 mmol/mol] to 6.90% [51.9 mmol/mol]; adjusted 8.72% [71.8 mmol/mol] to 7.36% [56.9 mmol/mol]). In the adjusted model, users with T1D had an average 0.41% (F=4.38, P=.04) higher HbA_{1c} than users with T2D, but there was no time by diabetes type interaction. After adjusting for a priori covariates and insulin status, users with T2D had a 1.27% (13.9 mmol/mol) HbA_{1c} reduction (F=364.50, P<.001; 8.16% [65.7 mmol/mol] to 6.89% [51.8 mmol/mol]).

Finally, using the app to track carbohydrates was associated with greater HbA_{1c} improvement even after adjusting for covariates and insulin status for users with T2D (all P<.01).

Table 1. Sample characteristics with tests of difference by diabetes type.

User characteristics	Total (N=256)	Type 1 diabetes (n=65)	Type 2 diabetes (n=191)	*P* [a]
Gender, n (%)				
Male	161 (62.9)	40 (61.5)	121 (63.4)	.91
Female	58 (22.7)	14 (21.5)	44 (23.0)	
Location, n (%)				
United States	217 (84.8)	54 (83.1)	163 (85.4)	.66
Europe	27 (10.5)	9 (13.8)	18 (9.4)	
Asia	8 (3.1)	2 (3.1)	6 (3.1)	
Pacific	2 (0.8)	0	2 (1.0)	
Africa	2 (0.8)	0	2 (1.0)	
Diabetes duration (years), mean (SD)	8.3 (8.8)	13.3 (11.6)	7.1 (7.7)	<.001
Insulin status (yes), n (%)	136 (53.1)	65 (100)	71 (37.2)	<.001
Self-care, n (%)				
App self-care entries	1439.5 (1809)	2055.0 (4264)	1318.0 (1463)	.002
Food entries	17.0 (166)	15.0 (150)	18.0 (178)	.67
Activity entries	628.5 (1049)	470.0 (1170)	664.0 (966)	.31
Blood glucose entries	115.0 (243)	193.0 (567)	94.0 (210)	.02
Medication entries	221.0 (452)	279.0 (3657)	207.0 (367)	.06
Glycemic control				
Months between HbA_{1c} entries, median (IQR)	4.06 (2.82)	5.16 (4.29)	3.88 (2.66)	.003
First HbA_{1c} (%), mean (SD)	8.23 (2.27)	8.31 (2.47)	8.20 (2.20)	.87
Second HbA_{1c} (%), mean (SD)	6.80 (0.99)	7.09 (1.15)	6.70 (1.39)	.01

[a] From chi-square or Mann-Whitney *U* tests.

Discussion

We assessed the HbA_{1c} change of 256 people with diabetes using the One Drop | Mobile app on an iPhone and Apple Watch for up to one year. HbA_{1c} decreased by 1.36% (14.9 mmol/mol) in a median of approximately 4 months. Using the app to track carbohydrates was independently associated with HbA_{1c} improvement.

To our knowledge, this study is the first to evaluate the HbA_{1c} benefit of a tethered diabetes mobile phone and smartwatch app. One study asked people with T1D to use a phone and smartwatch app and give qualitative feedback [16]. Users appreciated entering and viewing data from their watch, the watch's connectivity to their phone, and viewing reminders on their watch. One Drop | Mobile on Apple Watch delivers all three benefits and, based on our findings, may improve glycemic control.

There are study limitations. This is not a randomized controlled trial, preventing causal conclusions. The sample was self-selected, limiting generalizability. HbA_{1c} measurements were self-reported rather than assessed with a laboratory assay. Passively collected data are less prone to social desirability biases, but have their own reliability and validity issues [17]. The One Drop | Mobile app has features we did not evaluate or adjust for in our analyses. Finally, we do not know users' age or socioeconomic status (eg, income, education, insurance status), preventing generalizability to all ages and socioeconomic groups.

Despite these limitations, people of all ages [18], race/ethnicities, and socioeconomic backgrounds [19] increasingly want to use smart devices to assist in the management of diabetes [20]. Research needs to critically evaluate diabetes apps, trackers, and smartwatches, especially as new devices enter the marketplace. Findings must be disseminated directly to consumers and to physicians who can assess these tools and make recommendations accordingly.

Acknowledgments

This work was funded by Informed Data Systems Inc.

Conflicts of Interest

CO, BH, and JD are full-time employees and have stock in Informed Data Systems Inc, manufacturer of the One Drop | Mobile mobile phone and smartwatch mobile app. Informed Data Systems Inc paid JRvG for statistical services required for this research. DM serves on a clinical advisory board for the One Drop | Experts program unrelated to this research. DR has been paid by Informed Data Systems Inc for consultant services unrelated to this research.

References

1. Heintzman ND. A digital ecosystem of diabetes data and technology: services, systems, and tools enabled by wearables, sensors, and apps. J Diabetes Sci Technol 2015 Dec 20;10(1):35-41 [FREE Full text] [doi: 10.1177/1932296815622453] [Medline: 26685994]
2. Klonoff DC, Kerr D. Digital diabetes communication: there's an app for that. J Diabetes Sci Technol 2016 Sep;10(5):1003-1005 [FREE Full text] [doi: 10.1177/1932296816660210] [Medline: 27464752]
3. Research2Guidance. 2016. Diabetes app market report 2016-2021 URL: https://research2guidance.com/product/diabetes-app-market-report-2016-2021/ [accessed 2017-08-04] [WebCite Cache ID 6uqu1t7hI]
4. 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]
5. Diabetes Control Complications Trial (DCCT)/Epidemiology of Diabetes InterventionsComplications (EDIC) Study Research Group. Intensive diabetes treatment and cardiovascular outcomes in type 1 diabetes: the DCCT/EDIC study 30-year follow-up. Diabetes Care 2016 May;39(5):686-693. [doi: 10.2337/dc15-1990] [Medline: 26861924]
6. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet 1998 Sep 12;352(9131):837-853. [Medline: 9742976]
7. Chavez S, Fedele D, Guo Y, Bernier A, Smith M, Warnick J, et al. Mobile apps for the management of diabetes. Diabetes Care 2017 Oct;40(10):e145-e146. [doi: 10.2337/dc17-0853] [Medline: 28774944]
8. Harvey C, Koubek R, Bégat V, Jacob S. Usability evaluation of a blood glucose monitoring system with a spill-resistant vial, easier strip handling, and connectivity to a mobile app: improvement of patient convenience and satisfaction. J Diabetes Sci Technol 2016 Sep;10(5):1136-1141 [FREE Full text] [doi: 10.1177/1932296816658058] [Medline: 27390222]
9. Fowles JB, Rosheim K, Fowler EJ, Craft C, Arrichiello L. The validity of self-reported diabetes quality of care measures. Int J Qual Health Care 1999 Oct;11(5):407-412. [Medline: 10561032]
10. Kumar S, Moseson H, Uppal J, Osborn CY, Heyman M, Juusola J. Impact of a diabetes mobile app with in-app coaching on glycemic control (63LB). 2017 Presented at: American Diabetes Association 77th Scientific Sessions; Jun 9-13, 2017; San Diego, CA.
11. Rohlfing CL, Wiedmeyer HM, Little RR, England JD, Tennill A, Goldstein DE. Defining the relationship between plasma glucose and HbA(1c): analysis of glucose profiles and HbA(1c) in the Diabetes Control and Complications Trial. Diabetes Care 2002 Feb;25(2):275-278. [Medline: 11815495]
12. Makris K, Spanou L. Is there a relationship between mean blood glucose and glycated hemoglobin? J Diabetes Sci Technol 2011 Nov 01;5(6):1572-1583 [FREE Full text] [doi: 10.1177/193229681100500634] [Medline: 22226280]
13. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley; 1987.
14. Little RJ. Missing-data adjustments in large surveys. J Bus Econ Stat 1988 Jul;6(3):287-296. [doi: 10.2307/1391878]
15. Rubin DB. Statistical matching using file concatenation with adjusted weights and multiple imputations. J Bus Econ Stat 1986 Jan;4(1):87-94. [doi: 10.2307/1391390]
16. Årsand E, Muzny M, Bradway M, Muzik J, Hartvigsen G. Performance of the first combined smartwatch and smartphone diabetes diary application study. J Diabetes Sci Technol 2015 May;9(3):556-563 [FREE Full text] [doi: 10.1177/1932296814567708] [Medline: 25591859]
17. Fokkema T, Kooiman TJ, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers depend on walking speed. Med Sci Sports Exerc 2017 Apr;49(4):793-800. [doi: 10.1249/MSS.0000000000001146] [Medline: 28319983]
18. 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]
19. Ramirez V, Johnson E, Gonzalez C, Ramirez V, Rubino B, Rossetti G. Assessing the use of mobile health technology by patients: an observational study in primary care clinics. JMIR Mhealth Uhealth 2016 Apr 19;4(2):e41. [doi: 10.2196/mhealth.4928] [Medline: 27095507]
20. Lithgow K, Edwards A, Rabi D. Smartphone app use for diabetes management: evaluating patient perspectives. JMIR Diabetes 2017 Jan 23;2(1):e2. [doi: 10.2196/diabetes.6643]

Abbreviations

EULA: end-user license agreement
HbA_{1c}: glycated hemoglobin A_{1c}
IDS: Informed Data Systems

Edited by G Eysenbach; submitted 18.08.17; peer-reviewed by A Agne, J Bollyky, A Cross; comments to author 15.09.17; revised version received 06.10.17; accepted 29.10.17; published 29.11.17.

Please cite as:
Osborn CY, van Ginkel JR, Marrero DG, Rodbard D, Huddleston B, Dachis J
One Drop | Mobile on iPhone and Apple Watch: An Evaluation of HbA1c Improvement Associated With Tracking Self-Care
JMIR Mhealth Uhealth 2017;5(11):e179
URL: http://mhealth.jmir.org/2017/11/e179/
doi:10.2196/mhealth.8781
PMID:29187344

Original Paper

Mobile Phone Multilevel and Multimedia Messaging Intervention for Breast Cancer Screening: Pilot Randomized Controlled Trial

Hee Lee[1], PhD; Rahel Ghebre[2], MD; Chap Le[3], PhD; Yoo Jeong Jang[4], MA; Monica Sharratt[1], MSW; Douglas Yee[5], MD

[1]School of Social Work, College of Education and Human Development, University of Minnesota, Twin Cities, St Paul, MN, United States

[2]Department of Obstetrics, Gynecology and Women's Health, School of Medicine, University of Minnesota, Twin Cities, Minneapolis, MN, United States

[3]Division of Biostatistics, School of Public Health, University of Minnesota, Twin Cities, Minneapolis, MN, United States

[4]Department of Educational Psychology, College of Education and Human Development, University of Minnesota, Twin Cities, Minneapolis, MN, United States

[5]Division of Hematology, Oncology and Transplantation, School of Medicine and Pharmacology, University of Minnesota, Twin Cities, Minneapolis, MN, United States

Corresponding Author:
Hee Lee, PhD
School of Social Work
College of Education and Human Development
University of Minnesota, Twin Cities
1404 Gortner Ave.
St Paul, MN, 55108
United States
Phone: 1 6126243689
Fax: 1 6126243744
Email: hylee@umn.edu

Abstract

Background: Despite the increasing breast cancer incidence and mortality rates, Korean American immigrant women have one of the lowest rates of breast cancer screening across racial groups in the United States. Mobile health (mHealth), defined as the delivery of health care information or services through mobile communication devices, has been utilized to successfully improve a variety of health outcomes.

Objective: This study adapted the principles of mHealth to advance breast cancer prevention efforts among Korean American immigrant women, an underserved community.

Methods: Using a randomized controlled trial design, 120 Korean American women aged 40 to 77 years were recruited and randomly assigned to either the mMammogram intervention group (n=60) to receive culturally and personally tailored multilevel and multimedia messages through a mobile phone app along with health navigator services or the usual care control group (n=60) to receive a printed brochure. Outcome measures included knowledge, attitudes, and beliefs about breast cancer screening, readiness for mammography, and mammogram receipt. The feasibility and acceptability of the mMammogram intervention was also assessed.

Results: The intervention group showed significantly greater change on scores of knowledge of breast cancer and screening guidelines (P=.01). The intervention group also showed significantly greater readiness for mammography use after the intervention compared with the control group. A significantly higher proportion of women who received the mMammogram intervention (75%, 45/60) completed mammograms by the 6-month follow-up compared with the control group (30%, 18/60; P<.001). In addition, the intervention group rated satisfaction with the intervention (P=.003), effectiveness of the intervention (P<.001), and increase of knowledge on breast cancer and screenings (P=.001) significantly higher than the control group.

Conclusions: A mobile phone app–based intervention combined with health navigator service was a feasible, acceptable, and effective intervention mechanism to promote breast cancer screening in Korean American immigrant women. A flexible, easily tailored approach that relies on recent technological advancements can reach underserved and hard-to-recruit populations that bear disproportionate cancer burdens.

Trial Registration: Clinicaltrials.gov NCT01972048; https://clinicaltrials.gov/show/NCT01972048 (Archived by WebCite at https://clinicaltrials.gov/archive/NCT01972048/2013_10_29)

(JMIR Mhealth Uhealth 2017;5(11):e154) doi:10.2196/mhealth.7091

KEYWORDS

breast cancer; mammogram; mobile health; mHealth; mobile app intervention; multimedia text message; tailored message; Korean immigrant women; breast cancer disparity

Introduction

Breast cancer remains the most commonly diagnosed form of cancer in women, with approximately 1 in every 8 women in the United States expected to receive this diagnosis in her lifetime [1]. From 2002 to 2011, the incidence of breast cancer increased significantly by 0.8% annually among Asian American and Pacific Islander women, a sharper rise than among any other racial or ethnic group [2]. The Korean American ethnic group constituted 8.2% of the national Asian American population in 2010, increasing by 38.9% from 2000 to 2010 [3]. Although Korean American women tend to have low rates of breast cancer, breast cancer represents the leading type of cancer in Korean women in the United States [4]. In addition, the incidence in foreign-born Korean women demonstrated the greatest increase among Asian American subgroups, at 4% per year [5] between an initial period from 1988-1992 to a second period from 1997-2002 [6]. Accordingly, breast cancer incidence is approximately two times higher in US Korean women compared with native Koreans [6].

Most concerning, Korean Americans have strikingly low rates of cancer screening, including mammography. Breast cancer screening can reduce mortality by detecting cancers at an earlier stage of disease progression when the likelihood of survival is high. Overall, mortality reductions from the routine use of mammograms have been estimated between 19% and 40% [7]. As mammography has shown to effectively detect signs of breast cancer before they can be seen or felt, the American Cancer Society (ACS) recommends annual breast cancer screening for women in the age group of 45 and 54 years, as well as biannual or annual screening depending on a patient's risk for women aged 55 years and older [8]. Whereas there is controversy over the screening interval and age to start screening for mammography [9], the recommendation of screening is warranted in a population with a rising incidence.

In a sample from the California Health Interview Survey in 2003, Korean Americans reported the lowest engagement in nearly every type of cancer screening test, with over half (57.4%) of Korean American women indicating that they had never received a mammogram or that their last mammogram had taken place over a year ago [10]. Among diverse Asian American ethnic groups, Koreans have repeatedly demonstrated the lowest rates of up-to-date mammography screening, ranging from 22% to 57% [10-18]. Even more alarming, in one study [17], the proportion of Korean American women aged 40 years and older who had never engaged in mammography screening was estimated at 85% for those aged 40 to 49 years and 71% for those aged 50 years and older—much higher proportions than any of the other Asian ethnic groups surveyed. The rates of mammography among Korean American women across samples fall well below the Healthy People 2020 target of at least 81.1% of women aged 50 to 74 years having received a mammogram within the past 2 years [19].

Numerous barriers to breast cancer screening among Korean American women have been identified through previous research, which can be categorized as related to health care access, immigration history, and culture. Health care access factors include low rates of health insurance coverage [10,11,14,18,20] and lack of a usual source of care [10,13,14]. As many Korean Americans are foreign born, their recent immigration status may inhibit screening behaviors, in part, because of limited English proficiency [10,11,20-22]. Attitudes influenced by culture present barriers to screening as well. Some Korean women believe there is no risk of getting breast cancer [22], especially if one eats a healthy diet, has no family history of cancer, does not think or worry about it, and has not had multiple sexual partners or abortions [23]. Furthermore, some beliefs include that Korean women only get breast cancer if they work outside the home and do not have time to breastfeed their children [21] or that the development of cancer depends solely on fate [23]. Perceptions of the purpose of seeing a health care provider may also influence screening behavior, including that receiving a mammogram is embarrassing [16] and that it is only necessary to visit a health care provider when ill [16,24]. Older Korean immigrant women have expressed significantly different health beliefs pertaining to breast cancer screening than their younger counterparts [25]. Finally, health literacy, especially knowledge about mammography, strongly impacts Korean American women's screening behaviors [12,20,22,26,27]. In one study, knowledge of screening guidelines emerged as the single most important predictor of regular mammography, with greater knowledge increasing the likelihood of mammography by over 10 times [22]. Many of these barriers are modifiable, including health literacy, health care system factors, and cultural barriers, signaling targets for cancer screening promotion interventions. Factors that facilitate uptake of breast cancer screening serve as valuable targets as well. Among Korean American women, these facilitators include higher perceived benefits [15,26], more confidence in screening techniques [26], greater perceived susceptibility to breast cancer [15], and lower perceived barriers [12,26,28,29].

Due to cultural variations among different Asian ethnic groups, there has been a call to develop tailored approaches to reduce barriers and promote screening [17,21,23], yet little intervention development has successfully addressed this issue in Korean American women. Available evidence suggests that although a number of cancer prevention approaches geared toward Korean American women have been introduced, including interactive

education sessions [12,27,30], a printed brochure [31], and a community intervention with church-based workshops, and financial incentives [24], these interventions have had limited impact on promoting receipt of breast cancer screening. Key reasons behind such limited success include Korean American women being a hard-to-reach population [24,31] and the lack of tailoring to overcome cultural and personal barriers in previous interventions [24]. Designs that have demonstrated positive results tend to be community focused and provide improved access to preventive health care [32,33]. However, studies reporting effective interventions, such as a Korean-language photonovel [34], have sometimes measured only changes in knowledge of screening guidelines or intention to receive a mammogram rather than the actual receipt [12,35]. Interventions that actually improve mammogram receipt, such as a class combined with lay health worker follow-up counseling and navigation assistance [32], tend to be resource-, labor-, and time-intensive with restricted feasibility for widespread dissemination across the nation. In addition, though reservations about screening vary even within a single ethnic group, past interventions have not personally tailored interventions to each individual's concerns.

Addressing the gaps from previous research, this project sought to harness mobile phone technology as a means to enhance preventive health care among the Korean American population. Innovative health interventions increasingly incorporate the use of the Internet for a variety of reasons, including low cost and resource needs, convenience, overcoming the isolation of patients, reducing stigma, and allowing greater user control [36]. Mobile health (mHealth), refers to the use of mobile technology for health information delivery or the improvement of health outcomes [37]. In recent years, mHealth has emerged as a direct and effective medium to change health behaviors, demonstrating success in improving weight loss, metabolic control, blood pressure, diabetes management, stress levels, physical activity, asthma symptoms, medication adherence, hemoglobin A1c levels, smoking cessation, and self-efficacy [38-40]. However, there has been criticism that previous mHealth interventions lacked methodological rigor [41], were not driven by established theories [42], and have rarely been customized to meet the needs of unique individuals [37]. This study incorporated individually and culturally tailored messages into an mHealth intervention, evaluated through a randomized controlled trial (RCT). To our best knowledge, to date, a mobile phone app has not been adapted for mammogram promotion. Shaped by the Fogg behavioral model (FBM) [43] and the concept of persuasive technology [44], this study developed the mMammogram app, a mobile phone app–based intervention designed to motivate Korean American women to undergo an annual mammogram. In response to the fact that all seven of the identified previous intervention studies to promote breast cancer screening in Korean Americans utilized quasi-experimental designs [35], this study employed a novel RCT design with a comprehensive approach that addresses individual, cultural, and system barriers. Through the mobile phone medium, the intervention covered broad content areas and specifically tailored messages to overcome known barriers. This study aims to assess the efficacy of the mMammogram intervention combined with health navigator services, which were designed to motivate Korean American women to undergo breast cancer screening, as compared with the control brochure group. The four hypotheses were as follows: compared with the control group participants, the participants who received the mMammogram app intervention (1) would show greater positive change in knowledge, attitudes, and beliefs about breast cancer screening; (2) would demonstrate greater readiness, or intent, for mammography; (3) would report having received a mammogram at a higher rate; and (4) would express greater acceptance of and satisfaction with the intervention. As no previous study has evaluated a mobile phone–based breast cancer screening intervention in this underserved group, our pilot study sought to provide important insights as to the feasibility and acceptability of the mMammogram intervention, with the ultimate objective to reduce breast cancer disparities by enhancing adherence to screening guidelines.

This study applied Fogg Behavioral Model (FBM) [43], which has originated from persuasive technology [44], to overcome attitudinal and behavioral barriers to screening. Persuasive technology refers to a type of computing system intentionally designed to influence individuals to change maladaptive attitudes or behaviors by giving social cues to elicit certain responses from users [43,45]. The principles of FBM, which have become commonly employed in preventive health care, were utilized to increase self-efficacy and steer a process of change. The FBM explains how persuasive technology can provide an effective mechanism for behavioral change; because behavior is a product of motivation, ability, and triggers, a person must be sufficiently motivated, have the requisite ability, and be appropriately prompted to perform a target behavior [43]. All three factors must simultaneously be present for the new behavior to occur, which can be facilitated through technological devices. The FBM guided the design and development of the mMammogram intervention in first identifying barriers, then creating customized motivators, and finally providing timely triggers. In addition, the health belief model (HBM) [46] provided direction on uncovering the factors, such as perceived susceptibility, perceived severity, perceived benefits, perceived barriers, cues to actions, and level of self-efficacy, to target for change for each individual.

Methods

Study Design

In this two-arm RCT, participants were enrolled and screened for eligibility and informed consent was obtained. All participants then completed the baseline assessment (pretest) through an in-person interview at her preferred place and time before being randomized into the mMammogram intervention group or the usual care control group. No blinding of participants or study personnel was implemented. Control group participants received a printed brochure written in Korean that informs guidelines for breast cancer screening. For participants assigned to the intervention group, the research team downloaded the mMammogram mobile app onto each individual's personal mobile phone or a mobile phone lent to the participant by the research team for the duration of the intervention. The intervention period lasted 1 week with a 6-month follow-up.

Postintervention assessment interviews that utilized an extended version of the baseline questionnaire with additional questions regarding acceptability of the intervention took place at 1 week and 6 months following intervention completion. The 1-week posttest was conducted in person, whereas the 6-month follow-up test was administered via phone. Questionnaires were first developed in English and then translated into Korean using a back-translation method. All interviews were carried out by trained bilingual interviewers experienced in conducting in-person interviews in the Korean language and certified through intensive training, including review of written interview protocols, critical observations, and mock interviews. The institutional review board of the University of Minnesota approved study procedures.

Participant Recruitment, Assignment, and Retention

Using a multipronged recruitment strategy, 149 Korean American women were recruited for participation in this RCT. Eligibility criteria included the following: (1) being a Korean American immigrant woman, (2) aged 40 to 79 years, (3) who had not received a mammogram in the past 2 years, (4) lived in Minnesota, (5) possessed an active email account, and (6) were willing to use their own mobile phone or a mobile phone borrowed from the research team for the mobile app intervention. The exclusion criteria included those who (1) were born in the United States or immigrated to the United States as minors (under 18 years), (2) received a mammogram in the past year, and (3) aged under 40 or 80 years and older. Participants were recruited using flyers and brochures in the Korean language that were distributed to churches, temples, clinics, social service agencies, ethnic community centers, beauty salons, and ethnic markets serving the Korean American community. These materials specified the purpose of the project, eligibility criteria, and study personnel contact information. Members of the research team also made presentations at Korean churches and community centers and generated coverage in the Korean American ethnic press.

To obtain an adequate sample size, the project aimed to enroll 150 women with 75 in each arm, assuming an 80% retention rate. It was anticipated that a two-sided two-sample *t* test would be used at the conventional 5% type I error rate and 80% statistical power. A final sample size of 60 in each arm would allow the detection of an effect size of approximately 0.5, a difference in the average score equal to half the standard deviation (SD), conventionally considered a large trial. As previous research enabled an assumption that 20% of the control population would receive mammograms, a group size of 60 allowed the detection of difference in the proportions at 25% using a chi-square test at a 5% type I error rate and 80% statistical power.

After enrolling 149 participants, 144 provided informed consent and completed the pretest. Before the next phase, 13 participants were automatically released from participation after realizing they had received a mammogram in the past 2 years, rendering them ineligible. A total of 131 participants were randomized to the intervention and control groups by an approximately 1:1 ratio (intervention: n=68; control: n=63). The method of sequentially numbered, opaque sealed envelopes (SNOSE) was used for randomization [47]. Sealed identical envelopes were given to participants with a code designating intervention or control group written on a piece of paper on the inside; there were no detectable differences between the envelopes. Over the intervention period, 3 participants dropped out from each group (intervention: 2 loss of contact, 1 cognitive impairment; control: 1 refused, 1 not eligible [remembered receipt of mammogram within past 2 years], and 1 incomplete data). Although 65 participants in the intervention group and 60 participants in the control group completed the intervention period and all measures, 60 participants from each group were analyzed, as 5 participants in the intervention group were dropped from the analyses because of ineligibility. The 5 participants reported that they actually received the mammogram in the past year; the research intervention reminded them of the receipt of the mammogram when watching a video of a mammogram procedure. In sum, among the initially recruited 149 participants, 19 participants were screened out from the study before or after the intervention because of ineligibility. Among the remaining 130 participants, 10 participants left the study, thus yielding a 7.7% attrition rate. Each participant received US $20 for each face-to-face interview, plus US $20 reimbursement for text message data fees over the 6-month period in the intervention group.

Community Advisory Board

Drawing on a community-based participatory research approach, a community advisory board (CAB) was formed to provide guidance throughout the process of study development, execution, and dissemination of research findings. The CAB consisted of 5 members of the local Korean American community, including representatives from the Korean Service Center; Korean American Association; Korean American Women's Association in Minnesota, a university student group; and a Korean ethnic church. In bimonthly meetings, members of the CAB provided input in generating the format and content of text and multilevel and multimedia messages, ensuring cultural relevance. The CAB also assisted in devising strategies for participant recruitment and retention, enhancing the accessibility of the website, interpreting preliminary findings, and suggesting approaches for dissemination in the community.

mMammogram Intervention Development

The process of development for the mobile phone app, mMammogram, involved five main steps: (1) forming a CAB, (2) identifying barriers and mobile phone usage patterns and preferences, (3) creating motivators, (4) tailoring message content, and (5) developing appropriate triggers. After CAB members had been identified, a series of focus groups with Korean American women in their 40s and 50s were conducted to ascertain barriers, motivators, and mobile phone usage patterns. Each session lasted 1.5 to 2 hours, during which participants discussed their current knowledge of breast cancer and screening guidelines; individual, structural, and cultural barriers to screening; current mobile phone usage habits, including text and picture messaging; short message service and multimedia messaging service subscriptions; and ideas regarding the most effective content, type, and frequency of messages to promote screening.

Utilizing data from the focus groups, feedback from the CAB, and input from persuasive technology consultants, the content of the text, multimedia messages, and follow-up schedule were designed and finalized. Special emphasis was given to cultural health beliefs and misconceptions about breast cancer screening, such as the assumption that the absence of symptoms means good health, profound embarrassment related to physical exams, and fatalistic views of cancer. The system was designed to be both personally tailored and interactive, with the content, number, and timing of daily messages adapted to each individual. To keep messaging fresh and nonrepetitive, a database was generated with an ample amount of messages so that the type and content of messages could be varied over the week. The overall computer system for the intervention consisted of five components: (1) a Web-based application to enroll participants, set user preferences, display the global positioning system (GPS) navigation system with area clinic information, and upload text and multimedia messages; (2) a database to store participant records, rules, and messages sent and received; (3) a program to establish the appropriate timing of messages, determine which messages to send, and process received replies; (4) a text-message delivery or reception platform; and (5) a health navigator for assistance navigating cancer screening information, addressing technical problems, and providing transportation and interpretation services. The system also had tools enabling continuous technical monitoring to recognize anomalies that might indicate an individual was having difficulties with the mobile app. In these cases, the health navigator contacted the participant to prevent user frustration and increase adherence and satisfaction.

A series of three usability tests of the mMammogram system prototype were conducted with 5 focus group participants before the RCT, with feedback incorporated into the final mobile app. Each participant was asked to describe her evaluations of the wording of text messages and delivery of accurate information, quality and length of the videos, ease of message delivery, quality of emoticons and the appropriateness of their locations in text message sequences, quantity of interactive messages each day and difficulty responding to each question, overall length of messaging each day, technical problems, and their general impressions of the app for learning about breast cancer screening. On the basis of this feedback, the app was revised. The second and third usability tests were conducted in a similar manner with the same 5 participants, leading to further refinement.

At the outset of the RCT, following initial recruitment, pertinent information about each participant was collected during baseline face-to-face interviews regarding current knowledge of breast cancer, structural or cultural barriers to screening, level of intention to receive a mammogram, and personal preferences around SMS and MMS. In addition, a true or false questionnaire was employed to assess each participant's personal risk for breast cancer. Participants' responses in interviews and to the questionnaire were used to tailor messages to each individual.

The actual intervention was delivered in Korean over a 7-day period. Each day we sent 8 to 21 messages to participants over the course of the 7-day intervention. In the last text message of each day, the specially designed mMammogram logo was included to symbolize the conclusion of the intervention for that day. The week-long program allowed sufficient time to highlight various topical areas, including breast cancer, screening guidelines, and types of screening; breast cancer risk factors; individual, structural, and cultural barriers to screening; communication strategies; follow-up for test results; and information on local clinics. Messages followed a trajectory from basic knowledge building to specific strategies aimed to enhance motivation for and access to mammography. Approximately half of the messages requested a reply, providing a balance between education and motivation. An incentive system was employed to increase participant engagement; for each response to a question or a prompt, regardless of whether a participant answered a knowledge question correctly, she could earn a digital pink ribbon and collect these ribbons throughout the intervention period. Recognizing that visual messages can be particularly persuasive, some messages included illustrations, reference photos, and video clips. Video messages featured, for example, Korean American women sharing their personal experiences with mammogram screening, including how they have handled issues related to their cultural beliefs.

To increase accessibility to screening services, a website was created containing a list of area clinics and indicating those that provide free or lost-cost mammograms. All participants received a link to the website that could be accessed by a mobile phone or computer. The list provided information about all clinics, including office hours, types of health insurance accepted, possible free or low-cost options, and physician profiles. In addition, an embedded GPS navigating system allowed participants to determine the distance of clinics from their residence and directions to their chosen clinic. At the end of topic-based message sequences, participants were sent questions as triggers to set up appointments for a mammogram. In order for a trigger to be effective, it had to be noticed, associated with the target behavior, and sent at a suitable time. Participants were sent triggers such as, "Would you like a list of clinics in your area that offer screening?" Those who replied *yes* were sent links to the website with the customized contact information for local clinics and a message with a motivational statement such as, "Call today for an appointment!" A health navigator was available to assist participants in obtaining the necessary resources, appointments, or transportation to receive a mammogram.

Control Condition

Participants assigned to the control group received usual care that consisted of the mailing of printed materials in the Korean language with contact information of health navigator for questions regarding information we provided in the brochure. This approach has traditionally been used by ethnic health service agencies to promote cancer screening. The materials included a brochure with information on breast cancer and relevant screening guidelines from the ACS, as well as a list of community clinics, indicating those that offer low-cost or free mammography. The control group completed the same assessment schedule (baseline, 1 week post intervention, and monthly follow-up test) with the exclusion of the intervention.

Measures

Mammogram receipt was the primary outcome measure, whereas breast cancer knowledge, health beliefs, cultural attitudes, level of intention, and participant's satisfaction and opinion about the effectiveness of the intervention constituted secondary outcome measures. Control variables included sociodemographic characteristics (age; educational attainment; employment, income, and financial status; marital status; family members; and residence); family cancer history; health status; health care access; immigration information; lifestyle variables related to exercise, drinking, and smoking; and past breast cancer screening experiences. Outcome measures were operationalized as follows:

Mammography Receipt

Mammography receipt or a scheduled appointment after the intervention was collapsed into one variable and assessed through self-report (yes or no), which has been found to be reliable in cancer screening research [48]. Participants' mammography receipt was tracked for 6 months after the intervention (up to follow-up period).

Breast Cancer Knowledge

Breast cancer knowledge was measured by the breast cancer knowledge test [49], which has been validated with women from diverse cultural groups [50-52]. The test was revised to reflect current ACS breast cancer screening guidelines. The final knowledge scale consisted of 28 true or false items, and the score was computed by the number of items the participant answered correctly. The internal consistency for the present sample was acceptable (alpha=.77 for the pretest, alpha=.75 for the posttest).

Health Beliefs

Health beliefs were measured by Champion's health belief model (HBM) scale [53,54]. Items in the HBM scale map to three main variables used in this study: perceived susceptibility (3 items), perceived benefits (5 items), and perceived barriers (11 items). The scales have demonstrated high reliability and validity in the past, with ethnically diverse sample populations [55-58]. The full list of health beliefs assessed consisted of perceived susceptibility (3 items), perceived benefits (5 items), and barriers to receiving mammogram (16 items), as well as prevention orientation (5 items), self-efficacy of breast cancer screening (8 items), and distrust toward health professionals (5 items). All items were on a 4-point scale ranging from strongly disagree to strongly agree or from unconfident to confident. Higher item scores were indicative of stronger belief for the given construct, and scores of each construct were computed by the sum of item scores. The internal consistencies for the present sample was as follows: perceived susceptibility: alpha=.87 for the pretest, alpha=.73 for the posttest; perceived benefits: alpha=.70 for the pretest, alpha=.75 for the posttest; barriers to receiving mammogram: alpha=.89 for the pretest, alpha=.90 for the posttest; prevention orientation: alpha=.40 for the pretest, alpha=.55 for the posttest; self-efficacy of breast cancer screening: alpha=.89 for the pretest, alpha=.93 for the posttest; and distrust toward health professional: alpha=.72 for the pretest, alpha=.70 for the posttest.

Cultural Beliefs and Attitudes

Cultural beliefs and attitudes toward breast cancer screening were captured through 6 items from Tang et al's inventory [59] of cultural barriers to screening among Asian American women and 3 items regarding fatalism from a questionnaire developed by Taylor et al [60]. Besides fatalism (3 items), other attitudes measured were modesty (5 items), social support (6 items), and fear of discovery (1 item). All items were on a 4-point scale ranging from strongly disagree to strongly agree, with higher item scores indicative of stronger belief for the given construct. Scores of each construct were computed by the sum of item scores. The internal consistencies for the present sample was as follows: modesty: alpha=.70 for the pretest, alpha=.69 for the posttest; and social support: alpha=.57 for the pretest, alpha=.72 for the posttest. The internal consistency for the fear of discovery is not computable because it is a single-item scale.

Level of Intention

Level of intention to obtain a mammogram was informed by the stages of change in the transtheoretical model [61], which suggests that people move through a series of progressively more committed stages toward adoption of a new behavior. Adapting the stages of change to intention for mammography, participants were asked to indicate their level of intention to receive a mammography in the future on a 4-point scale (1=not within a year, 2=within a year, 3=within 3 months, and 4=within 1 month). One week after the intervention, the intention was reassessed among participants who had not received a mammography since participation in the study.

Participant Satisfaction

Participant satisfaction regarding the intervention they received was assessed using a 4-point scale item ranging from very dissatisfied to very satisfied 1 week after the intervention. In addition to general satisfaction, participants' willingness to recommend the intervention they received and intention to receive a mammography after this study were also measured using yes-or-no items 1 week after the intervention.

Intervention Effectiveness

Intervention effectiveness was measured by a 4-point scale item ranging from very ineffectual to very effectual. In addition to the general effectiveness of the intervention, participants' perceived level of knowledge about mammography was measured on a 3-point scale item (1=same, 2=improved, and 3=very improved).

Data Analysis

The data analysis included 60 participants in the intervention (ie, mMammogram app) group and 60 in the control group (ie, brochure) who completed pre- and posttest questionnaires. Before addressing proposed hypotheses, group equivalence in terms of baseline characteristics (ie, sociodemographics, family cancer history, health status, health care access, and past breast cancer screening experiences) was examined using *t* test and chi-square tests. For hypotheses 1 and 2, group equivalence at the pretest was first examined using the two-sample *t* test. Then, group differences in terms of changes in the given constructs were tested using a mixed-effect analysis of variance (ANOVA).

The mixed-effect ANOVA includes both within-subject (ie, repeated measures) and between-subject factors (ie, independent variable for which participants are assigned to one of the different conditions) and aims to examine whether there is an interaction between these two factors on the dependent variable. In the context of this study, time (pre- and posttest) represented the within-subject factor, whereas group (app vs brochure) represented the between-subject factor. For hypothesis 3, the percentage of participants from each arm who received mammograms or scheduled an appointment was compared using the chi-square test. Finally, for hypothesis 4 **,** averages of general satisfaction and effectiveness scores from each group were compared using the two-sample *t* test. Also, the percentage of participants from each group who endorsed *yes* for the intention and recommendation items were compared using the chi-square test. All the data were analyzed using the Statistical Package for the Social Sciences Statistics version 22 (IBM Corp).

Results

Sociodemographics of the Sample

Tables 1 and 2 summarize sociodemographics for continuous and categorical variables, respectively. The mean age of all participants was 51.60 years (SD 9.55). On average, they had lived in the United States for 18.43 years (SD 10.80), and their mean age at the time of immigration to the United States was 33.5 years (SD 8.76). In terms of educational background, participants had received an average of 15.14 years of education (SD 3.27), and 72.5% (87/120) of participants reported completion of college or university or beyond. With regard to employment and income, about half of the participants (50.8%, 61/120) were currently employed, and 42.5% (51/120) reported their household monthly income including tax as US $7000 or more. Most participants (86.7%, 104/120) reported that their financial condition was fair, good, or very good. With regard to families and residences, the majority of participants were married or cohabitating (86.7%, 104/120) and were living with their spouse or children (90.0%, 108/120). The majority (90.0%, 108/120) also lived in their own or leased house or condominium. In terms of current health conditions and health-related behaviors, 36.7% (44/120) of participants reported to be in good or very good health, and 70.8% (85/120) of the participants reported to exercise at least once a week. In addition, the majority of participants were nonsmokers (95.8%, 115/120) and nondrinkers (80.8%, 97/120). More importantly, app and brochure groups were not significantly different in these baseline characteristics, as indicated by insignificant *t* test and chi-square test results.

Experience of Breast Cancer Screening Before Intervention

Table 3 summarizes participants' previous experiences related to three types of breast cancer screening: breast self-examination (BSE), clinical breast examination (CBE), and mammography. When asked about their awareness of the given screening exams at baseline assessment, 92.5% (111/120), 57.5% (69/120), and 77.5% (93/120) of participants responded that they had heard of the BSE, CBE, and mammography, respectively. In terms of procedure knowledge (ie, whether participants knew how the given screening exam is performed), 81.7% (98/120), 53.3% (64/120), and 75.8% (91/120) of participants reported to know the procedures of the BSE, CBE, and mammography, respectively. In addition, 78.3% (94/120), 61.7% (74/120), and 70.0% (84/120) of participants reported that they had previously performed or received the BSE, CBE, and mammogram, respectively. The rate of having performed or received each screening exam at least once every 6 months was 45.8% (55/120), 61.7% (74/120), and 70.0% (84/120) for the BSE, CBE, and mammogram, respectively. Finally, time since the last performance or receipt of each screening exam was on average 0.91 years (SD 1.83), 3.59 years (SD 4.06), and 4.30 years (SD 4.05) for the BSE, CBE, and mammography, respectively. The app and brochure groups were not significantly different in any of these previous experiences as indicated by insignificant chi-square and t test results.

Table 1. Sociodemographics for continuous variables by group.

Variable	App (N=60), mean (SD[a])	Brochure (N=60), mean (SD)	All (N=120), break/>mean (SD)	Group difference *t* (degrees of freedom)	*P* value
Age, in years	51.38 (8.74)	51.82 (10.36)	51.60 (9.55)	−0.25 (118)	.81
Years living in the United States	17.90 (9.65)	18.97 (11.89)	18.43 (10.80)	−0.54 (118)	.59
Age at the time of immigration to the United States	33.80 (9.78)	33.19 (7.68)	33.50 (8.76)	0.38 (118)	.71
Years of education	15.00 (2.73)	15.28 (3.75)	15.14 (3.27)	−0.47 (118)	.64

[a]SD: standard deviation.

Table 2. Sociodemographics for categorical variables by group.

Variable	App (N=60), n (%)	Brochure (N=60), n (%)	All (N=120), n (%)	Group difference χ^2 (degrees of freedom)	*P* value
Highest level of education					
Middle school and less	4 (7)	5 (8)	9 (7.5)	2.7 (3)	.44
Completed high school	10 (17)	14 (23)	24 (20.0)		
Completed college or university	36 (60)	27 (45)	63 (52.5)		
Completed graduate school	10 (17)	14 (23)	24 (20.0)		
Employment					
No	28 (47)	31 (52)	59 (49.2)	0.3 (1)	.58
Yes	32 (53)	29 (48)	61 (50.8)		
Household monthly income in US dollars (including tax)					
Up to US $2999	13 (22)	15 (25)	28 (23.3)	1.3 (3)	.74
US $3000-$6999	21 (35)	19 (32)	40 (33.3)		
US $7000-$11,999	16 (27)	13 (22)	29 (24.2)		
US $12,000 or more	9 (15)	13 (22)	22 (18.3)		
Financial status					
Very bad	1 (2)	2 (3)	3 (2.5)	1.3 (4)	.86
Bad	5 (8)	7 (12)	12 (10.0)		
Fair	40 (67)	35 (58)	75 (62.5)		
Good	13 (22)	13 (22)	26 (21.7)		
Very good	1 (2)	2 (3)	3 (2.5)		
Marital status					
Single	3 (5)	2 (3)	5 (4.2)	1.1 (3)	.78
Married or cohabitating	51 (85)	53 (88)	104 (86.7)		
Separated or divorced	4 (7)	2 (3)	6 (5.0)		
Widowed	2 (3)	3 (5)	5 (4.2)		
Family members					
Alone	6 (10)	6 (10)	12 (10.0)	1.1 (3)	.78
With spouse	13 (22)	17 (28)	30 (25.0)		
With spouse and children	36 (60)	34 (57)	70 (58.3)		
With children (no spouse)	5 (8)	3 (5)	8 (6.7)		
Type of residence					
Own house or condominium	42 (70)	37 (62)	79 (65.8)	3.6 (4)	.46
Leased house or condominium	11 (18)	18 (30)	29 (24.2)		
Public housing	5 (8)	2 (3)	7 (5.8)		
Rented room	1 (2)	2 (3)	3 (2.5)		
Others	1 (2)	1 (2)	2 (1.7)		
Current health condition					
Very bad	1 (2)	0 (0)	1 (0.8)	7.7 (4)	.10
Bad	7 (12)	9 (15)	16 (13.3)		
Fair	29 (48)	30 (50)	59 (49.2)		
Good	13 (22)	19 (32)	32 (26.7)		
Very good	10 (17)	2 (3)	12 (10.0)		

Variable	App (N=60), n (%)	Brochure (N=60), n (%)	All (N=120), n (%)	Group difference χ^2 (degrees of freedom)	*P* value
Number of exercises per week					
0	19 (32)	16 (27)	35 (29.2)	1.6 (4)	.81
1-2	19 (32)	22 (37)	41 (34.2)		
3-4	15 (25)	18 (30)	33 (27.5)		
5-6	5 (8)	3 (5)	8 (6.7)		
7+	2 (3)	1 (2)	3 (2.5)		
Smoking					
No	57 (95)	58 (97)	115 (95.8)	.[a] (1)	>.99
Yes	3 (5)	2 (3)	5 (4.2)		
Drinking					
No	51 (85)	46 (77)	97 (80.8)	1.4 (1)	.25
Yes	9 (15)	14 (23)	23 (19.2)		

[a]Dot signifies that no numeric value is available. Instead of Pearson chi-square test, Fisher exact test was performed, given that the expected count for some cells is less than 5. In the Statistical Package for the Social Sciences (SPSS) software, only the *P* value of the Fisher exact test is reported rather than the test statistic.

Table 3. Summary of previous experience of breast cancer screening by group.

Screening	App (N=60), n (%)	Brochure (N=60), n (%)	All (N=120), n (%)	Group difference χ^2 (degrees of freedom)	*P* value
BSE [a]					
Awareness					
Yes	54 (90)	57 (95)	111 (92.5)	. [b] (1)	.49
No	6 (10)	3 (5)	9 (7.5)		
Procedure knowledge					
Yes	48 (80)	50 (83)	98 (81.7)	0.2 (1)	.64
No	12 (20)	10 (17)	22 (18.3)		
Previous performance					
Yes	49 (82)	45 (75)	94 (78.3)	0.8 (1)	.38
No	11 (18)	15 (25)	26 (21.7)		
Years since the last performance, mean (SD[c])	1.04 (2.24)	0.78 (1.28)	0.91 (1.83)	−1.15 (90)[d]	.25
CBE [e]					
Awareness					
Yes	34 (57)	35 (58)	69 (57.5)	0.0 (1)	.85
No	26 (43)	25 (42)	51 (42.5)		
Procedure knowledge					
Yes	29 (48)	35 (58)	64 (53.3)	1.2 (1)	.27
No	31 (52)	25 (42)	56 (46.7)		
Previous receipt					
Yes	37 (62)	37 (62)	74 (61.7)	0.0 (1)	>.99
No	23 (38)	23 (38)	46 (38.3)		
Years since the last receipt, mean (SD)	4.22 (4.22)	2.97 (3.86)	3.59 (4.06)	−1.51 (72)[d]	.13
Mammography					
Awareness					
Yes	47 (78)	46 (77)	93 (77.5)	0.1 (1)	.83
No	13 (22)	14 (23)	27 (22.5)		
Procedure knowledge					
Yes	45 (75)	46 (77)	91 (75.8)	0.1 (1)	.83
No	15 (25)	14 (23)	29 (24.2)		
Previous receipt					
Yes	44 (73)	40 (67)	84 (70.0)	0.6 (1)	.43
No	16 (27)	20 (33)	36 (30.0)		
Years since the last receipt, mean (SD)	4.23 (3.15)	4.38 (4.89)	4.30 (4.05)	−0.02 (82)[d]	.98

[a]BSE: breast self-examination.

[b]Dot signifies that no numeric value is available. Instead of Pearson chi-square test, Fisher exact test was performed, given that the expected count for some cells is less than 5. In the Statistical Package for the Social Sciences (SPSS) software, only the *P* value of the Fisher exact test is reported rather than the test statistic.

[c]SD: standard deviation.

[d]Signifies *t* (degrees of freedom).

[e]CBE: clinical breast examination.

Change in Knowledge, Attitudes, and Beliefs About Breast Cancer and Screening After the Intervention (Hypothesis 1)

Table 4 summarizes the changes in knowledge, attitudes, and beliefs about breast cancer screening by group as well as results of mixed-design ANOVA . Independent-samples *t* test results were insignificant at the pretest, indicating that the two groups were not significantly different in any of these constructs at baseline. Demonstrating within-subjects effects, participants showed significant improvement in knowledge ($F_{1,118}$=209.74, *P*<.001, effect size=0.64), reduction in fatalism ($F_{1,118}$=15.19 , *P*<.001, effect size=0.11), increased perceived benefits ($F_{1,118}$=20.16 , *P*<.001, effect size=0.15), and increased self-efficacy ($F_{1,118}$=18.79 , *P*<.001, effect size=0.14) related to breast cancer and screening. However, none of the between-subjects effects were significant, indicating that for both the pre- and posttests, the participants' scores on knowledge, attitudes, and beliefs about breast cancer and screening were not substantially different between the app and brochure groups. With regard to interaction between time and group, only the knowledge construct was found to be significant ($F_{1,118}$=6.24, *P*=.01, effect size=0.05) indicating that the increase in knowledge between pre- and posttest was significantly larger for the app group compared to the brochure group. The nature of this interaction is displayed in Figure 1.

Table 4. Summary of change in knowledge, attitude, and belief about breast cancer screening by group.

Variables	App (N=60)		Brochure (N=60)			Group difference	All (N=120)		Mixed-design ANOVA[a]	
	Pretest, mean (SD[b])	Posttest, mean (SD)	Pretest, mean (SD)	Posttest, mean (SD)	*t* (degrees of freedom)	Pretest, mean (SD)	Posttest, mean (SD)	Within-group, *F*[c] (effect size[d])	Between-group, *F* (effect size)	Interaction time × group, *F* (effect size)
Knowledge on breast cancer and screening	16.55 (5.13)	22.95 (3.52)	17.32 (4.38)	21.83 (3.48)	−0.88 (118)	16.93 (4.77)	22.39 (3.53)	209.74[e] (0.64)	0.069 (0.00)	6.24[f] (0.05)
Barriers to receiving mammography	27.22 (8.69)	26.10 (7.27)	26.60 (6.43)	25.88 (6.63)	0.44 (118)	26.91 (7.62)	25.99 (6.93)	2.66 (0.02)	0.12 (0.00)	0.13 (0.00)
Distrust of health professionals	10.10 (2.47)	9.57 (1.82)	10.23 (2.27)	9.98 (1.94)	−0.31 (118)	10.17 (2.36)	9.78 (1.88)	3.66 (0.03)	0.68 (0.01)	0.48 (0.00)
Fatalism	6.28 (1.54)	5.52 (1.24)	5.88 (1.56)	5.48 (1.61)	1.41 (118)	6.08 (1.56)	5.50 (1.43)	15.19[e] (0.11)	0.90 (0.01)	1.50 (0.01)
Fear of discovery	1.55 (0.67)	1.65 (0.66)	1.52 (0.62)	1.52 (0.57)	0.28 (118)	1.53 (0.65)	1.58 (0.62)	0.78 (0.01)	0.69 (0.01)	0.78 (0.01)
Modesty	10.82 (2.72)	11.05 (2.56)	10.87 (2.71)	10.63 (2.23)	−0.10 (118)	10.84 (2.70)	10.84 (2.40)	0.00 (0.00)	0.20 (0.00)	1.09 (0.01)
Perceived benefits	15.12 (2.85)	16.42 (2.12)	15.07 (1.89)	15.82 (2.21)	0.11 (118)	15.09 (2.40)	16.12 (2.18)	20. 16[e] (0.15)	0.86 (0.01)	1.45 (0.01)
Perceived susceptibility	5.62 (2.03)	5.43 (1.37)	5.32 (1.14)	5.35 (1.12)	1.00 (118)	5.47 (1.64)	5.39 (1.25)	0.27 (0.00)	0.73 (0.01)	0.56 (0.01)
Prevention orientation	15.77 (1.97)	16.00 (2.17)	15.98 (1.93)	15.93 (2.15)	−0.61 (118)	15.88 (1.94)	15.97 (2.15)	0.24 (0.00)	0.05 (0.00)	0.57 (0.01)
Self-efficacy on breast cancer screening	23.82 (4.54)	25.20 (4.27)	22.98 (3.98)	24.53 (4.06)	1.07 (118)	23.40 (4.27)	24.87 (4.16)	18.79[e] (0.14)	1.18 (0.01)	0.06 (0.00)
Social support	15.77 (2.94)	16.23 (3.03)	15.65 (2.85)	15.98 (2.70)	0.22 (118)	15.71 (2.89)	16.11 (2.86)	3.07 (0.03)	0.15 (0.00)	0.09 (0.00)

[a]ANOVA: analysis of variance.

[b]SD: standard deviation.

[c]Degrees of freedom for *F* test=1,118.

[d]For the effect size, partial eta squared (η_p^2) was computed.

[e]*P*<.001.

[f]*P*<.05.

Figure 1. Interaction between time and group.

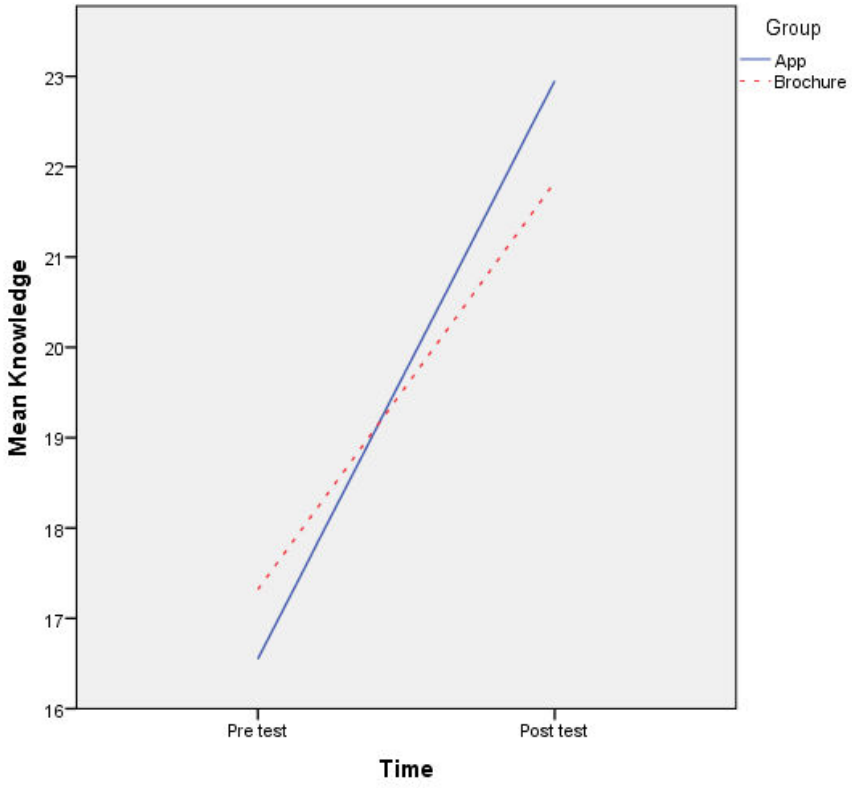


Table 5. Intention for mammography use by group.

Response	App (N=60)		Brochure (N=60)		All (N=120)		Group difference	
	Pretest, n (%)	Posttest, n (%)	Pretest, n (%)	Posttest, n (%)	Pretest, n (%)	Posttest, n (%)	Pretest, *t* (degrees of freedom)	Posttest, *t* (degrees of freedom)
No plan to do within 1 year	8 (13)	9 (15)	10 (17)	10 (17)	18 (15.0)	19 (15.8)	−0.64 (118)	3.48[a] (118)
Plan to do within 1 year	8 (13)	24 (40)	8 (13)	40 (67)	16 (13.3)	64 (53.3)		
Plan to do within the next 3 months	0 (0)	12 (20)	1 (2)	7 (12)	1 (0.8)	19 (15.8)		
Plan to do within 1 month	0 (0)	14 (23)	1 (2)	1 (2)	1 (0.8)	15 (12.5)		

[a]*P*<.01.

Change in Intention to Receive Breast Cancer Screening After the Intervention (Hypothesis 2)

Table 5 summarizes intention for mammography use by group. Independent-samples *t* test results were insignificant at the pretest, indicating that intention for mammography use was not substantially different between groups at the pretest (t_{118}=−0.64, *P*=.53). Therefore, instead of mixed-design ANOVA , an independent-samples *t* test was performed solely for the posttest scores to examine group differences in intention for mammography use after the intervention. As shown in Table 5, a significant group difference was found (t_{118}=3.48, *P*=.001) with a combined sample size of 120.

Receipt of Mammography After the Intervention (Hypothesis 3)

The app group was found to receive mammograms significantly more than the brochure group after the intervention, as indicated by chi-square test results (χ^2_1=24.4, *P*<.001). Specifically, 75% (45/60) of app group participants versus 30% (18/60) of brochure group participants received a mammogram after the intervention. About 8% (5/60) of participants in the app group received a mammogram through health navigators who arranged appointments and provided transportation and interpretation services, whereas 33% (20/60) received a mammogram at the Mammo a-go-go program, a free mammogram event that health navigator arranged in conjunction with a local health care system for participants who did not have health insurance or were underinsured. The rest of the participants (20/60, 33%) received a mammogram by themselves without a health navigator's help.

Table 6. Satisfaction, effectiveness, and acceptability of the intervention by group.

Variable	App (N=60) n (%)	Brochure (N=60) n (%)	All (N=120) n (%)	Group difference *t* (*P* value)
Effectiveness				
Very ineffectual	2 (3)	1 (2)	3 (2.5)	3.73 (<.001)
Ineffectual	0 (0)	2 (3)	2 (1.7)	
Effectual	26 (43)	49 (82)	75 (62.5)	
Very Effectual	32 (53)	8 (13)	40 (33.3)	
Increase of knowledge				
Same	0 (0)	4 (7)	4 (3.3)	3.52 (.001)
Improved	38 (63)	48 (80)	86 (71.7)	
Very improved	22 (37)	8 (13)	30 (25.0)	
Satisfaction with intervention				
Very dissatisfied	0 (0)	0 (0)	0 (0.0)	3.03 (.003)
Dissatisfied	0 (0)	1 (2)	1 (0.8)	
Satisfied	36 (60)	49 (82)	85 (70.8)	
Very satisfied	24 (40)	10 (17)	34 (28.3)	
Intention to receive a mammography in the future				
Yes	57 (95)	54 (90)	111 (92.5)	.[a] (.49)
No	3 (5)	6 (10)	9 (7.5)	
Recommendation of mammography				
Yes	59 (98)	55 (92)	114 (95.0)	.[a] (.21)
No	1 (2)	5 (8)	6 (5.0)	

[a]Dot signifies that no numeric value is available. Instead of Pearson chi-square test, Fisher exact test was performed given that the expected count for some cells is less than 5. In Statistical Package for the Social Sciences software, only the *P* value of the Fisher exact test is reported rather than the test statistic.

Satisfaction With and Effectiveness of the Intervention (Hypothesis 4)

To examine group differences in satisfaction with and effectiveness of the intervention, independent-samples *t* test and chi-square test were performed for Likert-type items and dichotomous items, respectively. As shown in Table 6, compared with the brochure group, the app group reported significantly higher ratings on perceived effectiveness of the intervention (t_{118}=3.73, *P*<.001), increase in knowledge (t_{118}=3.52, *P*=.001), and satisfaction with the intervention (t_{118}=3.03, *P*=.003). Although the app group also expressed greater intention to receive a mammogram in the future when it is due (95%, 57/60 vs 90%, 54/60) and were more willing to recommend the intervention they received to their friends (98%, 59/60 vs 92%, 55/60) compared with the brochure group, these differences were not statistically significant.

Discussion

Principal Findings

This pilot study offers initial evidence for the feasibility and effectiveness of a mobile app intervention with health navigation services as compared with the control group to increase participation in mammography among Korean American women, a hard-to-reach community with low rates of breast cancer screening. The main finding that the intervention group received mammograms at a significantly higher rate than the control group highlights breast cancer screening as another area in which innovative mobile phone app interventions can positively influence health behaviors. As such, this study further diversifies the list of outcomes shown to be compatible with an mHealth approach, contributing to a list that already includes fatigue [62], diabetes management [63], blood pressure control [64], and physical activity [65], among others [66].

Part of the differential effect of the intervention may be explained by the intervention group's higher ratings of perceived effectiveness and satisfaction with the mobile app with health navigation services compared with the control group's perspectives regarding the brochure. The substantial positive association between participants' perceptions of effectiveness and the actual effect of the intervention on attitudes and behaviors has been previously established through meta-analysis, with evidence that perceived effectiveness stands as a causal influence for actual effectiveness [67,68]. There is also strong evidence that patient satisfaction with care (*care* represented in this case by the mobile app intervention) impacts clinical

outcomes, including adherence to recommended behavior regimens and use of preventive care services [69]. Beyond greater perceived effectiveness and satisfaction with the intervention, the mobile app group exhibited greater increases in their knowledge levels about breast cancer screening than the control group, suggesting superior effectiveness of the mobile app for education on this topic. The link between knowledge of risk factors and screening procedures and receipt of breast cancer screening has been well established [70-72]. In addition, this finding supports the proliferation of mobile phone apps for the delivery of cancer-related information, including the effort to ensure the inclusion of scientifically validated data [73-75].

Despite an overall promising result, several of the secondary outcomes demonstrated a lack of significant differences between the intervention and control groups, contrary to hypotheses. Both the participants who received the mobile app intervention with health navigation services and those who received the usual care brochure demonstrated gains in attitudes toward screening and beliefs about barriers, self-efficacy, and health professionals over the study period; however, these gains were roughly equivalent between groups. Similarly, the percentage of women in each group who intended to receive a mammogram approximately doubled following the intervention period; however, the intervention group showed significantly higher readiness for mammography. Potential contributors to the lack of substantial differences in these outcomes include the influence of social desirability and potential contamination or spillover effects. Not only has research shown that women tend to score higher on measures of social desirability bias [76-78] but that participants from Asian cultures may score higher than groups of European descent as well [79,80]. Participants in both groups may have felt pressure to respond in positive ways on posttest measures. In addition, because participants were recruited from a local community and may have interacted outside the study context, participants in the control condition may have indirectly been exposed to contents from the intervention. The difficulty of preventing contamination has been cited as a unique barrier to conducting RCTs among Asian American populations in cancer screening research [81].

Limitations

The interpretation of findings from this study should take certain limitations into consideration. First, some of the construct measures had low reliability, as demonstrated by coefficient alphas below .7 for prevention orientation, fatalism, and social support. Because unreliable scales decrease the statistical power of instruments, the low reliability in these measures may have contributed to the insignificant differences found between the intervention and control groups. Second, there are multiple potential confounding factors that influenced mammogram receipt in both groups, including the monthly phone calls to check receipt of mammogram over the 6-month follow-up period, a sense of obligation that participants may have felt based on the rapport developed with the research team, and the pressure to comply with expectations based on a Korean cultural norm that makes women reluctant to give a direct negative response to a request. Third, the provision of health navigator services (eg, providing interpretation services and transportation services and arranging a free mammogram event such as Mammo-a-go-go program) to the intervention group may have been responsible for part of the differential effect of the intervention on the primary and secondary outcomes. However, for an immigrant group that lacks English proficiency and health care accessibility, provision of health navigation services is critical, combined with mobile app program to promote mammography. This study design renders it impossible to parse out the additive effect of these factors. Future studies, therefore, should use a three-arm design (app vs app plus health navigation services vs usual care) to tease out the pure effectiveness of mobile app intervention as compared with the mobile app intervention with health navigation services and usual care.

Implications for Practice, Policy, and Future Research

Highlights from postintervention focus groups shed light on ways that the information garnered from this research can more broadly enhance cancer prevention efforts that rely on mobile technology as an intervention medium. Overall, participants in the intervention group provided feedback that the mobile app helped to increase their knowledge about breast cancer and screening methods, reminding them of the importance of receiving regular mammograms. Interestingly, their participation in the study also primed them to be more attentive to information regarding breast cancer when it incidentally arose during their consumption of other media such as television and radio. This feedback highlights the potential for mobile phone messaging to act as a conduit of other sources of cancer-related information and screening motivation. Participants' comments also suggest that simply increasing knowledge of breast cancer risk factors may induce lifestyle changes to promote cancer prevention, such as regular exercise and diet. In addition, incorporating media into direct education about the procedures for screening methods, such as a breast self-exam video and a detailed procedural video for mammography embedded in the mobile app, may promote greater engagement in self-screening methods and efficiently alleviate fears toward screening. Health care clinics may capitalize on this information by showing such videos during patient visits or sending out intermittent educational and interactive multimedia messages during the interim between visits. Finally, the culturally unique aspects of the app, such as having information available in participants' native language and having testimonials from peers who share the same ethnic background, appear to have been particularly important components, reinforcing the notion that tailoring is essential in mHealth research and outreach efforts.

On the basis of the ubiquity of mobile phones in society, including widespread use among minority communities, a multilevel and multimedia messaging intervention such as mMammogram combined with health navigator services and locally available free mammogram program (eg, Mammo-a-go-go program) holds promise to be an effective method in reaching hard-to-recruit populations with high breast cancer burdens. The use of tailored digital messages that cover broad content areas overcomes restrictions based on place and time of delivery, as well as resource and financial limits. The format and contents of mMammogram could be easily translated and disseminated to various ethnic groups who face barriers to cancer screening, with each iteration of the model programmed

to tailor its approach to the unique needs of the cultural group and the individual. Along the same lines, the mMammogram model could be modified to target promotion of multiple different preventive screening behaviors to protect against other cancers such as colorectal cancer. In addition, apps could be designed to cover the full spectrum of cancer prevention, treatment, and survivorship, encompassing diagnosis, treatment options, decision making, communication strategies, psychosocial care, and wellness. By refining and expanding technologies that target disadvantaged populations, mHealth initiatives offer an encouraging strategy to reduce disparities in breast cancer and other health conditions.

Acknowledgments

The study was supported by funding from the Susan G Komen for the Cure Foundation (IIR12223971). The research team appreciates the Foundation's generous funding to pilot this intervention study and the Korean American immigrant women who participated in this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V.1.6.1).

[PDF File (Adobe PDF File), 656KB - mhealth_v5i11e154_app1.pdf]

References

1. American Cancer Society. Cancer. How common is breast cancer? URL: https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html [accessed 2017-09-24] [WebCite Cache ID 6tm0o3jC4]
2. Kohler BA, Sherman RL, Howlader N, Jemal A, Ryerson AB, Henry KA, et al. Annual report to the nation on the status of cancer, 1975–2011, featuring incidence of breast cancer subtypes by race/ethnicity, poverty, and state. J Natl Cancer Inst 2015 Mar 30;107(6):djv048 [FREE Full text] [doi: 10.1093/jnci/djv048] [Medline: 25825511]
3. Hoeffel EM, Rastogi S, Kim MO, Shahid H. Census. 2012. The Asian population: 2010 URL: https://www.census.gov/prod/cen2010/briefs/c2010br-11.pdf [accessed 2017-10-05] [WebCite Cache ID 6tz9BGRRx]
4. Miller BA, Chu KC, Hankey BF, Ries LA. Cancer incidence mortality patterns among specific Asian Pacific Islander populations in the U.S. Cancer Causes Control 2008 Apr;19(3):227-256. [doi: 10.1007/s10552-007-9088-3]
5. Gomez SL, Quach T, Horn-Ross PL, Pham JT, Cockburn M, Chang ET, et al. Hidden breast cancer disparities in Asian women: disaggregating incidence rates by ethnicity and migrant status. Am J Public Health 2010 Apr;100(Suppl 1):S125-S131. [Medline: 20147696]
6. Lee J, Demissie K, Lu SE, Rhoads GG. Cancer Incidence among Korean-American Immigrants in the United States and Native Koreans in South Korea. Cancer Control 2017 Aug 30;14(1):78-85. [doi: 10.1177/107327480701400111]
7. Yaffe MJ. Point: mammography screening-sticking to the science. Curr Oncol 2015 Jun;22(3):174-176. [doi: 10.3747/co.22.2371] [Medline: 26089713]
8. American Cancer Society. Cancer. Breast cancer early detection and diagnosis URL: https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection.html [accessed 2017-09-20] [WebCite Cache ID 6tm8c1eRo]
9. Siu AL, U.S. Preventive Services Task Force. Screening for breast cancer: U.S. preventive services task force recommendation statement. Ann Intern Med 2016 Feb 16;164(4):279-296. [doi: 10.7326/M15-2886] [Medline: 26757170]
10. McCracken M, Olsen M, Chen Jr MS, Jemal A, Thun M, Cokkinides V, et al. Cancer incidence, mortality, and associated risk factors among Asian Americans of Chinese, Filipino, Vietnamese, Korean, and Japanese ethnicities. CA Cancer J Clin 2007;57(4):190-205. [Medline: 17626117]
11. Kagawa-Singer M, Pourat N, Breen N, Coughlin S, Abend McLean T, McNeel TS, et al. Breast and cervical cancer screening rates of subgroups of Asian American women in California. Med Care Res Rev 2007 Dec;64(6):706-730. [Medline: 17804823]
12. Kim JH, Menon U. Pre- and postintervention differences in acculturation, knowledge, beliefs, and stages of readiness for mammograms among Korean American women. Oncol Nurs Forum 2009 Mar;36(2):E80-E92. [Medline: 19273397]
13. Lee EE, Fogg LF, Sadler GR. Factors of breast cancer screening among Korean immigrants in the United States. J Immigr Minor Health 2006 Jul;8(3):223-233. [Medline: 16791532]
14. Lee HY, Ju E, Vang PD, Lundquist M. Breast and cervical cancer screening among Asian American women and Latinas: Does race/ethnicity matter? J Womens Health (Larchmt) 2010 Oct;19(10):1877-1884. [Medline: 20831437]
15. Lee H, Kim J, Han HR. Do cultural factors predict mammography behaviour among Korean immigrants in the USA? J Adv Nurs 2009 Dec;65(12):2574-2584. [Medline: 19941544]

16. Maxwell AE, Bastani R, Warda US. Mammography utilization and related attitudes among Korean-American women. Women Health 1998;27(3):89-107. [Medline: 9698640]
17. Sadler GR, Ryujin L, Nguyen T, Oh G, Paik G, Kustin B. Heterogeneity within the Asian American community. Int J Equity Health 2003 Dec 29;2(1):12. [Medline: 14697098]
18. Kandula NR, Wen M, Jacobs EA, Lauderdale DS. Low rates of colorectal, cervical, and breast cancer screening in Asian Americans compared with non‐Hispanic Whites: cultural influences or access to care? Cancer 2006 Jul 1;107(1):184-192. [Medline: 16721803]
19. CDC. Healthy People 2020 URL: http://www.cdc.gov/aging/help/DPH-Aging/healthy-people.html [accessed 2017-09-26] [WebCite Cache ID 6tmAZZ0i8]
20. Yu MY, Hong OS, Seetoo AD. Uncovering factors contributing to under-utilization of breast cancer screening by Chinese and Korean women living in the United States. Ethn Dis 2003;13(2):213-219. [Medline: 12785418]
21. Gany FM, Herrera AP, Avallone M, Changrani J. Attitudes, knowledge, and health-seeking behaviors of five immigrant minority communities in the prevention and screening of cancer: a focus group approach. Ethn Health 2006 Feb;11(1):19-39. [Medline: 16338753]
22. Juon HS, Kim M, Shankar S, Han W. Predictors of adherence to screening mammography among Korean American women. Prev Med 2004 Sep;39(3):474-481. [Medline: 15313086]
23. Lee EE, Tripp-Reimer T, Miller AM, Sadler GR, Lee SY. Korean American women's beliefs about breast and cervical cancer and associated symbolic meanings. Oncol Nurs Forum 2007 May;34(3):713-720. [Medline: 17573330]
24. Moskowitz J, Kazinets G, Wong J, Tager IB. “Health is strength”: a community health education program to improve breast and cervical cancer screening among Korean American Women in Alameda County, California. Cancer Detect Prev 2007;31(2):173-183. [Medline: 17418978]
25. Eun Y, Lee EE, Kim MJ, Fogg L. Breast cancer screening beliefs among older Korean American women. J Gerontol Nurs 2009 Sep;35(9):40-50. [Medline: 19715259]
26. Lee HY, Stange MJ, Ahluwalia JS. Breast cancer screening behaviors among Korean American immigrant women: findings from the Health Belief Model. J Transcult Nurs 2015 Nov;26(5):450-457. [Medline: 24848345]
27. Sadler GR, Ryujin LT, Ko CM, Nguyen E. Korean women: breast cancer knowledge, attitudes and behaviors. BMC Public Health 2001;1:7. [Medline: 11553321]
28. Han Y, Williams RD, Harrison RA. Breast cancer screening knowledge, attitudes, and practices among Korean American women. Oncol Nurs Forum 2000;27(10):1585-1591. [Medline: 11103377]
29. Le E, Menon U, Nandy K, Szalacha L, Kviz F, Cho Y, et al. The effect of a couples intervention to increase breast cancer screening among Korean Americans. Oncol Nurs Forum 2014 May;41(3):E185-E193. [Medline: 24769601]
30. Kim JH, Menon U, Wang E, Szalacha L. Assess the effects of culturally relevant intervention on breast cancer knowledge, beliefs, and mammography use among Korean American women. J Immigr Minor Health 2010 Aug;12(4):586-597. [Medline: 19373555]
31. Maxwell AE, Jo AM, Chin SY, Lee KS, Bastani R. Impact of a print intervention to increase annual mammography screening among Korean American women enrolled in the National Breast and Cervical Cancer Early Detection Program. Cancer Detect Prev 2008;32(3):229-235. [Medline: 18799271]
32. Han HR, Lee H, Kim MT, Kim KB. Tailored lay health worker intervention improves breast cancer screening outcomes in non-adherent Korean-American women. Health Educ Res 2009 Apr;24(2):318-329. [Medline: 18463411]
33. Kim YH, Sarna L. An intervention to increase mammography use by Korean American women. Oncol Nurs Forum 2004;31(1):105-110. [Medline: 14722594]
34. Juon HS, Choi S, Klassen A, Roter D. Impact of breast cancer screening intervention on Korean-American women in Maryland. Cancer Detect Prev 2006;30(3):297-305. [Medline: 16870356]
35. Lee SY. Cultural factors associated with breast and cervical cancer screening in Korean American women in the US: an integrative literature review. Asian Nurs Res (Korean Soc Nurs Sci) 2015 Jun;9(2):81-90. [Medline: 26160234]
36. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006 Jun 23;8(2):e10. [Medline: 16867965]
37. US Department of Health and Human Services. Hrsa. 2016. Using health text messages to improve consumer health knowledge, behaviors, and outcomes: an environmental scan URL: https://www.hrsa.gov/sites/default/files/archive/healthit/txt4tots/environmentalscan.pdf [accessed 2017-10-05] [WebCite Cache ID 6tzF2kR8x]
38. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69. [Medline: 20354039]
39. 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. [Medline: 19135907]
40. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [Medline: 19382860]
41. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236. [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]

42. 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. [doi: 10.1007/s13142-011-0021-7]
43. Fogg BJ. A behavior model for persuasive design. 2009 Presented at: The 4th International Conference on Persuasive Technology; April 26-29; Claremont, California. [doi: 10.1145/1541948.1541999]
44. Fogg BJ. Persuasive Technologies. Communications of the ACM 1999;42(5):26-29. [Medline: 10.1145/301353.301396]
45. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco, CA: Morgan Kaufmann Publishers; 2002:89-120.
46. Champion VL, Skinner CS. The health belief model. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass; 2008:45-65.
47. Doig GS, Simpson F. Randomization and allocation concealment: a practical guide for researchers. J Crit Care 2005 Jun;20(2):187-191. [Medline: 16139163]
48. Caplan LS, Mcqueen DV, Qualters JR, Leff M, Garrett C, Calonge N. Validity of women's self-reports of cancer screening test utilization in a managed care population. Cancer Epidemiol Biomarkers Prev 2003 Nov;12(11):1182-1187. [Medline: 14652278]
49. McCance KL, Mooney KH, Smith KR, Field R. Validity and reliability of a breast cancer knowledge test. Am J Prev Med 1990;6(2):93-98. [Medline: 2363955]
50. Hall CP, Wimberley PD, Hall JD, Pfriemer JT, Hubbard E, Stacy AS, et al. Teaching breast cancer screening to African American women in the Arkansas Mississippi river delta. Oncol Nurs Forum 2005 Jul 1;32(4):857-863. [Medline: 15990915]
51. Madanat H, Merrill RM. Breast cancer risk-factor and screening awareness among women nurses and teachers in Amman, Jordan. Cancer Nurs 2002 Aug;25(4):276-282. [Medline: 12181496]
52. Padela A, Murrar AS, Adviento B, Liao C, Hosseinian Z, Peek M, et al. Associations between religion-related factors and breast cancer screening among American Muslims. J Immigr Minor Health 2015 Jun;17(3):660-669. [doi: 10.1007/s10903-014-0014-y] [Medline: 24700026]
53. Champion VL. Instrument refinement for breast cancer screening behaviors. Nurs Res 1993;42(3):139-143. [Medline: 8506161]
54. Champion VL. Revised susceptibility, benefits, and barriers scale for mammography screening. Res Nurs Health 1999 Aug;22(4):341-348. [Medline: 10435551]
55. Champion V, Maraj M, Hui S, Perkins AJ, Tierney W, Menon U, et al. Comparison of tailored interventions to increase mammography screening in nonadherent older women. Prev Med 2003 Feb;36(2):150-158. [Medline: 12590989]
56. Karayurt O, Dramali A. Adaptation of Champion's Health Belief Model Scale for Turkish women and evaluation of the selected variables associated with breast self-examination. Cancer Nurs 2007;30(1):69-77. [Medline: 17235224]
57. Parsa P, Kandiah M, Mohd Nasir MT, Hejar AR, Nor Afiah MZ. Reliability and validity of Champion's Health Belief Model Scale for breast cancer screening among Malaysian women. Singapore Med J 2008 Nov;49(11):897-903. [Medline: 19037556]
58. Taymoori P, Berry T. The validity and reliability of Champion's Health Belief Model Scale for breast cancer screening behaviors among Iranian women. Cancer Nurs 2009;32(6):465-472. [Medline: 19816165]
59. Tang TS, Solomon LJ, McCracken LM. Cultural barriers to mammography, clinical breast exam, and breast self-exam among Chinese-American women 60 and older. Prev Med 2000 Nov;31(5):575-583. [doi: 10.1006/pmed.2000.0753] [Medline: 11071839]
60. Taylor VM, Jackson JC, Tu SP, Yasui Y, Schwartz SM, Kuniyuki A, et al. Cervical cancer screening among Chinese Americans. Cancer Detect Prev 2002;26(2):139-145. [Medline: 12102148]
61. Prochaska JO, Redding 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; 1997:61-83.
62. van Drongelen A, Boot CR, Hlobil H, Twisk JW, Smid T, van der Beek AJ. Evaluation of an mHealth intervention aiming to improve health-related behavior and sleep and reduce fatigue among airline pilots. Scand J Work Environ Health 2014 Nov;40(6):557-568. [Medline: 25121620]
63. Arora S, Peters AL, Agy C, Menchine M. A mobile health intervention for inner city patients with poorly controlled diabetes: proof-of-concept of the TExT-MED program. Diabetes Technol Ther 2012 Jun;14(6):492-496. [Medline: 22524591]
64. 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 Sep 4;2(2):e32. [Medline: 24004517]
65. 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). [Medline: 26553211]
66. 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;10(1):e1001362. [Medline: 23349621]

67. Dillard JP, Weber KM, Vail RG. The relationship between the perceived and actual effectiveness of persuasive messages: a meta-analysis with implications for formative campaign research. J Commun 2007;57(4):613-631. [doi: 10.1111/j.1460-2466.2007.00360.x]
68. Dillard JP, Shen L, Vail RG. Does perceived message effectiveness cause persuasion or vice versa? 17 consistent answers. Hum Commun Res 2007;33(4):467-488. [doi: 10.1111/j.1468-2958.2007.00308.x]
69. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. Br Med J Open 2013;3(1):e001570. [doi: 10.1136/bmjopen-2012-001570]
70. Harris DM, Miller JE, Davis DM. Ncbi.nlm.nih. 2003. Racial differences in breast cancer screening, knowledge and compliance URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2594576/pdf/jnma00312-0043.pdf [accessed 2017-10-05] [WebCite Cache ID 6tzIav6Rc]
71. Pearlman DN, Clark MA, Rakowski W, Ehrich B. Screening for breast and cervical cancers: the importance of knowledge and perceived cancer survivability. Women Health 1999;28(4):93-112. [Medline: 10378347]
72. Sim HL, Seah M, Tan SM. Breast cancer knowledge and screening practices: a survey of 1,000 Asian women. Singapore Med J 2009 Feb;50(2):132-138. [Medline: 19296027]
73. Lee HY, Koopmeiners JS, Rhee TG, Raveis VH, Ahluwalia JS. Mobile phone text messaging intervention for cervical cancer screening: changes in knowledge and behavior pre-post intervention. J Med Internet Res 2014 Aug 27;16(8):e196. [Medline: 25164545]
74. Lee HY, Koopmeiners JS, McHugh J, Raveis VH, Ahluwalia JS. mHealth pilot study: text messaging intervention to promote HPV vaccination. Am J Health Behav 2016 Jan;40(1):67-76. [Medline: 26685815]
75. Pandey A, Hasan S, Dubey D, Sarangi S. Smartphone apps as a source of cancer information: changing trends in health information-seeking behavior. J Cancer Educ 2013 Mar;28(1):138-142. [Medline: 23275239]
76. Bernardi RA, Guptill ST. Social desirability response bias, gender, and factors influencing organizational commitment: an international study. J Bus Ethics 2008;81(4):797-809. [doi: 10.1007/s10551-007-9548-4]
77. Chung J, Monroe GS. Exploring social desirability bias. J Bus Ethics 2003;44(4):291-302. [doi: 10.1023/A:1023648703356]
78. Dalton D, Ortegren M. Gender differences in ethics research: the importance of controlling for the social desirability response bias. J Bus Ethics 2011;103(1):73-93. [doi: 10.1007/s10551-011-0843-8]
79. Keillor BD, Owens D, Pettijohn C. Mrs. A cross-cultural/cross-national study of influencing factors and socially desirable response biases URL: https://www.mrs.org.uk/ijmr_article/article/34842 [accessed 2017-10-05] [WebCite Cache ID 6tzJrn3mD]
80. Middleton KL, Jones JL. Socially desirable response sets: the impact of country culture. Psychol Mark 2000;17(2):149-163. [doi: 10.1002/(SICI)1520-6793(200002)17:2<149::AID-MAR6>3.0.CO;2-L]
81. Hou SI, Sealy DA, Kabiru CW. Closing the disparity gap: cancer screening interventions among Asians--a systematic literature review. Asian Pac J Cancer Prev 2011;12(11):3133-3139. [Medline: 22394003]

Abbreviations

ACS: American Cancer Society
ANOVA: analysis of variance
BSE: breast self-examination
CAB: community advisory board
CBE: clinical breast examination
FBM: Fogg behavioral model
GPS: global positioning system
HBM: health belief model
mHealth: mobile health
RCT: randomized controlled trial
SD: standard deviation

Edited by G Eysenbach; submitted 01.12.16; peer-reviewed by N Azevedo, G Clarke Hillyer; comments to author 10.02.17; revised version received 08.07.17; accepted 29.07.17; published 07.11.17.

Please cite as:
Lee H, Ghebre R, Le C, Jang YJ, Sharratt M, Yee D
Mobile Phone Multilevel and Multimedia Messaging Intervention for Breast Cancer Screening: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2017;5(11):e154
URL: http://mhealth.jmir.org/2017/11/e154/
doi:10.2196/mhealth.7091
PMID:29113961

Original Paper

User Acceptance of Wrist-Worn Activity Trackers Among Community-Dwelling Older Adults: Mixed Method Study

Arjun Puri[1], MSc; Ben Kim[1], MHI; Olivier Nguyen[1,2], BEng; Paul Stolee[3], PhD; James Tung[4], PhD; Joon Lee[1], PhD

[1]Health Data Science Lab, School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada

[2]Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada

[3]School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada

[4]Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Joon Lee, PhD
Health Data Science Lab
School of Public Health and Health Systems
University of Waterloo
Lyle Hallman North
200 University Avenue West
Waterloo, ON, N2L3G1
Canada
Phone: 1 5198884567 ext 31567
Fax: 1 5197466776
Email: joon.lee@uwaterloo.ca

Abstract

Background: Wearable activity trackers are newly emerging technologies with the anticipation for successfully supporting aging-in-place. Consumer-grade wearable activity trackers are increasingly ubiquitous in the market, but the attitudes toward, as well as acceptance and voluntary use of, these trackers in older population are poorly understood.

Objective: The aim of this study was to assess acceptance and usage of wearable activity trackers in Canadian community-dwelling older adults, using the potentially influential factors as identified in literature and technology acceptance model.

Methods: A mixed methods design was used. A total of 20 older adults aged 55 years and older were recruited from Southwestern Ontario. Participants used 2 different wearable activity trackers (Xiaomi Mi Band and Microsoft Band) separately for each segment in the crossover design study for 21 days (ie, 42 days total). A questionnaire was developed to capture acceptance and experience at the end of each segment, representing 2 different devices. Semistructured interviews were conducted with 4 participants, and a content analysis was performed.

Results: Participants ranged in age from 55 years to 84 years (mean age: 64 years). The Mi Band gained higher levels of acceptance (16/20, 80%) compared with the Microsoft Band (10/20, 50%). The equipment characteristics dimension scored significantly higher for the Mi Band (P<.05). The amount a participant was willing to pay for the device was highly associated with technology acceptance (P<.05). Multivariate logistic regression with 3 covariates resulted in an area under the curve of 0.79. Content analysis resulted in the formation of the following main themes: (1) smartphones as facilitators of wearable activity trackers; (2) privacy is less of a concern for wearable activity trackers, (3) value proposition: self-awareness and motivation; (4) subjective norm, social support, and sense of independence; and (5) equipment characteristics matter: display, battery, comfort, and aesthetics.

Conclusions: Older adults were mostly accepting of wearable activity trackers, and they had a clear understanding of its value for their lives. Wearable activity trackers were uniquely considered more personal than other types of technologies, thereby the equipment characteristics including comfort, aesthetics, and price had a significant impact on the acceptance. Results indicated that privacy was less of concern for older adults, but it may have stemmed from a lack of understanding of the privacy risks and implications. These findings add to emerging research that investigates acceptance and factors that may influence acceptance of wearable activity trackers among older adults.

(JMIR Mhealth Uhealth 2017;5(11):e173) doi:10.2196/mhealth.8211

KEYWORDS

health; mHealth; fitness trackers; older adults

Introduction

Smart Wearable Devices and Older Adults

Today, Canadian older adults are leading longer, healthier, and more active lives compared with older adults from previous decades [1-3]. Successful aging is achieved through aging-in-place, a concept that depicts the continued living at home and to do so while maintaining independence, social contact, and dignity [4-6]. Aging-in-place has profound health and mental benefits [7-9]; is more cost-effective than institutionalized care [7,8,10]; and is perceived as more desirable, graceful, and fulfilling among the aging cohort [4,11].

With the increased desirability to age in place, numerous technologies have emerged with the aim of supporting aging-in-place with diverse purposes, including enhancing safety through providing medication reminders [12,13], improving social interactivity through video telephony [12,14,15], and maintaining capacity to carry out daily activities functions via electronic memory aids [16,17]. In recent years, off-the-shelf smart wearable devices such as heart rate monitors and physical activity trackers have seen tremendous growth [18]. Consumer research indicates that baby boomers are the next primary users of smart wearable technology as they are the fastest growing wearable activity tracker users [18,19]. More importantly, patients and providers equally anticipate a greater role of wearable activity trackers in managing health and achieving high quality of care and patient satisfaction [20]. The rapid growth of wearable activity trackers may have originated from the opportunity it provides for aging-in-place as a tool that can enable self-management of chronic diseases, remote monitoring by clinicians, and collecting clinically relevant data that can fuel big data analytics [21,22]. To seize this opportunity, it is critical to understand the factors and processes involved in adopting and using smart wearable activity trackers among older adults. Understanding the acceptance and use of smart wearable activity trackers are especially important for enabling aging-in-place, as ongoing and voluntary use is critical to accurate and comprehensive data collection for technology-driven interventions [23]. However, there currently exists little research that appropriately and adequately explores older adults' attitudes toward, as well as acceptance and usage of, smart wearable activity trackers [24,25].

Literature Review on Wearable Acceptance

Research activities for smart wrist activity trackers tend to focus on younger populations [26,27]. An increase in gerontological research has been observed because of the increased adoption of smart wearable activity trackers among older adults and the recognition of their greater potential health benefits to older adults [28-30]. Previous research studies identified that older adults in general perceived smart wearable activity trackers as easy to use, useful, comfortable, and acceptable in the short term [28] and long term [29]. These 2 studies used wearable activity trackers that clip onto the belt or pocket, which can affect its acceptance differently than wearable activity trackers that are worn on wrists like a bracelet. One study examined the acceptance of both wearable activity trackers that are worn on wrists and clip-ons and traditional pedometers by older adults [30]. This study reported higher acceptance of wearable activity trackers, either clip-on or wrist worn, over pedometers among older adults [30]. Furthermore, older adults expressed that wrist-worn activity trackers are preferred over clip-ons because of a less likelihood of losing or breaking them [30].

The first impression is often determined by the style, but a long-term adoption of wearable activity trackers is often influenced by an array of factors such as comfort and usability [31]. The short study duration of the previous smart wearable device research study left the investigation of the technology acceptance deficient and limited to the first impression [30]. The duration of the study ranged from 3 days to 7 days for participants [30]. Short-term technology acceptance may not be indicative of long-term acceptance as research indicates that use of smart wearables such as activity trackers tends to drop after the first few weeks of ownership [18]. This rapid drop of adherence to wearable activity trackers is also apparent in a younger population where more than 50% stopped using the device after 14 days and 75% at around 30 days [32]. Therefore, there is a need for a study with a longer duration than 7 days.

Existing research studies [28-30] examined the acceptance level solely based on the technology acceptance model (TAM) [33]. The theoretical constructs from TAM, including perceived usefulness, perceived ease of use, and other external variables such as comfort, were highly associated with older adults' acceptance [28-30]. However, investigation of technology acceptance with TAM leaves out other critical factors that are important. TAM received an update (ie, TAM2) to put greater emphasis on technology acceptance within organizational settings consisting of additional theoretical constructs that describe the perceived usefulness that incorporates social influences, output quality, and result demonstrability [34]. Further iteration of update resulted in TAM3, which refined and added more theoretical constructs related to the perceived ease of use [35]. Another popular theoretical model for technology acceptance, the unified theory of acceptance and use of technology (UTAUT), which combined multiple technology acceptance and behavior change theories, also emphasizes the theoretical constructs related to social influences as well as individual characteristics such as age and gender as a moderator to behavior intention [36]. These newer theoretical constructs are missing in the TAM that the existing research studies used, yet the significant increase in these models' performance in explaining technology acceptance is attributed to these newly added constructs [34-36].

These theoretical models aim to provide general determinants of technology acceptance. Subsequently, they lack context specificity and artifact specificity, but they can provide more relevant key predictors for technology acceptance [35]. There has been development of technology acceptance theoretical models that are specific to older adults such as social agent

technologies [37] and Internet uses [38] or for emerging technologies such as wearable devices [39,40]. Gao et al [40] identified privacy as an important barrier for older adults to accept wearable technology. Equipment characteristics such as battery longevity, ergonomics, and aesthetics have also been identified as important factors for older adults [30,41]. Cost was found to be a significant factor in a systematic review that examined technologies for aging-in-place [17]. A similar pattern was noted in a previous wearable activity tracker study in which participants considered cost as one of the major barriers for future purchase [30]. This raises the question of sensitivity toward the cost of equipment as the tested devices' prices ranged from US $60 to US $150, which is often considered low cost in the eyes of researchers [28]. To the best of our knowledge, no studies have examined the factors for acceptance of wrist-worn activity trackers among older adults using potential determinants outside TAM.

Research Objective

The objective of this study was to assess acceptance and usage of wearable activity trackers in Canadian community-dwelling older adults in a free-living environment. This study extends the current literature by investigating additional potentially influential factors to TAM, unique to emerging technologies and older adults, including privacy concerns, facilitating conditions, perceived risks, subjective norm, and equipment characteristics.

Methods

Study Design

A sequential explanatory mixed method design was used to explore the study objective. The quantitative data were collected in the first phase of the study through a questionnaire that was developed specifically for this study. The questionnaire results were analyzed and a semistructured interview guideline was developed. In the second phase, semistructured interviews were conducted to further probe older adults' experience with wrist-worn activity trackers and complement the quantitative analyses.

Research ethics approval for this study was obtained from the University of Waterloo Office of Research Ethics. All participants gave written informed consent.

Procedures

Phase 1

This was a 6-week-long phase focused on quantitatively assessing older adults' technology acceptance through the questionnaire. This phase applied a crossover study design in which participants were randomly assigned to 2 groups, one for the Microsoft Band and the other for the Xiaomi Mi Band. The crossover between the 2 groups happened at the end of the third week. In other words, each participant tried both wrist bands, for 3 weeks each, but the order of the devices was randomized. The questionnaire was administered twice to capture technology acceptance toward each device, at the end of the third and sixth week. Demographical information and information regarding previous experience with technology and wearable devices were also collected.

Phase 2

Semistructured interviews were conducted at the participant's own residence in a private setting. A reflexive interview process, which allows for modification of the interview technique and content as needed, was adopted to overcome the researcher's own postulations or presumptions of wearable activity trackers [42,43]. The first interview was reviewed to note early instances of potential biases in the form of leading questions, which was readjusted and corrected in the remaining 3 interviews.

Recruitment

A convenience sampling technique was used to recruit 20 older adults aged 55 years and older from the cities of Kitchener, Waterloo, Cambridge, and Guelph in Ontario, Canada. Flyers were posted in at local community centers and recreational facilities with approval. Interested participants were instructed to contact the researcher via phone or email. Recruitment started in March 2016 and lasted for 2 months. Participants must have been able to wear 2 wearable activity trackers for 21 days each.

A purposive and criterion sampling was used to recruit 4 participants for the subsequent semistructured interviews from the 20 participants. The criteria for purposive sampling were determined by age and the degree of wearable activity tracker acceptance (ie, high acceptance, medium acceptance, and low acceptance).

Equipment

Wearable Activity Trackers

Two consumer-level wearable activity trackers, the Microsoft Band and Mi Band (Figure 1), were selected based on their features (accelerometer, gyroscope, display), functionalities (step count, heart rate monitoring, recommendations), ergonomics (size and flexibility), and price. These 2 wearable activity trackers have immense differences in features, sensors, and price. As a result, these 2 devices represent the far ends of the consumer market spectrum, and the selection of devices was based on these differences.

The Microsoft Band offers an extensive sensor array and a touch screen liquid-crystal display (LCD; Table 1). The manufacturer's quoted battery life is approximately 48 hours but varies based on individual usage. It is considerably more expensive than basic activity and heart rate trackers available at a cost of US $299 at launch.

The Mi Band offers a basic triaxial accelerometer and an optical heart rate sensor at a cost of US $20 at launch (Table 1). The Mi Band has an estimated battery life of almost 30 days, with real-world tests performance ranging from 45 days to 50 days. The Mi Band has 3 individual light-emitting diodes (LEDs) that provide activity progress by lighting up 1, 2, or all 3. The device is lightweight and can be worn on either the wrist or neck as a pendant, offering versatile placement. The placement of the Mi Band was limited to the wrist for this study.

Figure 1. The Xiaomi Mi Band (left) and Microsoft Band (right).

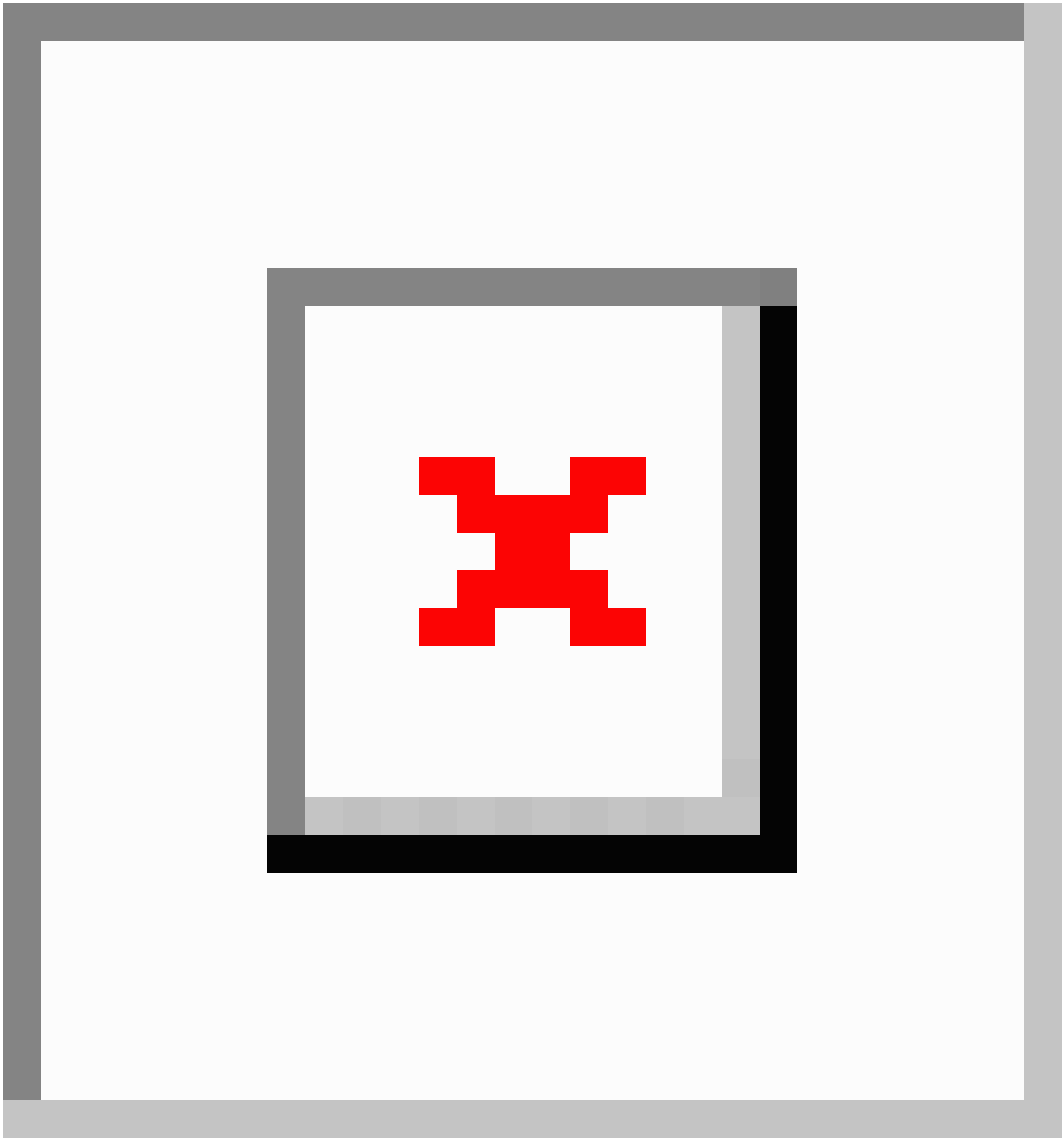


Table 1. Two wearable activity trackers' characteristics.

Characteristic	Mi Band	Microsoft Band
Display	LED[a] dots	Touch LCD[b] display
Battery life	30 days	48 hours
Cost	US $20	US $299
Sensors	Accelerometer, optical heart rate sensor	Accelerometer, optical heart rate sensor, gyroscope, galvanic skin response sensor, global positioning system, UV[c] sensor, microphone

[a]LED: light-emitting diode.

[b]LCD: liquid-crystal display.

[c]UV: ultraviolet.

Table 2. Description of dimensions related to technology acceptance.

Dimensions	Description
Perceived usefulness	Perceived usefulness refers to improvements in one's job performance but in the context of this study; it was adapted to refer to the degree to which using a technology can help monitor older adults' health and support aging-in-place.
Perceived ease of use	Perceived ease of use has been established as a key indicator for user acceptance and is defined as "the degree to which a person believes that using a technology will be free from effort" [45].
Subjective norm	Subjective norm is another dimension that is representative of user acceptance, and it is defined as the likelihood of recommending the use of the said technology to individuals who are influential in the lives of the technology user [45] .
Facilitating conditions	A review of literature revealed varying classifications of facilitating conditions [38]. Facilitating factors are factors that can increase or decrease the effort required to use a technology such as availability, affordability, availability of training resources, and so on [46].
Privacy concerns	Privacy concerns is a novel dimension in the framework and has been included because of the emergent tendency of smart device and technology manufacturers to use Internet communication protocols to store and analyze data in the cloud, rather than on the particular device.
Perceived risks	Perceived risks have been established to be influential to consumer behavior and important when evaluating user acceptance of technology [47].
Equipment characteristics	Finally, equipment characteristics that can influence the technology acceptance were deemed an important dimension and described as one of the major factors in another study [41].

Smartphone

Both wearable activity trackers offer companion smartphone apps, and all participants were provided with an accompanying smartphone, the Motorola Moto E, with the app preinstalled and set up for use. The collected data from wearable activity trackers were transferred to the smartphones. The apps displayed the progress, patterns, and summaries. Participants were trained on how to use the wearable activity trackers and smartphone apps, but smartphone use was not mandatory.

Data Collection and Analysis

Technology Acceptance Questionnaire

A 31-item, 5-point Likert scale, and an additional 6 multiple-choice items, self-reported, paper-based questionnaire for older adults was developed based on the fundamental dimensions that influence user acceptance of technology from TAM [33] and the sensor acceptance model [44] (Multimedia Appendix 1). The investigated key dimensions for wearable activity tracker acceptance were perceived usefulness, perceived ease of use, privacy concerns, perceived risks, facilitating conditions, subjective norm, and equipment characteristics (Table 2). Detailed description of the technology acceptance questionnaire, including each dimension and corresponding questions, is summarized in Multimedia Appendix 1.

Semistructured Interview

A crossover approach to data collection through the questionnaire captured the technology acceptance and finer granular information about various factors that may influence the acceptance. However, this design of study does not fully explain the contextual factors related to technology acceptance and the participants' cognitive rationale behind the questionnaire results. To overcome this shortcoming, one-to-one semistructured interviews were conducted in the second phase [48]. A semistructured interview guide was developed to gain deeper insight into each dimension of technology acceptance.

Analyses

Phase 1

The demographic information, previous technology use, and characteristics related to smart wearable devices were analyzed through descriptive statistics. For all quantitative analyses, cases with missing data were excluded, which occurred because some participants refused to answer some of the questions.

The acceptance of the wearable activity tracker was measured by the question L33: *Would you use the device you used during the last 21 days to continue to monitor or track your physical activity or health?* In a univariate analysis, several statistical tests were used to analyze participants' responses to each Likert-scale question, with respect to user acceptance. First, the Student independent *t* test was performed to investigate the differences in the total mean score for the technology acceptance questionnaire between the Microsoft and Mi Bands. Second, the Wilcoxon signed-rank test was conducted to test for the differences in the dimensions in the technology acceptance questionnaire between the 2 devices. Third, Spearman rho was calculated to test whether a correlation between participants' responses to the Likert-scale questions and user acceptance exists. Finally, Spearman rho was used to assess the relationship between the order in which the devices were provided and the acceptance so as to ensure that the order did not influence the acceptance.

In a multivariable regression analysis, the association between the 7 dimensions from the technology acceptance questionnaire (Table 2) and user acceptance was assessed using logistic regression. Each dimension was represented by the sum of the responses to the questions that belong to that dimension. Because there were fewer than 40 questionnaire responses after excluding missing data (each of the 20 participants completed 2 questionnaires for the 2 wrist bands), only 3 of the 7 dimensions were selected using backward stepwise feature selection and were included as covariates in the logistic

regression model, so that the number of covariates did not exceed one-tenth of the number of cases. Using the 3 selected dimensions, leave-one-out cross-validation was performed to evaluate the logistic regression model's user acceptance classification performance. As the performance metric, the area under the receiver operating characteristic curve (AUC) was calculated. Statistical significance was set at alpha=.05 for all statistical results.

Phase 2

Semistructured interviews were coded and themed using a directed content analysis strategy whereby the themes explored follow structure determined by concepts reviewed in literature, while also allowing for the discovery of the previously undiscovered or unmentioned data and themes [43]. All interviews were transcribed, and 2 researchers independently read the transcripts to familiarize themselves with the data. Predetermined codes based on the concepts and variables discovered during literature review were used. New data that were not represented by preexisting categories were then identified and analyzed. Codes were reviewed and combined into themes that appropriately and accurately described the interview data. This process was iterative as the researchers' knowledge of the interview data increased. The themes generated by the 2 researchers were compared with to note similarities and differences and transformed into the overarching themes.

Results

Participant Characteristics

All 20 participants who enrolled completed the study. The respondents ranged in age from 55 years to 84 years (mean 64 years), and 60% (12/20) were female. Of the total, 18 participants (90%) used a computer on a daily basis and 14 (70%) personally owned a smartphone. Seventeen participants (85%) had heard of smart wearable devices, indicating a high degree of awareness among the group, though only one used a store-bought wearable activity tracker to monitor their health (Table 3).

Table 3. Participant characteristics and previous technology experience.

Characteristic	Value
Age in years, mean (range)	64 (55-84)
Sex, n (%)	
Male	8 (40%)
Female	12 (60%)
Marital status, n (%)	
Married	15 (75%)
Divorced	2 (10%)
Separated	1 (5%)
Widowed	2 (10%)
Never married	0 (0%)
Education level, n (%)	
High school	3 (15%)
Some postsecondary	3 (15%)
Completed postsecondary	6 (30%)
Some postgraduate	1 (5%)
Completed postgraduate	7 (35%)
Income in CAD dollars [a], n (%)	
Less than $20,000	2 (10%)
$20,000-$39,999	5 (25%)
$40,000-$69,999	3 (15%)
$70,000-$99,999	7 (35%)
$100,000-$149,999	2 (10%)
$150,000 or more	0 (0%)
Computer use, n (%)	
None	1 (5%)
Once a month	0 (0%)
Once or twice a week	1 (5%)
Daily	18 (90%)
Smartphone ownership, n (%)	
Own smartphone	14 (70%)
Do not own smartphone	6 (30%)
Smart wearable device ownership, n (%)	
Own smart wearable device	1 (5%)
Do not own smart wearable device	19 (95%)
Heard of smart wearable devices, n (%)	
Yes	17 (85%)
No	3 (15%)

[a]Missing n=1.

Table 4. Acceptance per device.

Acceptance	Mi Band, n (%)	Microsoft Band, n (%)	Combined, n (%)
Yes	16 (80)	10 (50)	26 (65)
No	4 (20)	10 (50)	14 (35)

Table 5. Differences in the technology acceptance questionnaire scores per dimension for each device.

Dimensions (maximum score)	Mi Band score, mean (SD)	Microsoft Band score, mean (SD)	*P* value
Equipment characteristics (30)	29.05 (3.27)	22.35 (4.20)	<.001
Perceived ease of use (35)	25.45 (4.70)	27.55 (3.24)	.11
Facilitating conditions (10)	12.1 (1.59)	11.55 (1.50)	.12
Perceived usefulness (25)	18.3 (2.72)	17.25 (2.69)	.13
Privacy concerns (15)	11.1 (2.71)	11.55 (2.67)	.18
Perceived risks (15)	4.75 (1.29)	4.9 (1.55)	.72
Subjective norm (15)	11.10 (2.00)	11.35 (2.21)	.94

Wearable Activity Tracker Acceptance

Overall, the wearable activity trackers had a moderate level of acceptance (26/40, 65%) among the community-dwelling older adults. The Mi Band had higher acceptance rate (16/20, 80%) than the Microsoft Band (10/20, 50%) (Table 4). The order in which participants received the devices was not correlated with the acceptance for the Mi Band and Microsoft Band (*P* value of .67 and .73, respectively).

Technology Acceptance Questionnaire

Overall, participants' experiences were similar between the Mi and Microsoft Bands when the individual dimensions were compared (Table 5). The participants' total score to the technology acceptance questionnaire were higher for the Mi Band (72.16) than Microsoft Band (68.71), but the Student *t* test was statistically not significant (*P*=.16). The Wilcoxon signed-rank test revealed that the Mi Band scored significantly higher in the equipment characteristics dimension (*P*<.001; Table 5). Statistical significance was not found for the perceived ease of use dimension between the devices (*P*=.11), despite the Microsoft Band scoring significantly higher in 3 questions that measured perceived ease of use (*P*<.05; Multimedia Appendix 2). No other dimensions were significantly different between the devices.

Dimensions Associated With Technology Acceptance

According to Spearman rho, technology acceptance was moderately correlated with 8 questions with correlation coefficients ranging from .31 to .49 (Table 6). Technology acceptance was most highly correlated with question L35, which asked participants about the price they were willing to pay for wearable activity trackers. Privacy concerns (L19 and L20) were negatively correlated in moderate strengths to technology acceptance.

Table 6. Correlation between technology acceptance questionnaire and technology acceptance.

Question item #	Corresponding dimension	Correlation coefficient (rho)	*P* value
L35: How much would you be willing to pay for the device you wore during the last 21 days?	N/A[a]	.49	.001
L19: I had no concerns about my privacy while wearing the device.	Privacy concerns	−.43	.006
L21: I have the knowledge necessary to use the device.	Facilitating conditions	.43	.006
L20: I am comfortable with my health data being shared with equipment manufacturers as long as it is shared anonymously.	Privacy concerns	−.38	.02
L13: I find the device easy to use.	Perceived ease of use	.35	.03
L12: The device's smartphone application was easy to use.	Equipment characteristics	.34	.03
L2: I was afraid that the device would discover a major health issue.	Perceived risks	.32	.04
L18: I was able to put the device on in a reasonable amount of time.	Perceived ease of use	−.31	.05

[a]N/A: not applicable

Table 7. Multivariate logistic regression model with 3 selected features.

Dimensions	Odds ratio (95% CI)	*P* value
Facilitating conditions	2.51 (1.20-5.27)	.02
Privacy concerns	0.64 (0.36-1.13)	.12
Perceived risks	1.82 (0.85-3.90)	.13

Table 8. Semistructured participant characteristics.

Participant name	Age in years	Gender	Band acceptance		Current smart device ownership
			Mi Band	Microsoft Band	
Anita	65	Female	No	No	Blackberry
Paula	84	Female	Yes	No	None
Francine	65	Female	Yes	No	iPad
Greg	83	Male	No	No	Yes[a]

[a]Exact device model was not identified.

Predictors of Wearable Activity Tracker Acceptance

Feature selection resulted in the 3 dimensions (facilitating conditions, privacy concerns, and perceived risks) to be used as covariates in the multivariate logistic regression model. The logistic regression results are shown in Table 7. AUC for this model was 0.79.

Directed Content Analysis

Participant Characteristics

The 4 participants recruited for the semistructured interviews and their acceptance of each wearable activity tracker are described in Table 8. Unfortunately, we could not recruit a participant who accepted both wearable activity trackers to satisfy the criterion sampling. Pseudonyms were assigned to protect confidentiality. The directed content analysis resulted in 5 overarching themes.

Theme 1: Smartphones as Facilitators of Wearable Activity Trackers

The acceptance of a wearable activity tracker could be affected as the wearable activity trackers relied heavily on smartphones for visualizing the data. When participants were asked about their experience and attitudes related to smartphones, they reverberated perceived ease of use as an important factor for future intention to use. Participants voiced their concerns with regard to perceived ease of use through fear of forgetting to use, losing, and breaking smartphones, as well as the inconvenience of carrying them, which influenced their future intention to use. Participants stated:

> *To tell you the truth, I was afraid to use it just in case I broke it because I didn't know anything about it.* [Paula]

> *I'm not one to take a phone with me.* [Francine]

Lack of experience with smartphones among participants led to two different outcomes in older adults' desire to use smartphones in the future. First, Anita and Francine, who had been exposed to smartphone previous to the study, responded positively toward their experience of using smartphones. Anita noted that she liked "the convenience" of a smartphone and "the fact that you just swipe it...I like that." Francine appreciated the immediacy and convenience with which smartphones can provide information:

> *I like the idea that the information is right there when you want it...I like the idea that it's a source of information that is easily accessible…* [Francine]

They expressed their future intention to use smartphones and willingness to learn as follows:

> *Well, I have tried to use it more. So I guess it helped to—made it decide that maybe we need to—I need to work more on it and try and figure out what exactly I can do with it.* [Anita]

> *I would think that they are the way of the future.* [Francine]

On the other hand, Greg and Paula described their prior experience with smartphone as negative and casted their pessimistic attitudes toward future intentions as follows:

> *We are so far behind in the e-world that trying to cope with those type of things, out of our ignorance, is really sort of impossible at times.* [Greg]

> *I don't know if I'm ready to have a smartphone and get rid of my house phone.* [Paula]

Extending the negative perceptions toward smartphones, they speculated the harmful effect of smartphones by describing them as a "distraction," "deterrent to conversation," and "enslaving of our young people."

Theme 2: Privacy Is Less of a Concern for Wearable Activity Trackers

Participants perceived the wearable activity tracker data such as step counts, sleep hours and efficiency, physical activity level, and heart rate as not private and were open to sharing them. Additionally, there was an overall diminished sense of privacy due to widespread data sharing in other aspects of their lives, as shown below:

> *I mean privacy—I would not like somebody to be able to go into my bank account or into personal details like that. But privacy; how I live or what I do, that's not a—not bothering me. No.* [Paula]

> *It's like the information you have about your salary, how much you pay for your house, how much you pay for rent, how much you paid for your car, whether you had sex last night or anticipate having it next year [laughs] you know, all that stuff…I can share anything in my life and I don't get the feeling that I shouldn't be sharing that with somebody.* [Greg]

Francine displayed a high level of trust in the system and compared it with her trust in the security of electronic medical records as follows:

> *So, if it's connected with my name, great. I mean there's all that information in my doctor's computer, which is linked to the hospitals and various other places I'm sure. Can't see why it can't be there.* [Francine]

Participants' perception of wearable activity tracker data differs from traditional notions of private data. Participants had minimal privacy concerns that were specific to wearable activity trackers and believed that they would have no bearing on acceptance and future use.

Theme 3: Value Proposition—Self-Awareness and Motivation

Participants perceived the information from wearable activity trackers useful for understanding the level and intensity of physical activity they are getting, as echoed by the quotes below:

> *I'd like to know how inactive I am. I'd also like to know how much sleep I really do get.* [Francine]

> *But it has helped in that I know when I do certain walks, approximately how many steps and when I do my classes, how many steps. Then I can get a rough idea but it's the competitiveness in me I guess to.* [Anita]

The increased self-awareness motivated lifestyle changes. Anita attested to her increased motivation to increase physical activity level, whereas Paula used the wearable activity tracker to recognize when to slow down:

> *To make sure that I'm doing my 8000 steps a day or whether the number so I'm getting enough exercise…But I would still I think be more likely to keep up my step count if I was wearing it.* [Anita]

> *Moving and that kind of stuff, so that's made me realized yeah you should stop and you should settle down and take it easy.* [Paula]

Wearable activity trackers triggered them to educate themselves on health topics. Greg stated:

> *The only thing I was surprised about, I know about REM sleep, I've done a lot of reading on it. I learned from this that out of 5½ or 6 hours sleep, or what I thought was sleep, an hour and 20 minutes was in REM.* [Greg]

Not all information was valuable, available, or presented in the right format. Information lost its value when participants already understood their current health status. Furthermore, wearable activity trackers did not present resting heart rate in an easy-to-understand manner, which diminished its value. Greg said:

> *…interesting, but not necessarily useful…But I wasn't worried about whether I was doing the right thing or not…I never did learn anything about my resting heart rate when I was sleeping.* [Greg]

Theme 4: Subjective Norm, Social Support, and Sense of Independence

Anita and Francine expected that their friends and family would support their decision to use wearable activity trackers and did not foresee reasons for discouragement from them. Francine had revealed that she received support from her husband with the use of wearable activity trackers, demonstrating the availability of immediate social support. However, Greg and Paula indicated possibly disheartening opinions from friends and family, followed by positive outlooks:

> *...my youngest daughter might think it is stupid…but I don't think my oldest daughter would have anything against it.* [Paula]

The concept of subjective norm was closely tied to the sense of independence. This was implicitly dissolved in answers that reassured that it was ultimately their decision to use:

> *…at first, they [friends and family] might think that I'm a, what do you call the, hypochondriac. But I don't care. And they will eventually come around to seeing that I'm taking it as an adult self-interest, a self-directive interest in my own being, my wellbeing.* [Greg]

> *That's my opinion, if I want to do it, that's up to me to do it. If I want to walk…what I want to do, I do.* [Paula]

The significance of the sense of independence to participants surfaced more explicitly when they were asked how much social support their family and friends can provide. Greg stated:

> *I'm sure it's there [the support] but it means taking their time, and making my problem their problem. And that's hard for me to do because of my own attitudes about independence I think. I really resent supervision, which is intrusive and demanding; kinds of stuff like that within the family.* [Greg]

The importance of older adults' independence was highlighted and was demonstrative of an impediment in reaching out for loved one's support in using wearable activity trackers. As a result of this hesitance, older adults may face reduced social approval and support, which may potentially affect the acceptance and future use of wearable activity trackers.

Theme 5: Equipment Characteristics Matter—Display, Battery life, Comfort, and Aesthetics

The equipment characteristics of the 2 wearable activity trackers varied significantly. Participants confirmed that display, battery

life, aesthetics, and comfort had a significant influence on device acceptance.

All participants approved of the significant value of LCD display on the Microsoft Band for quickly accessing information compared with the Mi Band. The Mi Band relied on smartphones for data retrieval and it was a deterrent to acceptance. Furthermore, the noninteractive nature of the Mi Band display demotivated participants from exploring the device functionalities and caused frustration with the device, which is evident from the following quote:

> *The Mi Band, obviously it didn't offer as many options. And it didn't encourage me to do as much exploring, maybe there just wasn't—it wasn't there. I don't know, I kind of gave up because I couldn't figure it out.* [Anita]

A majority of participants preferred Mi Band's longer battery life of 30 days and expressed that the short battery life of the Microsoft Band negatively impacted their use:

> *I was able to wear it [the Microsoft Band] 2, 3 days and charge it and, you know, like you didn't feel you had to do this all the time or you had to be home…because of my lifestyle I am not usually home at certain times.* [Francine]

Participants described the Microsoft Band, which is larger and inflexible, as "uncomfortable" and "rigid." The Mi Band's strap was preferred for its thinness and flexibility. The comfort affected participants' decision to acceptance and use, as shown below:

> *Well, I have very small wrists. So, if it doesn't fit nicely, then it's uncomfortable and is an irritation because it's flying around slipping down onto my hand. It's not comfortable. So yes, it has to be comfortable to wear.* [Paula]

All female participants valued the aesthetics of wearable activity tracker highly in determining inappropriate settings to have them on. On the other hand, the only male participant had an indifferent opinion on the aesthetics of the device, as shown below:

> *Well. I wouldn't wear them out for the evening…Not if I was going out—depending on where I'm going, but they're definitely not formal wear.* [Anita]
>
> *If it looked more like jewelry, I think more people would wear it.* [Paula]
>
> *No [aesthetics don't matter]. I don't see it as a fashion thing.* [Greg]

Equipment characteristics were important factors for participants and they were closely tied to the perceived ease of use and perceived usefulness.

Discussion

Principal Findings

The aim of this study was to explore the attitudes toward and acceptance of 2 specific wearable activity trackers among a sample of Canadian community-dwelling older adults with a mixed-methods study design. A total of 20 older adults were recruited for phase 1 and 4 for phase 2. Most of the participants were frequent computer and smartphone users and had high awareness of wearable activity trackers.

Overall, participants indicated a significantly higher acceptance rate for the Mi Band (16/20, 80%) in comparison with the Microsoft Band (50%). This is also reflected in the statistical analysis where the Mi Band recorded a significantly higher score in the equipment characteristics dimension. This result emphasizes the significance of equipment characteristics in determining the acceptance for wearable activity trackers, especially as the battery life and comfort when compared between the 2 devices were substantially different. Wearable activity trackers are often regarded more personal than other technologies such as computers. This notion was well illustrated in the semistructured interview in which all female participants highly regarded wearable activity trackers as a fashion item, and the ability to conceal it was important in future decisions to use. Mercer and colleagues [30] also reported the aesthetics and subtlety of a device as major reasons for device preference.

A higher rate of acceptance for the Mi Band over the Microsoft Band was in conflict with the findings from the semistructured interviews in which the participants expressed difficulties with accessing data with the Mi Band. Mercer and colleagues [30] also reported a similar finding where older adults who were not familiar with smartphones preferred a wearable activity tracker with a clear display rather than accessing information via smartphones. Additionally, the final acceptance was higher for the device with greater comfort and style rather than its ability to display information. This may indicate that comfort, usability, battery life, and price are more crucial in accepting a wearable activity tracker than the ease of accessibility of data.

These findings have important implications for initiatives aimed at aging-in-place and the selection of technologies. To ensure long-term and continuous usage of wearable activity trackers selected to enable aging-in-place, researchers should take the steps to ensure equipment characteristics such as aesthetics, comfort, and battery life are in line with older adults' expectations, as they have the potential to deter usage and acceptance. In this study, despite the preference of the Microsoft Band for easy data accessibility, participants reported high self-awareness of their physical activity level and health status for both devices. This warrants future research studies to investigate how varying ease of data access impacts self-awareness in finer granularity and ultimately its contribution to behavior changes.

In this study, price or how much a participant was willing to pay for a wearable activity tracker demonstrated the highest correlation (rho=.49, P=.001) with acceptance for both the Mi Band and the Microsoft Band. This was further confirmed through the semistructured interviews in which participants supported a greater likelihood of using a wearable activity tracker if it were free. In a systematic review that assessed technology for aging-in-place, high cost negatively influenced preimplementation technology acceptance [17]. Another study that examined the appropriateness of price in relation to the quality of product (ie, price reasonability) also identified a

significant influence on technology acceptance for fitness-oriented wearable activity trackers for the general population [40].

This study is one of the first studies to confirm the significance of device cost post implementation for older adults. Understanding the relationship between the price and postimplementation acceptance is important as it indicates that the cognitive trade-off between perceived value of wearable activity trackers and monetary cost are not likely to occur for older adults. This result provides an important consideration for health promotion efforts (such as increasing physical activity, going outdoors, or increasing awareness of one's own health) aimed toward older adults who use wearable activity trackers. Older adults who have purchased wearable activity trackers out of pocket are likely to have accepted the technology, and thus, promotion efforts should consider allocating the budget on delivering value-enhancing contents. On the other hand, health efforts that target older adults who have not considered smart wearable devices should consider subsidizing or providing the devices free of cost to increase acceptance.

Although participants claimed in the semistructured interviews that privacy was not a determining factor for using the wearable activity trackers, the quantitative results indicate the opposite. Out of the 3 privacy questions, 2 questions (L19 and L20) demonstrated a correlation between perceived privacy risk and technology rejection (rho=−.43 and −.38, respectively; $P<.05$). In other words, participants with high perceived privacy concern were moderately associated with more technology acceptance. This is in contrast with the previous research study by Gao and colleagues [40] in which perceived privacy concerns were significantly associated with wearable technology acceptance among the general population. Such inconclusive and contradicting findings may indicate that older adults had a lack of understanding of potential privacy implications. Uncertainties around privacy implications for emerging technologies were frequently identified themes among older adults [49,50]. Despite the uncertainties over privacy implications, older adults are willing to share information and compromise privacy when the technology is perceived as beneficial [17], improve or maintain independence [51], and valuable [52]. On the basis of these studies, one plausible explanation for the findings of this study is that the increased perceived benefits of wearable activity trackers may have outweighed the high privacy concerns. This may have led to the association between high technology acceptance and high privacy concern. Furthermore, Gao and colleagues [40] have also reported a weaker relationship between privacy and acceptance for medical devices compared with fitness-oriented devices. This trend indicates that older adults are willing to share fitness data when the perceived value is high. Notwithstanding, little is known about the potential privacy threats related to wearable activity trackers among older adults. In a study that investigated older adults' perception around privacy for wearable device, a strong relationship was found between reduced perceived privacy risk and heightened and transparent legal consequences for companies [50]. The same study also discovered data sharing control as an important aspect for ensuring privacy [50]. Collection of wearable activity trackers also poses questions to public health with regard to data ownership, classification of the data as health information, and anonymity and reidentification [53]. The topic of data privacy of smart wearables and ubiquitous health will require a major attention as, currently, they are not regulated and solely rely on self-governing and oversight.

Limitations

Several limitations are present in this study. A small sample of 20 older adults was recruited using a convenience sampling method, and the participants in this study were selected from a small number of geographical locations. This study investigated only 2 wearable activity trackers, and they may not be representative of the full range of devices and functionalities currently available. Smart wearables evolve at a very rapid rate, and the functionalities can change enormously with the introduction of new sensor technologies. While the study period was lengthier than most previous research studies of wrist-worn activity tracker acceptance, extended-term acceptance could not be comprehensively explored over a period of 6 weeks. Another limitation of this study is that it did not investigate any aspect of the data collected from the wearable activity trackers, as the purpose of this study was to understand the technology acceptance for older adults. Only 4 semistructured interviews with purposive participant recruitment were conducted. The aim of the interviews was to provide depth and context to the data collected by the questionnaires, and as such, saturation was not an end-goal. Finally, no participants of the semistructured interviews had accepted the Microsoft Band, which may have biased the qualitative results favorable to the Mi Bands.

Conclusions

This exploratory study generated several important findings about older adults' acceptance of and attitudes toward the 2 wrist-worn activity trackers. The acceptance of wearable activity trackers, by nature of how they are worn, is highly influenced by the equipment characteristics, including comfort, battery life, and especially aesthetics. The cost of the device is a strong indicator of technology acceptance, and its importance continues post implementation as well. Older adults were open to sharing health information generated by wearable activity trackers as long as their strengths related to health benefits and maintaining independence are clearly demonstrated. However, a lack of understanding of potential privacy risks was evident, hindering informed decision making by older adults. These findings provide guidance for future health promotion efforts that plan to target wearable activity tracker usage. Future research should focus on not only solidifying the value of smart wearables for older adults but also identifying its potential negative impact such as privacy risks. This study is a small but important step toward understanding the acceptance and usage of wrist-worn activity trackers in older adults. The findings will provide guidance for future large-scale studies.

Acknowledgments

This work was supported by a grant from the AGE-WELL Network of Centres of Excellence.

Authors' Contributions

AP and JL formulated and designed the research. The design of the research was refined with the help of PS and JT. AP carried out the data collection. AP, ON, BK, and JL conducted the quantitative analyses. AP and BK performed the qualitative analyses. BK wrote the manuscript, with input from AP. JL contributed to the development and refinement of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Technology acceptance questionnaire.

[PDF File (Adobe PDF File), 23KB - mhealth_v5i11e173_app1.pdf]

Multimedia Appendix 2

Wilcoxon signed-ranked test for individual items from the technology acceptance questionnaire.

[PDF File (Adobe PDF File), 13KB - mhealth_v5i11e173_app2.pdf]

References

1. Statistics Canada. 2015 Sep 29. Canada's population estimates: age and sex, July 1, 2015 URL: http://www.statcan.gc.ca/daily-quotidien/150929/dq150929b-eng.htm [WebCite Cache ID 6nu5EF64q]
2. Canadian Institute for Health Information. National health expenditure trends, 1975 to 2013. Ottawa, ON: CIHI; 2013. URL: https://secure.cihi.ca/free_products/NHEXTrendsReport_EN.pdf [WebCite Cache ID 6r2w1bJrk]
3. Canadian Institute for Health Information. Health care in Canada, 2011: a focus on seniors and aging. Ottawa, ON: CIHI; 2011. URL: https://secure.cihi.ca/free_products/HCIC_2011_seniors_report_en.pdf [WebCite Cache ID 6FIMyTE9L]
4. Wiles JL, Leibing A, Guberman N, Reeve J, Allen RE. The meaning of "aging in place" to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: 10.1093/geront/gnr098] [Medline: 21983126]
5. Bacsu JR, Jeffery B, Johnson S, Martz D, Novik N, Abonyi S. Healthy aging in place: supporting rural seniors' health needs. Online J Rural Nurs Health Care 2012;12(2):77-87.
6. Chapin R, Dobbs-Kepper D. Aging in place in assisted living: philosophy versus policy. Gerontologist 2001 Feb;41(1):43-50. [Medline: 11220814]
7. Hollander MJ, Chappell NL. A comparative analysis of costs to government for home care and long-term residential care services, standardized for client care needs. Can J Aging 2007;26 Suppl 1:149-161. [doi: 10.3138/cja.26.suppl_1.149] [Medline: 18089532]
8. Chappell NL, Dlitt BH, Hollander MJ, Miller JA, McWilliam C. Comparative costs of home care and residential care. Gerontologist 2004 Jun;44(3):389-400. [Medline: 15197293]
9. Graybill EM, McMeekin P, Wildman J. Can aging in place be cost effective? A systematic review. PLoS One 2014;9(7):e102705 [FREE Full text] [doi: 10.1371/journal.pone.0102705] [Medline: 25058505]
10. Marek KD, Stetzer F, Adams SJ, Popejoy LL, Rantz M. Aging in place versus nursing home care: comparison of costs to Medicare and Medicaid. Res Gerontol Nurs 2012 Apr;5(2):123-129. [doi: 10.3928/19404921-20110802-01] [Medline: 21846081]
11. Davey J. "Ageing in place": The views of older homeowners on home maintenance, renovation and adaptation. Soc Policy J New Zealand 2006;27:128 [FREE Full text]
12. Mylively. 2015. Live!y - Personal Emergency Response - Reimagined URL: http://www.mylively.com/ [accessed 2017-06-07] [WebCite Cache ID 6r2wkJqSr]
13. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile phone apps to improve medication adherence: a systematic stepwise process to identify high-quality apps. JMIR Mhealth Uhealth 2016 Dec 02;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
14. Preventicesolutions. 2015. Bodyguardian heart URL: http://www.preventicesolutions.com/technologies/body-guardian-heart.html [accessed 2017-06-07] [WebCite Cache ID 6r2wu8j4q]
15. Carepredict. 2015. Empowering better senior care URL: https://www.carepredict.com [accessed 2017-06-07] [WebCite Cache ID 6r2x4mCJ3]

16. Demiris G, Rantz M, Aud M, Marek K, Tyrer H, Skubic M, et al. Older adults' attitudes towards and perceptions of "smart home" technologies: a pilot study. Med Inform Internet Med 2004 Jun;29(2):87-94. [doi: 10.1080/14639230410001684387] [Medline: 15370989]
17. Peek ST, Wouters EJ, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJ. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
18. Salah H, MacIntosh E, Rajakulendran N. Marsdd. Toronto: MaRS; 2014. Wearable tech: leveraging Canadian innovation to improve health URL: https://www.marsdd.com/wp-content/uploads/2015/02/MaRSReport-WearableTech.pdf [accessed 2017-06-07] [WebCite Cache ID 6r2xIGIIk]
19. Accenture. Insight driven health: silver surfers continue to make waves in ehealth. But are healthcare organizations swimming against the tide? URL: https://www.accenture.com/t20160829T222305__w__/us-en/_acnmedia/PDF-29/Accenture-Silver-Surfers-Continue-to-Make-Waves.pdf [accessed 2017-03-30] [WebCite Cache ID 6pMLflNvl]
20. Accenture. 2016. Insight driven health: Accenture 2016 consumer survey on patient engagement URL: https://www.accenture.com/t20160629T045303__w__/us-en/_acnmedia/PDF-6/Accenture-Patients-Want-A-Heavy-Dose-of-Digital-Research.pdf [accessed 2017-06-07] [WebCite Cache ID 6r2xQtgJi]
21. 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]
22. Clifton L, Clifton DA, Pimentel MA, Watkinson PJ, Tarassenko L. Predictive monitoring of mobile patients by combining clinical observations with data from wearable sensors. IEEE J Biomed Health Inform 2014 May;18(3):722-730. [doi: 10.1109/JBHI.2013.2293059] [Medline: 24808218]
23. Kim BY, Lee J. Smart devices for older adults managing chronic disease: a scoping review. JMIR Mhealth Uhealth 2017 May 23;5(5):e69 [FREE Full text] [doi: 10.2196/mhealth.7141] [Medline: 28536089]
24. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014 Apr 09;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
25. Chen CC, Shih HS. A study of the acceptance of wearable technology for consumers: an analytical network process perspective. Int J Anal Hierarchy Process 2014:1-5.
26. 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;12(1):159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
27. 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]
28. Vooijs M, Alpay LL, Snoeck-Stroband JB, Beerthuizen T, Siemonsma PC, Abbink JJ, et al. Validity and usability of low-cost accelerometers for internet-based self-monitoring of physical activity in patients with chronic obstructive pulmonary disease. Interact J Med Res 2014;3(4):e14 [FREE Full text] [doi: 10.2196/ijmr.3056] [Medline: 25347989]
29. 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;4(2):e35 [FREE Full text] [doi: 10.2196/mhealth.5120] [Medline: 27076486]
30. 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;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
31. Karahanoğlu A, Erbuğ C. Perceived qualities of smart wearables: determinants of user acceptance. : ACM Press; 2011 Presented at: Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces - DPPI'11; June 22-25, 2011; Milano, Italy. [doi: 10.1145/2347504.2347533]
32. Shih PC, Han K, Poole ES, Rosson MB, Carroll JM. Use and adoption challenges of wearable activity trackers. 2015 Presented at: iConference 2015 Proceedings; 2015; Illinois p. 1 URL: https://www.ideals.illinois.edu/handle/2142/73649
33. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340 [FREE Full text] [doi: 10.2307/249008]
34. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
35. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
36. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 01:425-478 [FREE Full text]
37. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the almere model. Int J Soc Robot 2010 Dec 01;2(4):361-375. [doi: 10.1007/s12369-010-0068-5]

38. Pan S, Jordan-Marsh M. Internet use intention and adoption among Chinese older adults: from the expanded technology acceptance model perspective. Comput Human Behav 2010 Sep;26(5):1111-1119 [FREE Full text] [doi: 10.1016/j.chb.2010.03.015]
39. Yang H, Yu J, Zo H, Choi M. User acceptance of wearable devices: an extended perspective of perceived value. Telematics Informatics 2016 May;33(2):256-269. [doi: 10.1016/j.tele.2015.08.007]
40. Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. Ind Manag Data Syst 2015 Oct 19;115(9):1704-1723 [FREE Full text] [doi: 10.1108/IMDS-03-2015-0087]
41. Mihailidis A, Cockburn A, Longley C, Boger J. The acceptability of home monitoring technology among community-dwelling older adults and baby boomers. Assist Technol 2008;20(1):1-12. [doi: 10.1080/10400435.2008.10131927] [Medline: 18751575]
42. Hsiung PC. Utsc.utoronto.ca. 2010. Lives and legacies: a guide to qualitative interviewing URL: http://www.utsc.utoronto.ca/~pchsiung/LAL/reflexivity [accessed 2017-06-07] [WebCite Cache ID 6r2yuPKXi]
43. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
44. Fensli R, Pedersen PE, Gundersen T, Hejlesen O. Sensor acceptance model - measuring patient acceptance of wearable sensors. Methods Inf Med 2008;47(1):89-95. [Medline: 18213434]
45. Davis F, Bagozzi R, Warshaw P. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003 [FREE Full text] [doi: 10.1287/mnsc.35.8.982]
46. Thompson RL, Higgins CA, Howell JM. Personal computing: toward a conceptual model of utilization. MIS Q 1991 Mar;15(1):125-143. [doi: 10.2307/249443]
47. Taylor JW. The role of risk in consumer behavior. J Mark 1974 Apr;38(2):54-60 [FREE Full text] [doi: 10.2307/1250198]
48. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. London: Sage Publications, Inc; Mar 14, 2013.
49. Cheung C, Bietz MJ, Patrick K, Bloss CS. Privacy attitudes among early adopters of emerging health technologies. PLoS One 2016;11(11):e0166389 [FREE Full text] [doi: 10.1371/journal.pone.0166389] [Medline: 27832194]
50. Segura Anaya LH, Alsadoon A, Costadopoulos N, Prasad PW. Ethical implications of user perceptions of wearable devices. Sci Eng Ethics 2017 Feb 02:1-28. [doi: 10.1007/s11948-017-9872-8] [Medline: 28155094]
51. Mynatt ED, Melenhorst AS, Fisk AD, Rogers WA. Aware technologies for aging in place: understanding user needs and attitudes. IEEE Pervasive Comput 2004 Apr;3(2):36-41 [FREE Full text] [doi: 10.1109/MPRV.2004.1316816]
52. Courtney KL. Privacy and senior willingness to adopt smart home information technology in residential care facilities. Methods Inf Med 2008;47(1):76-81. [Medline: 18213432]
53. Paul G, Irvine J. Privacy implications of wearable health devices. New York: ACM; 2014 Sep Presented at: Proceedings of the 7th International Conference on Security of Information and Networks; September 9-11, 2014; Glasgow, Scotland p. 117 URL: http://dl.acm.org/citation.cfm?doid=2659651.2659683 [doi: 10.1145/2659651.2659683]

Abbreviations

AUC: area under the curve
LCD: liquid-crystal display
LED: light-emitting diodes
TAM: technology acceptance model
UTAUT: unified theory of acceptance and use of technology
UV: ultraviolet

Edited by G Eysenbach; submitted 14.06.17; peer-reviewed by S Berrouiguet, N Ridgers, M Lang, R Robinson; comments to author 25.08.17; revised version received 06.09.17; accepted 23.09.17; published 15.11.17.

Please cite as:
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;5(11):e173
URL: http://mhealth.jmir.org/2017/11/e173/
doi:10.2196/mhealth.8211
PMID:29141837

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org