JMIR mHealth and uHealth

Impact Factor (2022): 5.0
Volume 8 (2020), Issue 10 ISSN 2291-5222 Editor in Chief: Laurie Buis, PhD, MSI

Contents

Original Papers

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

Reviews

XSL•FO
RenderX

Corrigenda and Addendas

XSL•FO
RenderX

Original Paper

Symbol Digit Modalities Test Variant in a Smartphone App for Persons With Multiple Sclerosis: Validation Study

Pim van Oirschot[1], PhD; Marco Heerings[2,3], MANP; Karine Wendrich[4], MSc; Bram den Teuling[1], MSc; Marijn B Martens[5,6], PhD; Peter J Jongen[7,8], PhD

[1]Orikami Digital Health Products, Nijmegen, Netherlands
[2]Dutch National MS Foundation, Rotterdam, Netherlands
[3]Radboud University Medical Center, Nijmegen, Netherlands
[4]Faculty of Science, Institute for Science in Society, Radboud University, Nijmegen, Netherlands
[5]Drug Target ID, Nijmegen, Netherlands
[6]NeuroDrug Research BV, Nijmegen, Netherlands
[7]Department of Community & Occupational Medicine, University Medical Centre Groningen, Groningen, Netherlands
[8]MS4 Research Institute, Nijmegen, Netherlands

Corresponding Author:
Pim van Oirschot, PhD
Orikami Digital Health Products
Ridderstraat 29
Nijmegen, 6511TM
Netherlands
Phone: 31 243010100
Email: pim@mssherpa.nl

Abstract

Background: The decline of cognitive processing speed (CPS) is a common dysfunction in persons with multiple sclerosis (MS). The Symbol Digit Modalities Test (SDMT) is widely used to formally quantify CPS. We implemented a variant of the SDMT in MS sherpa, a smartphone app for persons with MS.

Objective: The aim of this study was to investigate the construct validity and test-retest reliability of the MS sherpa smartphone variant of the SDMT (sSDMT).

Methods: We performed a validation study with 25 persons with relapsing-remitting MS and 79 healthy control (HC) subjects. In the HC group, 21 subjects were matched to the persons with MS with regard to age, gender, and education and they followed the same assessment schedule as the persons with MS (the "HC matched" group) and 58 subjects had a less intense assessment schedule to determine reference values (the "HC normative" group). Intraclass correlation coefficients (ICCs) were determined between the paper-and-pencil SDMT and its smartphone variant (sSDMT) on 2 occasions, 4 weeks apart. Other ICCs were determined for test-retest reliability, which were derived from 10 smartphone tests per study participant, with 3 days in between each test. Seven study participants with MS were interviewed regarding their experiences with the sSDMT.

Results: The SDMT scores were on average 12.06% higher than the sSDMT scores, with a standard deviation of 10.68%. An ICC of 0.838 was found for the construct validity of the sSDMT in the combined analysis of persons with MS and HC subjects. Average ICCs for test-retest reliability of the sSDMT for persons with MS, the HC matched group, and the HC normative group were 0.874, 0.857, and 0.867, respectively. The practice effect was significant between the first and the second test of the persons with MS and the HC matched group and trivial for all other test-retests. The interviewed study participants expressed a positive attitude toward the sSDMT, but they also discussed the importance of adapting a smartphone cognition test in accordance with the needs of the individual persons with MS.

Conclusions: The high correlation between sSDMT and the conventional SDMT scores indicates a very good construct validity. Similarly, high correlations underpin a very good test-retest reliability of the sSDMT. We conclude that the sSDMT has the potential to be used as a tool to monitor CPS in persons with MS, both in clinical studies and in clinical practice.

(JMIR Mhealth Uhealth 2020;8(10):e18160) doi:10.2196/18160

XSL•FO
RenderX

KEYWORDS

relapsing-remitting multiple sclerosis; cognition; processing speed; mobile phone

Introduction

Background

Multiple sclerosis (MS) is a chronic disease in which the body's immune system mistakenly attacks the isolating sheath (myelin) that surrounds the nerve fibers in the central nervous system. MS may affect the brain, the spinal cord, and the optic nerves, and disrupt the information flow across the affected nerves. This may cause a variety of symptoms, including loss of vision (optic neuritis), muscle weakness, sensory symptoms, cognitive dysfunction, altered coordination, and fatigue. For most persons with MS, sudden relapses contribute to the unpredictability of the disease, which makes it difficult to devise a treatment plan specific to an individual patient.

The prevalence of cognitive dysfunction in persons with MS is estimated to be between 40% and 70%, depending on the research setting [1-4]. This can include dysfunction in visuospatial processing, cognitive processing speed (CPS), working memory, executive functioning, verbal and visual learning, as well as episodic memory. CPS decline is among the first cognitive signs and the most commonly observed cognitive deficit in persons with MS. Importantly, CPS decline has a significant impact on the quality of life [3,5]. CPS is usually assessed by measuring the amount of information processed in a unit of time or the time needed to process a given amount of information.

The Paced Auditory Serial Addition Test (PASAT) [6] and the Symbol Digit Modalities Test (SDMT) [7] are the most widely used tests to formally quantify CPS in MS, and they focus on auditory CPS and visuospatial CPS, respectively [8-10]. Both are included in the Rao Brief Repeatable Battery of Neuropsychological tests, amongst tests for learning and executive function. Sometimes, the Brief Repeatable Battery of Neuropsychological tests is administered by clinicians during check-ups [2,11]. Both tests are part of the Minimal Assessment of Cognitive Function [12]. Notably, the more compact Brief International Cognitive Assessment for MS (BICAMS) test battery prefers SDMT over PASAT [13]. The SDMT score has a strong correlation with the Expanded Disability Status Scale (EDSS) score [14]—a stronger correlation than the PASAT scores [15,16]—as well as with abnormalities seen in magnetic resonance images such as in brain lesion volume, cerebral atrophy, diffusion tensor indices of normal appearing brain tissue, and retinal nerve fiber layer thickness [17]. Although both the PASAT and SDMT are highly sensitive in detecting cognitive impairment, the SDMT has higher acceptability among patients, is easier to administer, and has slightly higher sensitivity than PASAT [13,15]. Moreover, findings suggest that the SDMT more accurately reflects the qualitative nature of self-reported cognitive impairment and should perhaps replace the PASAT as part of the MS Functional Composite (MSFC) [18]. Finally, the SDMT has been found to be relatively resistant to practice effects and is therefore useful for serial testing and screening [19].

Although the SDMT has been proven to be clinically relevant, it is not regularly used in standard care. This is partly because patients often do not complain about cognitive changes during consultation with a health care professional. It is also because of the time factor and administrative burden, despite the relatively short duration of one assessment, that is, 90 seconds. Paper-and-pencil variants of the SDMT have been existing for quite some time [20,21], but digital variants have been developed only in recent years for computers [18], tablets [22,23], or smartphones [24,25]. It is believed that e-versions will make it easier to monitor CPS. However, to our knowledge, no scientific papers about the validity of the smartphone variants of the SMDT have been published to date.

MS sherpa (Orikami Digital Health Products) is a smartphone app intended to support the monitoring of persons with MS, in order to give patients and their health care professionals personalized insight into the presence and progress of MS-related symptoms and signs. MS sherpa contains a smartphone variant of the SDMT (sSDMT). This sSDMT works as follows: a large black symbol present in the middle of the screen has to be matched to the correct digit. Matching can be done at the bottom of the screen. Once an answer is given, a new symbol appears. The sSDMT score is the number of correct answers in this 90-second test. A small stopwatch counting down is shown above the numbers at the bottom of the screen. The key is shown at the top of the screen. The same key is used during 1 assessment, but it varies between assessments; the symbols are randomly matched to digits in the key each time a test is done. These symbols are different from those in the original SDMT since these are protected by copyrights.

Note that contrary to the original SDMT, in sSDMT, one cannot look ahead as to which symbol is to be matched next. However, also in the original SDMT, skipping a symbol in the presented sequence is not allowed. There are 2 other aspects in which the sSDMT differs from the SDMT. First, in the original SDMT, the first 26 items are selected from the first 6 symbols in the key. In the sSDMT, symbols are selected from the full key during the whole assessment. This was also the case in the first version of other alternative versions of the SDMT [5]. Second, instead of using the first 10 "practice" items that can be matched with guidance in the original SDMT, it is possible to practice the sSDMT by pressing the corresponding button on the instruction screen. In a practice test, symbols are again randomly matched to digits in the key, which is therefore different from the key in the "real" sSDMT. The rationale for not matching exactly the "practice" phase to the original SDMT was that home monitoring can be done frequently and practicing before every test should not be mandatory, from a usability perspective.

Objective

We aimed to study the construct validity and test-retest reliability of the sSDMT that is implemented in MS sherpa as an assessment for CPS and present our findings in this paper, along with respondents' experiences with the sSDMT. The validation of the sSDMT was part of the MS Self study. The

XSL•FO
RenderX

MS Self study was a validation study during which study participants performed self-monitoring assessments during 4 weeks with a precursor of MS sherpa, which was called the "Mijn Kwik" app and a Fitbit Charge 2 wearable. Besides investigating the validity of the sSDMT, another research objective of the MS Self study was to investigate first the experiences with digital self-monitoring through smartphone apps and activity trackers of persons with MS by interviewing 7 study participants with MS. Furthermore, the study aimed to validate a smartphone variant of the 2-minute walking test (s2MWT) and a smartphone walking balance test (sWBT). In particular, we investigated if the outcomes of the smartphone tests were in agreement with the outcomes of corresponding clinical tests, namely, SDMT, the 2-minute walking test (2MWT), and the Timed Up and Go test (TUG). The other results from the interviews as well as a description of the methodology have recently been published [26]. Separate research papers on the s2MWT, the sWBT, and the Fitbit data are in preparation. A poster with the preliminary results of the MS Self study [27] can be found on the MS sherpa website [28]. This poster also contains an image of the sSDMT.

Methods

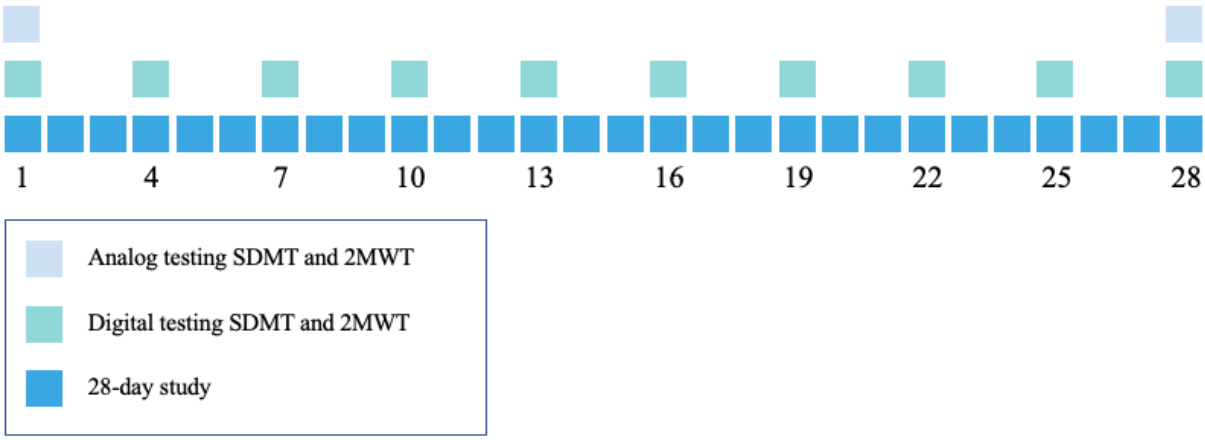
Study Design

This study was performed in 25 persons with relapsing-remitting MS and 2 groups of healthy control (HC) subjects (n=79). The HC subjects in the first control group (HC matched, n=21) were matched to the persons with MS with regard to age, gender, and education. Five education categories were defined: "secondary general education," "senior secondary general education and preuniversity education," "secondary vocational education," "higher professional education," and "scientific education". The second control group (HC normative, n=58) was set up to determine the normal distribution for the smartphone test results.

The inclusion criteria were as follows: (1) willing to participate and capable of doing all the tasks mentioned in the protocol, (2) able and willing to use own smartphone, which should be an iPhone 5 or a newer Apple device, or a phone with an Android operating system version 6 or higher, and (3) aged between 20 years and 50 years. For persons with MS, the following additional inclusion criteria were set: (1) diagnosis of relapsing-remitting MS for more than one year and (2) having an EDSS score between 1.5 and 6.5. A maximal EDSS score of 6.5 is needed to perform the 2MWT. A minimal EDSS score of 1.5 was chosen to find a difference in the 2MWT results for persons with MS and HC subjects.

On the first and the last day of the study, persons with MS and the HC matched group came to the premises of the Dutch National MS Foundation (Nationaal MS Fonds [NMSF]) to perform a paper-and-pencil SDMT, followed by a simultaneously performed 2MWT and a s2MWT in the open air. Later that day, and at most 24 hours later, they also performed the sSDMT in their home environment. During the 4-week follow-up, these study participants performed the s2MWT and sSDMT tests at home, once every 3 days, that is, 10 times in total. A schematic overview of the study design is given in Figure 1. The EDSS was administered to every person with MS by MH on the first day of the study.

Figure 1. Overview of the study design and assessment scheme. SDMT: Symbol Digit Modalities Test; 2MWT: 2-minute walking test.



The HC subjects in the HC normative group were instructed to do the s2MWT and the sSDMT 3 times in total, with 1 week in between the assessments. From these tests and from the 10 home assessments of the other study participants, the test-retest reliability of the sSDMT was determined. The combined test results "SDMT-sSDMT" for the persons with MS and the HC matched group were used to determine the construct validity of the sSDMT. This combination will hereafter be referred to as the validation assessment. Note that there are 2 validation assessments per study participant: one in the beginning and one at the end of the study.

Recruitment

The persons with MS and HC matched group were recruited via the NMSF (Rotterdam, the Netherlands). The HC normative group was recruited via the social network of the app manufacturer.

XSL•FO
RenderX

Ethical Approval and Informed Consent

This study was approved by the medical ethical committee METC Brabant (Tilburg, the Netherlands) under protocol number NL61291.028.17. All study participants agreed to the privacy statement of the Mijn Kwik app prior to first use. Persons with MS and the HC matched group signed an informed consent letter on paper.

Data Collection

Data were collected between May 2017 and May 2018. The ages of the HC subjects in the HC matched group were matched as closely as possible to that of the persons with MS, which led to at most 2 years difference for 86% (18/21) of the pairs. Approximately 90% (19/21) of the pairs matched in gender. An exact match in education category was obtained for 29% (6/21) of the pairs.

Data Analysis

In the analyses performed in this study, the SDMT score on the first day of the study was compared with the first sSDMT score and the SDMT score on the last day of the study was compared with the last sSDMT score. In the data cleaning process of the sSDMT data, the home assessments in which the study participants had a score below 20 were removed because these were outliers. It was assumed that the users were distracted during these tests. A score below 20 occurred in less than 1% (3/423) of all the sSDMTs that were done.

Statistical Analysis

Following the study protocol, the significance level α was set to 5%. Two-sided *t* tests were conducted to compare the SDMT score with the sSDMT score. For the smartphone tests, the median test score of each study participant was taken in the *t* test analysis. Internal consistency was evaluated and quantified using Cronbach α, in which an α value larger than .7 was defined to be acceptable [29,30]. Test-retest reliability was determined by calculating the intraclass correlation coefficient (ICC) between measurements at different times. ICCs smaller than 0.59 were considered to be "moderate," ICCs between 0.60 and 0.79 were considered "good," and ICCs above 0.80 were considered "very good" [31,32]. The effect size Cohen *d* was determined to investigate the practice effect. A Cohen *d* below 0.20 was seen as trivial [33].

To investigate the construct validity of the sSDMT, we calculated the (1) ICC between the SDMT and the sSDMT and the (2) Pearson correlation coefficient between the SDMT and the sSDMT (the distributions were checked for normality by applying Shapiro-Wilk tests).

For these statistical tests, we used the same acceptance criteria as for the test-retest reliability: values smaller than 0.59 were considered to be "moderate," values between 0.60 and 0.79 were considered "good," and values above 0.80 were considered "very good."

We calculated different ICCs for the construct validity and the test-retest reliability, following the McGraw and Wong [34] convention. Although the "model" selection (two-way mixed effects) and the "type" selection (single rater/measurement) of both ICCs are the same, the "definition" (consistency or absolute agreement) is different [35]. ICCs of type "ICC(3,1)" were calculated to compare the SDMT and the sSDMT. These ICCs have definition "consistency," since these 2 test scores were not expected to exactly match. ICCs of type "ICC(A,1)" were calculated to determine test-retest reliability. These ICCs have definition "absolute agreement," since the test scores per individual were not expected to change within the 1-month follow-up of the study.

We checked if the distributions of the number of correct answers on the first sSDMT that was done for the 3 populations in this study were normally distributed via Shapiro-Wilk tests and if the 2 groups of the HC subjects had the same underlying distribution via a 2-sample Kolmogorov-Smirnov test. Finally, we did a two-sided *t* test to confirm the difference between persons with MS and HC subjects.

Results

Participant Demographics

The mean, median, and standard deviation in the ages for the various groups are listed in Table 1, as well as the gender, the number of participants in each education category (for the HC normative group, no education information was collected), and the mean, median, and standard deviation in the EDSS score (persons with MS).

XSL•FO
RenderX

Table 1. Participant demographics.

Characteristics	Persons with MS[a], n=25	HC[b] matched, n=21	HC normative, n=58
Age (years)			
Mean (SD)	40 (8)	37 (8)	34 (8)
Median	43	36	32
Gender			
Female, n (%)	23 (92)	17 (81)	29 (50)
Male, n (%)	2 (8)	4 (19)	29 (50)
Education			
Secondary general education, n (%)	1 (4)	3 (14)	—[c]
Senior secondary general education and preuniversity education, n (%)	3 (12)	3 (14)	—
Secondary vocational education, n (%)	8 (32)	3 (14)	—
Higher professional education, n (%)	9 (36)	6 (29)	—
Scientific education, n (%)	4 (16)	6 (29)	—
EDSS[d] score			
Mean (SD)	3.1 (1.4)	N/A[e]	N/A
Median	3.0	N/A	N/A

[a]MS: multiple sclerosis.

[b]HC: healthy control.

[c]Not available.

[d]EDSS: Expanded Disability Status Scale.

[e]N/A: not applicable.

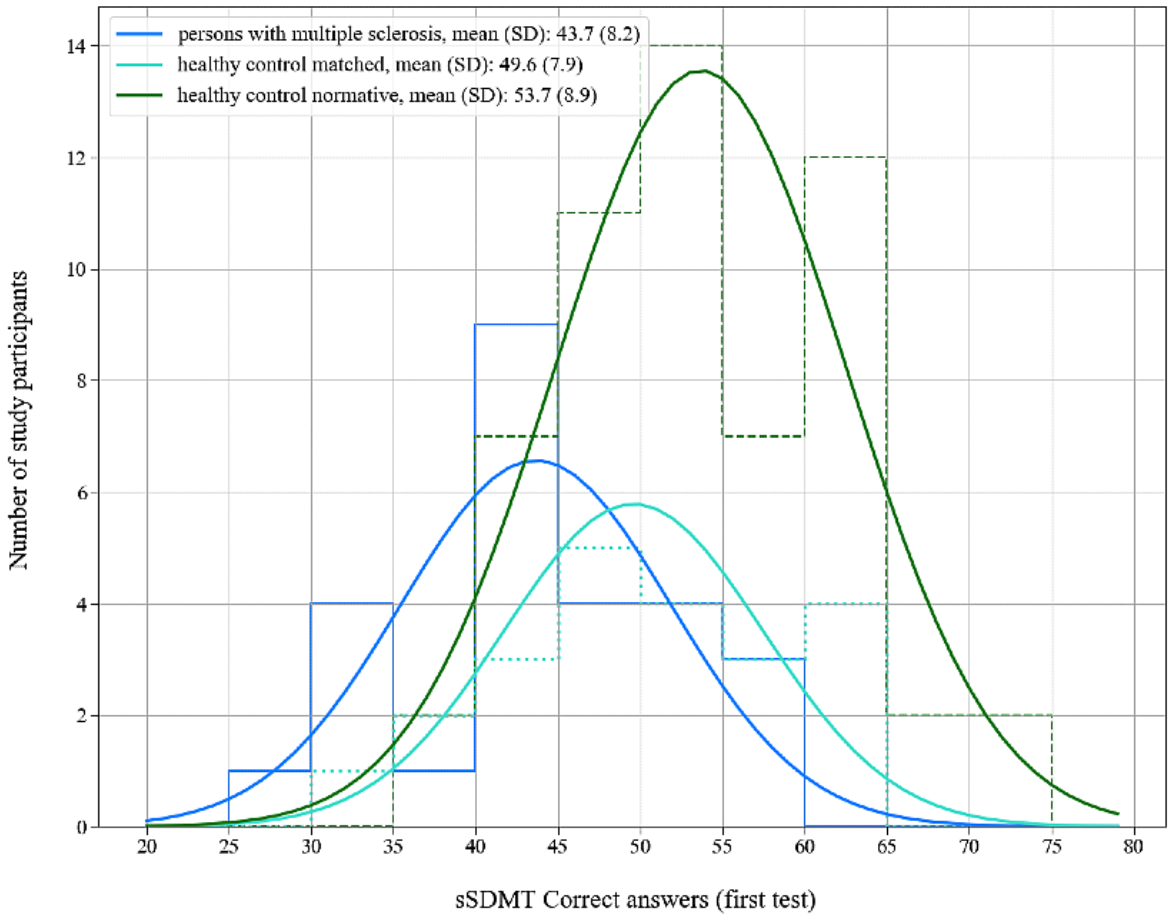
Besides these 104 study participants, 2 more persons were recruited, but 1 person with MS dropped out of the study because she was experiencing a relapse and 1 HC dropped out because she did not like wearing Fitbit.

Distinctions Between Persons with MS and HC Subjects

Figure 2 shows the number of correct answers for the first sSDMT in the persons with MS and in the 2 control groups. The number of correct answers was normally distributed for all 3 groups as calculated by the Shapiro-Wilk tests (P=.49, P=.29, and P=.37 for persons with MS, the HC matched group, and the HC normative group, respectively). The 2 groups of HC subjects have the same underlying distribution as confirmed using a 2-sample Kolmogorov-Smirnov test (Kolmogorov-Smirnov statistic=0.26, P=.20). Independent 2-sample t tests between persons with MS versus the HC matched group and persons with MS versus the HC normative group confirmed that the sSDMT can distinguish between persons with MS and HC subjects at the group level (P=.02 and P<.001, respectively).

XSL•FO
RenderX

Figure 2. Distributions in the number of correct answers on the first test done with the sSDMT for the 3 groups in this study. The thin solid line shows the distribution for persons with multiple sclerosis, the dotted line shows the distribution for the healthy control matched group, and the dashed line shows the distribution for the healthy control normative group. The thick solid lines represent Gaussian fits to the distributions, of which the means (SD) are shown in the legend. sSDMT: smartphone variant of Symbol Digit Modalities Test.



Construct Validity

Two variants of the conventional SDMT were used in the validation assessments: the original paper version of the SDMT [7] (abbreviated as "SDMT") and a paper version of the sSDMT (a variant of the SDMT, hereafter referred to as "vSDMT"). Since the SDMT is copyright protected, the vSDMT was printed in the study protocol. This was misinterpreted by some members of the research team (PvO and MH) as an extra variant of the SDMT to be investigated and was randomly assigned to more than half of the study participants. Different variants could therefore be used in the 2 validation assessments of a given person. Contrary to the regular instructions for an SDMT, study participants did not get 10 "practice" items to be matched with guidance in this study neither for the vSDMT nor for the SDMT. In total, 37 SDMTs and 55 vSDMTs were conducted. In this section, we present the comparison between the SDMT and the sSDMT (raw data in Multimedia Appendix 1). The raw data of the validation assessments with vSDMTs are listed in Multimedia Appendix 2.

The construct validity was determined from 37 validation assessments with the SDMT. One of these assessments was the first validation assessments of a dropout. The vSDMT score of the first validation assessments of the other dropout is included in the Multimedia Appendix 2. However, including these assessments did not increase the total number of validation assessments since 1 HC subject did 2 invalid SDMTs. This person did not match the symbols with the digits in the order of appearance on the sheet of paper, but instead started matching all symbols of a specific kind first, both on the first and on the last day of the study.

It was found that, on average, 6.62 less correct answers were given on the sSDMT than on the SDMT (Figure 3). A Shapiro-Wilk test on the distribution of the differences accepted normality (P=.15). This allows us to determine the 95% CI, at 1.96*s around the mean difference, with SD=6.13 correct answers, which is visualized in the Bland-Altman plot (Figure 4). In this plot, the difference between the 2 tests is shown as a percentage of the mean of the 2 test outcomes on the vertical axis, and the mean is shown on the horizontal axis. The number of correct answers on the sSDMT is, on average, 12.06% (6.62/54.80) lower than the number of correct answers on the SDMT (P<.001). The 1.96*s above and below this average results in a difference of 8.87% more and 33.00% less correct answers on the sSDMT than on the SDMT, respectively.

XSL•FO
RenderX

Figure 3. Distribution of the differences between the number of correct answers on the sSDMT and the SDMT. The dashed line represents a normal distribution. sSDMT: smartphone variant of Symbol Digit Modalities Test; SDMT: Symbol Digit Modalities Test.

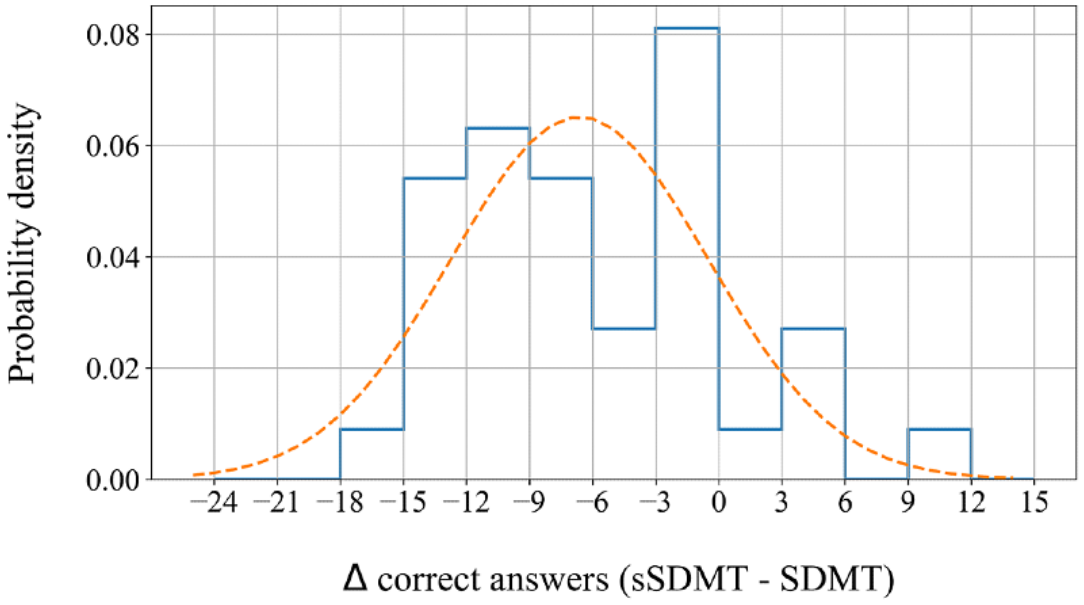
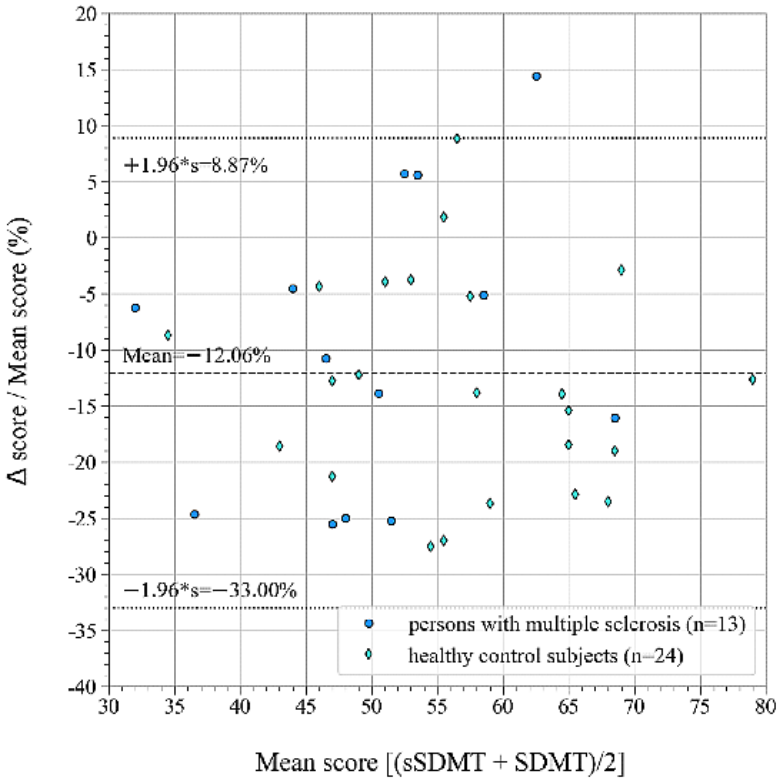


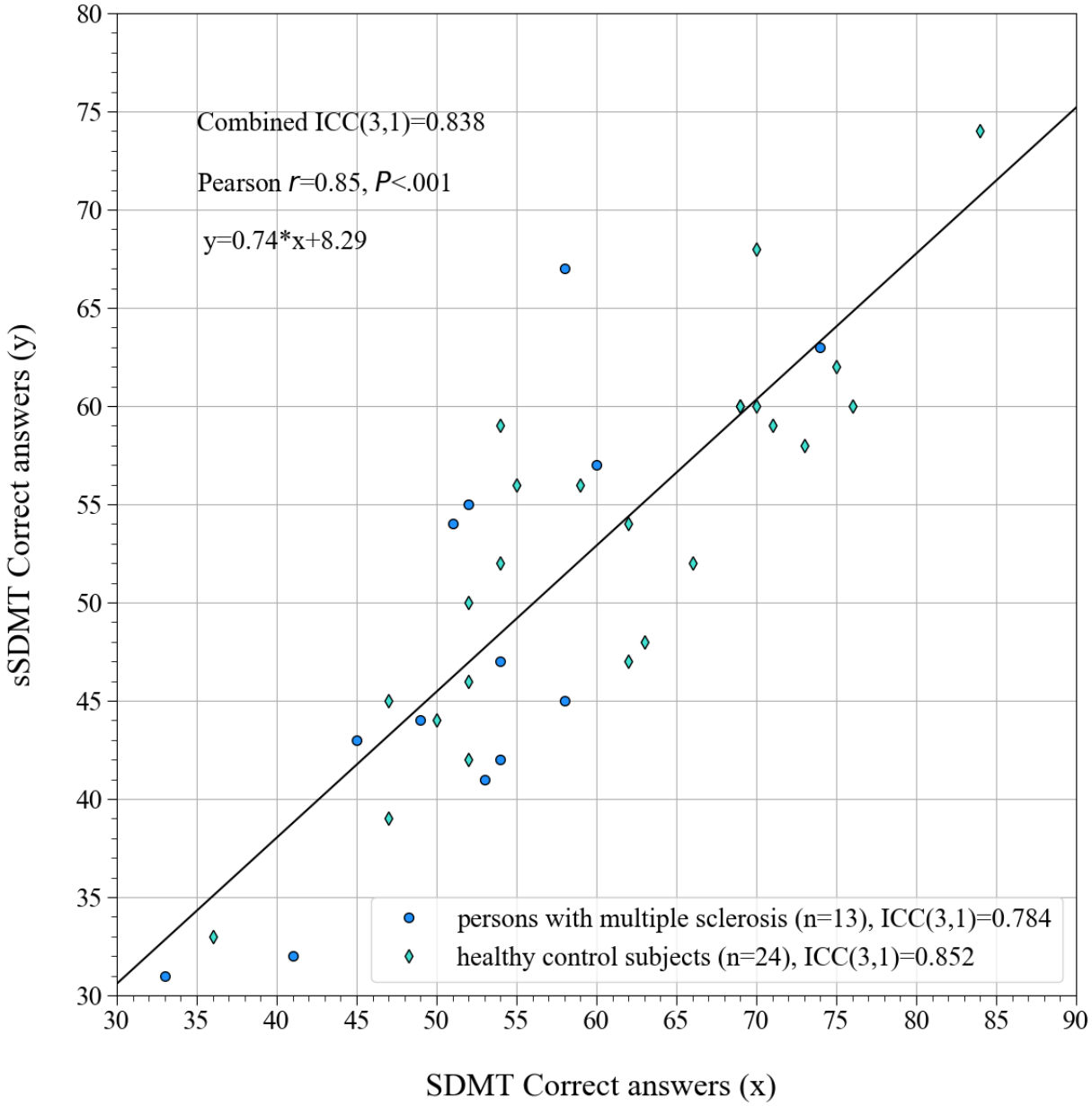
Figure 4. Bland-Altman plot of differences between the number of correct answers on the sSDMT and SDMT, expressed as percentages of the mean value (Δ/mean) versus the mean of the 2 measurements (raw data in Multimedia Appendix 1). The dashed line shows the mean percentage difference, and the dotted lines show the 95% confidence interval. sSDMT: smartphone variant of Symbol Digit Modalities Test; SDMT: Symbol Digit Modalities Test.



When investigating the agreement between the SDMT and the sSDMT using a *t* test, as was planned in the study protocol, a very small *P* value was obtained. A dependent *t* test for paired samples yielded a test statistic of 6.49 ($P<.001$). Since this was below the significance level of .05, we rejected the null hypothesis of identical average scores. In Figure 5, the number of correct answers on the sSDMT is plotted against the number of correct answers on the SDMT. ICCs(3,1) were determined for persons with MS and HC subjects separately and for the combined dataset, which are shown in the top left corner of Figure 5. The Pearson correlation coefficient and the formula corresponding to the linear fit through the data are also shown in the top left corner of Figure 5.

XSL•FO
RenderX

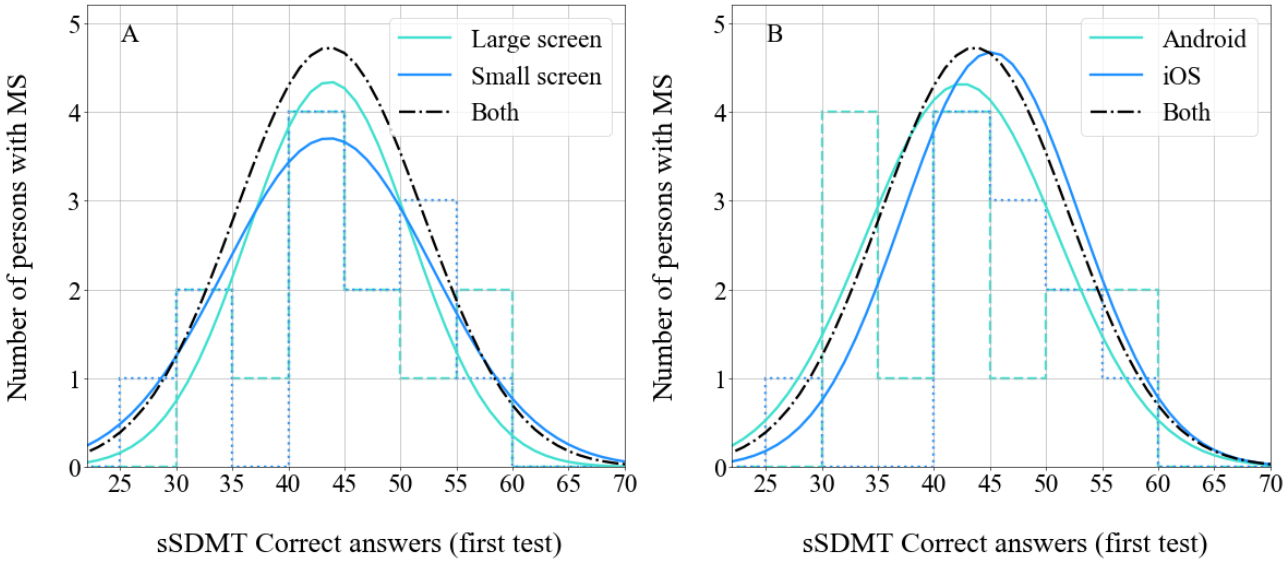
Figure 5. Scatter plot showing the ICCs(3,1) and the correlation (Pearson r, upper left corner) between the number of correct answers on the SDMT (horizontal axis, "x") and the sSDMT (vertical axis, "y"). A linear fit through the data points is visualized as a black solid line, for which the formula is also given in the upper left corner. ICC: intraclass correlation coefficient; sSDMT: smartphone variant of Symbol Digit Modalities Test; SDMT: Symbol Digit Modalities Test.



For the persons with MS, we checked if the number of correct answers on the first sSDMT was dependent on the screen size of their smartphones or their operating systems. The resulting distributions are respectively shown in the left and right panels of Figure 6. We took the total number of pixels as an estimate for screen size. The screens with less than 2 million pixels were considered small, while those with 2 million pixels or more were considered large. The number of correct answers was normally distributed for all 4 categories as calculated by the Shapiro-Wilk tests (P=.54, P=.48, P=.16, and P=.30 for a small screen, large screen, iPhone operating system, or Android operating system, respectively). Since we investigated the score on the first sSDMT, we also included the test results of the persons with MS who dropped out of the study. We cannot confirm if this dropout did not experience difficulties with perceptual motor abilities, but inclusion of this individual does not significantly affect the results.

XSL•FO
RenderX

Figure 6. A: Distribution of the number of correct answers on the first sSDMT of 13 persons with multiple sclerosis who used a smartphone with a large screen size (dashed line) and 13 persons with multiple sclerosis who used a smartphone with a small screen size (dotted line). The thick solid lines represent Gaussian fits to the distributions. B: Distribution of the number of correct answers on the first sSDMT of 11 persons with multiple sclerosis who used a smartphone with an iPhone operating system (iOS) (dotted line) and 15 persons with multiple sclerosis who used a smartphone with an Android operating system (dashed line). The thick solid lines represent Gaussian fits to the distributions. In both panels, the dot-dashed line represents a Gaussian fit to the full distribution (both categories). MS: multiple sclerosis; sSDMT: smartphone variant of Symbol Digit Modalities Test.



The 2 categories of screen size as well as the 2 categories of operating systems had the same underlying distribution as confirmed using 2-sample Kolmogorov-Smirnov tests (Kolmogorov-Smirnov statistic for screen size=0.23, *P*=.83 and Kolmogorov-Smirnov statistic for operating system=0.28, *P*=.60). These distributions are plotted with a dot-dashed line in both panels of Figure 6. Similarly, independent 2-sample *t* tests between the 2 categories of screen size and between the 2 categories of operating systems confirmed that neither the screen size nor the operating system had an effect on the test score (*P*=.98 and *P*=.46, respectively).

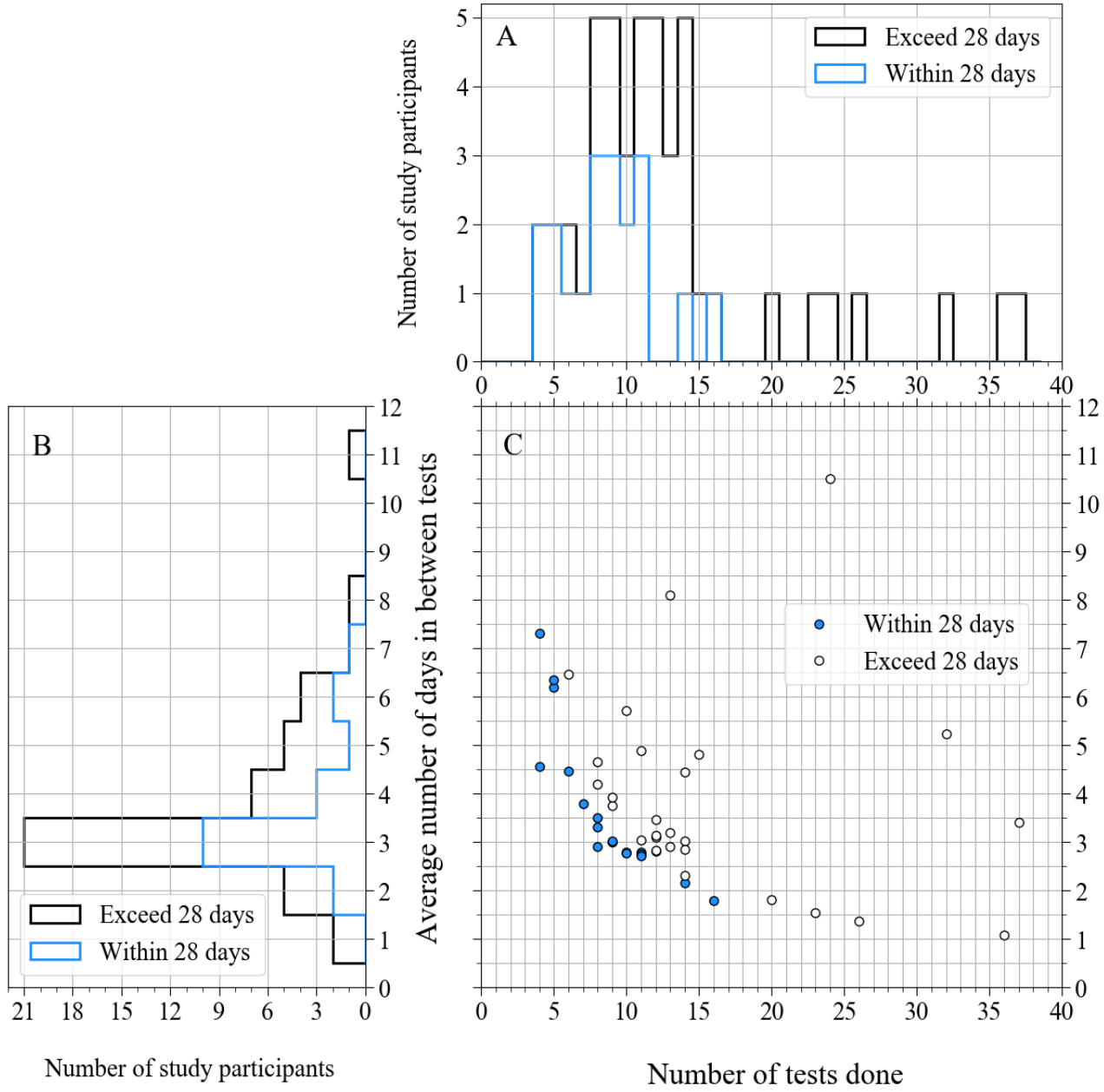
Test-Retest Reliability

One of the dropouts did 7 home assessments with the sSDMT. We consider this a sufficiently large number to calculate test-retest reliability; therefore, we included this dropout for the analyses.

Figure 7A shows the distribution of the total number of home assessments done for persons with MS and HC matched group. Even though 57% (27/47) of the persons with MS and HC subjects in the matched group did more than 10 tests in total, we included only their first 10 home assessments for the test-retest reliability calculations, as planned in the study protocol. Figure 7B shows the distribution of the average number of days the persons with MS and HC subjects in the matched group left between tests, which was averaged over all tests of a study participant. This distribution peaks at 3 days, but it is different from 3 days for 55% (26/47) of the study participants. In these histograms, we also show the distributions of 60% (28/47) of the persons with MS and HC subjects in the matched group who continued to do tests after 4 weeks. The scatter plot in the lower right panel of Figure 7 (Figure 7C) shows the relation between the number of tests done and the average number of days in between tests. Study participants who did all home assessments within 28 days are plotted with filled circles, and the others with open circles.

XSL•FO
RenderX

Figure 7. A: Histogram showing the distribution of total number of tests done. B: Histogram showing the distribution of number of days in between tests, averaged over all tests of a study participant. C: Scatter plot that shows the relation between the number of tests done (horizontal axis) and the average number of days in between tests, averaged over all tests of a study participant (vertical axis) for all study participants. Filled circles correspond to participants that finished the study within 28 days, open circles correspond to those for which the study duration exceeded 28 days. The blue lines in panels A and B correspond to these filled circles, and the black lines in panels A and B correspond to the open circles.



The ICCs(A,1) (with their 95% CI), Cronbach α, and Cohen *d* values that were derived from 9 test-retests of the sSDMT for the persons with MS and the HC matched group are listed in Table 2, and the values of those derived from the 2 test-retests performed by the HC normative group are shown in Table 3. The raw data from persons with MS and the HC matched group can be found in the Multimedia Appendix 3 and that of the HC normative group in Multimedia Appendix 4. The mean values of the ICC(A,1) for persons with MS, HC matched group and the HC normative group were 0.874, 0.857, and 0.867, respectively.

XSL•FO
RenderX

Table 2. Test-retest reliability scores ICC(A,1), Cronbach α, and Cohen d of the sSDMT for persons with multiple sclerosis and the healthy control matched group. The numbers in the parentheses after the ICCs(A,1) indicate lower and upper boundaries of the 95% confidence interval, respectively.

Test-retest	ICC[a](A,1) persons with MS[b]	ICC(A,1), HC[c] subjects	Cronbach α of persons with MS	Cronbach α of HC subjects	Cohen *d* of persons with MS	Cohen *d* of HC subjects
1	0.747 (0.297-0.902)	0.735 (0.401-0.888)	.903	.875	0.477	0.384
2	0.937 (0.863-0.972)	0.859 (0.688-0.940)	.967	.926	0.062	0.153
3	0.899 (0.788-0.953)	0.885 (0.740-0.952)	.945	.943	0.030	0.160
4	0.849 (0.690-0.930)	0.797 (0.554-0.916)	.922	.889	0.160	0.169
5	0.805 (0.611-0.909)	0.820 (0.567-0.931)	.893	.895	0.138	0.000
6	0.880 (0.752-0.944)	0.913 (0.764-0.970)	.934	.952	0.041	0.072
7	0.939 (0.867-0.972)	0.837 (0.586-0.942)	.968	.908	0.059	0.095
8	0.915 (0.814-0.963)	0.961 (0.871-0.989)	.956	.978	0.094	0.000
9	0.898 (0.769-0.957)	0.904 (0.644-0.977)	.946	.946	0.107	0.098
mean	0.874 (0.717-0.945)	0.857 (0.646-0.945)	.937	.924	0.130	0.126

[a]ICC: intraclass correlation coefficient.

[b]MS: multiple sclerosis.

[b]HC: healthy control.

Table 3. Test-retest reliability scores ICC(A,1), Cronbach α, and Cohen d of the sSDMT for the healthy control normative group. The numbers in the parentheses after the ICCs(A,1) indicate the lower and upper boundaries of the 95% CI, respectively.

Test-retest	ICC[a](A,1)	Cronbach α	Cohen *d*
sSDMT[b] test-retest 1	0.877 (0.772-0.931)	.943	0.201
sSDMT test-retest 2	0.857 (0.769-0.913)	.925	0.116
mean	0.867 (0.771-0.922)	.934	0.159

[a]ICC: intraclass correlation coefficient.

[b]sSDMT: smartphone variant of Symbol Digit Modalities Test.

As was planned, we quantified the practice effect using the effect size Cohen *d*, with a value below 0.20 being trivial (Cohen, 1988). The Cohen *d* values in Table 3 show that there is a significant practice effect after the first sSDMT for all study participants. When comparing the first and second sSDMT scores of the persons with MS and HC subjects in the matched group, we found that the mean increase in the number of correct answers in the second test was 3.38 points or 6.87% compared to the first test. Toward the end of the study, the practice effect on the sSDMT completely disappeared at group level. For example, there was, on average, no difference in the scores between the seventh and the eighth sSDMT.

Study participants performed better on the SDMT on the last day of the study, on average, compared to their score on the first day of the study. The average increase between these 2 tests was 5.57%, with a standard deviation of 11.40%, as can be calculated for the 14 study participants that did the SDMT both at the first and at the last day of the study, and even differences as high as 34.92% are reached. Comparing the first and last sSDMT for these 14 study participants yielded, on average, an increase in the number of correct answers of 11.77%, with a standard deviation of 7.84%. The Cohen *d* values quantifying the practice effect on the SDMT and the sSDMT between the first and the last day of the study were found to be 0.365 and 0.908, respectively. The practice effect for the HC normative group was also investigated. The Cohen *d* values corresponding to the first and the second retest were found to be 0.201 and 0.116 respectively (see Table 3). Although a mean increase in the test score of 3.32% at group level was found between the first and the second test with a standard deviation of 7.29%, the practice effect could be considered trivial for this group on the basis of the Cohen *d* values.

Interview Results

The 7 participants with MS who were interviewed about their experiences with the smartphone app and Fitbit activity tracker in general expressed a positive attitude regarding sSDMT. They liked doing the test, often describing it as a “game” or a “puzzle.” Moreover, 5 respondents imagined that this test could provide valuable information to health care providers about the health status of their patients, as this is relevant information about how their patients are doing. Despite this general positive attitude, respondents also expressed some remarks regarding the sSDMT. Two respondents noticed that sometimes the digits they pressed did not seem to respond. They were unsure whether this was caused by the app or their phone. Furthermore, several respondents discussed that in order to use a smartphone cognition test in their daily life, such a test should become personalized. Three respondents mentioned that every person with MS has different difficulties regarding cognition and felt

XSL•FO
RenderX

that these differences could not be captured by a single test. Rather, they envisioned a smartphone app with multiple cognition tests from which persons with MS can choose from, depending on their personal cognitive issues. Another point of personalization concerned the test frequency. Four respondents expected that during the stable periods of their MS, they would feel less need to perform the sSDMT than during periods of relapse. Therefore, they desired flexibility in the sSDMT test frequency when they would use the smartphone app in their daily life.

Discussion

Principal Findings

We found that the sSDMT can distinguish between persons with MS and HC subjects at the group level. The test scores on the sSDMT are, on average, 12.06% lower than that on the paper-and-pencil SDMT. Although there is no exact agreement between the sSDMT and the SDMT, the 2 tests are strongly correlated, with ICC(3,1) values of 0.784 and 0.852 for persons with MS and HC subjects, respectively. The sSDMT shows very good test-retest reliability, with average ICCs of 0.874, 0.857, and 0.867 for persons with MS, the HC matched group, and the HC normative group, respectively. The practice effect was significant between the first and the second test of the persons with MS and the HC matched group and trivial for all other test-retests. A positive attitude toward the sSDMT was found during the interviews with study participants with MS. Importantly, interview respondents expressed the desire to adapt smartphone cognition tests according to the individual needs of persons with MS, with regard to both the type and frequency of testing.

Limitations

The SDMT is a relevant tool both for screening and monitoring not only for MS but for many other clinical diseases such as Huntington disease, Parkinson disease, and stroke [36]. This study limits itself to persons with MS, and the validity of the sSDMT will still have to be demonstrated for applications outside MS. The validation and study population are limited to Dutch people and the Dutch language. We expect that the construct validity and test-retest reliability would have been equally good if this study had been done with persons with MS from various European countries, given the validity and reliability that were found in studies in populations from various countries with other digital assessment devices for CPS. The MS sherpa app now also supports an English version of the sSDMT, and a German version is planned. However, we recommend a separate validation for each country where the sSDMT will be used, for example, as done for BICAMS [37,38] because the SDMT is known to be affected by culture [39]. It was not expected that the number of correct answers on the sSDMT and the SDMT would exactly match, because the sSDMT was not designed to follow the SDMT as close as possible, and systematic differences between the tests exist. Therefore, the SDMT and sSDMT may not be used interchangeably for monitoring. The lower scores on the sSDMT compared to the SDMT are most likely caused by the fact that all items of the sSDMT are selected from the full key, whereas in the SDMT, the first 26 items are selected from the first 6 symbols in the key only.

A 10% difference between 2 SDMT scores of a patient is found to be an indication of a clinically meaningful change in the patients' CPS [8]. This study was not designed to derive the boundary for meaningful change on the sSDMT score. Even though there is a strong correlation between the SDMT and the sSDMT, we cannot claim that a 10% difference on 2 sSDMT scores also signifies a clinically meaningful change. Moreover, we found more than 10% increase in sSDMT performance between the first and the last day of the study, on average, even though the disability of the persons with MS was not expected to change during the 4 weeks of follow-up. An explanation for this increase and for the Cohen *d* value quantifying the practice effect on the sSDMT between the first and the last day of the study of 0.908 is that 60% (28/47) of the study participants had been learning on the sSDMT every 3 days or more often, leading to accumulations of learning effects.

One might wonder how often an sSDMT could be scheduled, to keep its value as an objective instrument for measuring CPS. Benedict et al [19] showed that the SDMT has good-to-excellent reproducibility over repeated testing when used in monthly successive examinations. The HC subjects in the normative group who had a time of 1 week between tests instead of 3 days had less of a practice effect than the persons with MS and HC subjects in the matched group. Therefore, when trying to minimize the practice effect, the scheduling frequency of self-assessments should be considered for clinical practice. Our current suggestion would be to schedule the sSDMT once a month to avoid inducing large practice effects when supporting monitoring for clinical practice.

This study was not set up to cross-device validate the sSDMT, but we have tried to give some sense of the validity over different screen sizes and the main software platforms, because we believed that these 2 factors could affect the results. In our study, we found an approximately equal number of study participants using the Android or iPhone operating system platform, which gave us the opportunity to significantly show that there is no structural bias in using any of the platforms. A more structural approach to cross-device validation and more detailed analysis of the performance over different devices could help to detect validity issues in different devices.

In our study, we did not explicitly give the instructions to self-monitor under the same conditions each time as much as possible (preferably in a quiet room). We expect that such instructions would improve the test-retest reliability. The fact that the ICCs and Cronbach α values of the first test-retest of the persons with MS and the HC matched group were lower than all following 8 can likely be explained by the fact that there is, in general, a large practice effect between the first and the second time a cognitive test is done, especially with only 3 days in between the tests. This is also reflected in the Cohen *d* values, which are the highest for the first test-retest. We explain the triviality of the Cohen *d* values corresponding to the first and the second retest of the HC normative group with the larger time in between successive tests (approximately 1 week).

XSL•FO
RenderX

The fact that 60% (28/47) of the persons with MS and HC subjects in the matched group continued to do home assessments after 4 weeks might be because the end day of the study could often not be planned 4 weeks after the first day. It could also be a sign of high acceptance of the sSDMT, as expressed by the interviewed study participants. These results, together with the qualitative findings that participants liked doing the sSDMT, often describing it as a "game" or a "puzzle," support that the sSDMT was viewed as "game-like" and confirms that the sSDMT is well-designed for the user.

Since persons with MS were asked to write down their answers on the paper-and-pencil SDMT themselves, persons with MS who had limited hand dexterity might perform lower on the paper-and-pencil SDMT than those with the same CPS without hand dexterity problems. There was no Nine Hole Peg Test scheduled in this study to investigate this possible bias. During the sSDMT, persons with MS are asked to tap the correct answer. We expect hand dexterity problems to be less of a problem for tapping on a smartphone than for writing down a number on paper. An alternative solution for this bias would be an oral sSDMT, which should then be validated against an oral SDMT. However, persons with MS might not find this a natural way of self-monitoring.

The number of study participants for which the test-retest reliability of the paper-and-pencil SDMT could be studied was relatively low, because 2 different versions of the paper-and-pencil SDMT were used. Although the reliability of the paper-and-pencil SDMT is well known in literature, it would have been interesting to study the practice effect of frequently performing our sSDMT on the paper-and-pencil SDMT in more detail.

Finally, an important limitation is the omission of the 10 mandatory practice items on the paper-and-pencil SDMT. This could have introduced a bias toward a better correlation with the sSDMT, because in the digital variant, practicing is optional, and therefore it might affect our construct validity. The construct validity should therefore ideally be reevaluated in a follow-up study in which study participants do the standard 10 practice items on the paper-and-pencil SDMT.

As mentioned in the Introduction, originally our rationale for an optional practice assessment was that practicing should not be mandatory for persons with MS that do the sSDMT repeatedly from a usability perspective. Practicing was optional but could be done unlimitedly by pressing the corresponding button on the instruction screen. However, our results have let us to reconsider this perspective and we now believe that the sSDMT nor a practice session should be accessible unlimitedly, because of the accumulative nature of the practice effect. However, when the sSDMT is scheduled only once a month, users might want to practice before each digital assessment. Therefore, we currently suggest 10 mandatory practice items before each digital assessment, which is also more similar to the paper-and-pencil SDMT, and therefore might improve the construct validity.

Comparison With Prior Work

A fair number of computerized neuropsychological assessment devices for monitoring cognitive impairment in MS has been developed in recent years. A recent systematic review of the literature on test batteries and single tests with good evidence for reliability and validity yielded 44 CPS tests, of which all computerized tests based on SDMT correlated with the conventional SDMT, with correlation coefficients ranging from 0.75 to 0.88 [40]. Our correlation coefficient of 0.85 is in line with these findings. Most CPS tests also showed acceptable reliability, for example, ICC values of 0.88 and 0.97 were reported for the Processing Speed Test (PST) [41] and the computerized version of SDMT [18], respectively. The 95% CI on the mean ICC that we find ranges from 0.717 to 0.945. This is thus also similar to that reported in prior work.

To our knowledge, there are 2 other smartphone-based SDMT apps for MS that should be mentioned as alternative solutions to monitor CPS. One is MSCopilot [24], which contains a digital SDMT variant and 3 other digital variants of tests in the MSFC. However, results were only reported for the combined digital MSFC assessment in comparison to MSFC z-scores. Data on the validity and reliability of their digital SDMT variants specifically have never been published to our knowledge. FLOODLIGHT [25] is another smartphone monitoring app for persons with MS, which was developed by Roche. At a poster presented at the European Committee for Treatment and Research in Multiple Sclerosis in 2018, Montalban et al [25] reported a Spearman's correlation of 0.615 between the FLOODLIGHT smartphone-based SDMT and the conventional in-clinic outcome measure (oral SDMT). This is considerably lower than our correlation coefficients (ie, Pearson r=0.85).

Well-validated digital assessment devices for CPS screening include the PST and the recently introduced Multiple Screener tool, which contains a digital SDMT for which an ICC of 0.79 between digital and paper-and-pencil-based assessment was reported [42]. Rudick et al [23] reported a Pearson correlation coefficient of 0.80 between the PST and analogous technician tests and concordance correlation coefficients to quantify test-retest reproducibility of 0.853 for the technician and 0.867 for persons with MS. Our results are similar to theirs. However, both the PST and Multiple Screener are only available for iPads and are intended to be used for monitoring CPS in a more controlled setting (in the clinic), whereas the MS sherpa app is intended to be used for home monitoring. The importance of personalization and customization of smartphone apps for persons with MS has been noted in other studies [43,44]. It is therefore recommended that the needs and context of the individual with MS are taken into account in the design of apps for persons with MS.

To our knowledge, this is the first study in which a smartphone variant of the SDMT that can be done unsupervised was cross-platform validated. We obtained valuable novel insights into frequent home monitoring with the SDMT such as the observation that the practice effect was only nontrivial between the first and second sSDMT (with 10 assessments scheduled and approximately 3 days in between each assessment) but also about the cumulative practice effects that are involved.

XSL•FO
RenderX

Furthermore, we believe the qualitative insights obtained from patient interviews on the needs and wishes of smartphone-based home monitoring using sSDMT can inspire developers, caregivers, and researchers for future developments.

Conclusion

This study shows the construct validity of the sSDMT since the ICCs(3,1) between the SDMT and the sSDMT for persons with MS and HC subjects were 0.78 and 0.85, respectively. The sSDMT does have very good test-retest reliability because only the first test-retest of the persons with MS and the HC matched group yielded an ICC(A,1) smaller than 0.80. All the other ICCs were higher than 0.80, both for persons with MS and for HC subjects. We conclude that the sSDMT has the potential to be used as a tool to monitor CPS in persons with MS, both in clinical studies and in clinical practice.

Acknowledgments

The study received financial support from the European Regional Development Fund (EFRO OP Oost Project 00217) and the NMSF, Rotterdam, the Netherlands. We thank the anonymous referees for valuable comments.

Conflicts of Interest

PvO is employed by Orikami Digital Health Products. BdT is a founder and owner of Orikami Digital Health Products. PJJ is advisor to Orikami Digital Health Products and has received honoraria from Bayer Netherlands for consultancy activities.

Multimedia Appendix 1
Number of correct answers on the SDMT and the sSDMT and the agreement between the two methods, for the 37 validation assessments that were done with the SDMT.
[DOCX File , 18 KB - mhealth_v8i10e18160_app1.docx]

Multimedia Appendix 2
Number of correct answers on the vSDMT and the sSDMT and the agreement between the two methods, for the 54 validation assessments that were done with the vSDMT.
[DOCX File , 21 KB - mhealth_v8i10e18160_app2.docx]

Multimedia Appendix 3
Number of correct answers on the sSDMT on 10 occasions (labeled 1 to 10) for the persons with multiple sclerosis and the healthy controls in the matched group.
[DOCX File , 22 KB - mhealth_v8i10e18160_app3.docx]

Multimedia Appendix 4
Number of correct answers on the sSDMT on three occasions (labeled 1, 2 and 3) for the healthy control normative group.
[DOCX File , 17 KB - mhealth_v8i10e18160_app4.docx]

References

1. Amato MP, Zipoli V, Portaccio E. Multiple sclerosis-related cognitive changes: a review of cross-sectional and longitudinal studies. J Neurol Sci 2006 Jun 15;245(1-2):41-46. [doi: 10.1016/j.jns.2005.08.019] [Medline: 16643953]
2. Foley, Benedict RHB, Gromisch ES, Deluca J. The Need for Screening, Assessment, and Treatment for Cognitive Dysfunction in Multiple Sclerosis: Results of a Multidisciplinary CMSC Consensus Conference, September 24, 2010. Int J MS Care 2012;14(2):58-64 [FREE Full text] [doi: 10.7224/1537-2073-14.2.58] [Medline: 24453735]
3. Manca R, Sharrack B, Paling D, Wilkinson ID, Venneri A. Brain connectivity and cognitive processing speed in multiple sclerosis: A systematic review. J Neurol Sci 2018 May 15;388:115-127. [doi: 10.1016/j.jns.2018.03.003] [Medline: 29627004]
4. Rocca MA, Amato MP, De Stefano N, Enzinger C, Geurts JJ, Penner I, et al. Clinical and imaging assessment of cognitive dysfunction in multiple sclerosis. The Lancet Neurology 2015 Mar;14(3):302-317. [doi: 10.1016/s1474-4422(14)70250-9]
5. van Geest Q, Douw L, van 't Klooster S, Leurs C, Genova H, Wylie G, et al. Information processing speed in multiple sclerosis: Relevance of default mode network dynamics. Neuroimage Clin 2018;19:507-515 [FREE Full text] [doi: 10.1016/j.nicl.2018.05.015] [Medline: 29984159]
6. Gronwall DMA. Paced auditory serial-addition task: a measure of recovery from concussion. Percept Mot Skills 1977 Apr;44(2):367-373. [doi: 10.2466/pms.1977.44.2.367] [Medline: 866038]
7. Smith A. Symbol digit modalities test. Los Angeles, CA: Western Psychological Services; 1982.
8. Benedict RH, DeLuca J, Phillips G, LaRocca N, Hudson LD, Rudick R, Multiple Sclerosis Outcome Assessments Consortium. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. Mult Scler 2017 Apr;23(5):721-733 [FREE Full text] [doi: 10.1177/1352458517690821] [Medline: 28206827]

XSL•FO
RenderX

9. Learmonth YC, Motl RW, Sandroff BM, Pula JH, Cadavid D. Validation of patient determined disease steps (PDDS) scale scores in persons with multiple sclerosis. BMC Neurol 2013 Apr 25;13:37 [FREE Full text] [doi: 10.1186/1471-2377-13-37] [Medline: 23617555]
10. Rao SM, Leo GJ, Haughton VM, St Aubin-Faubert P, Bernardin L. Correlation of magnetic resonance imaging with neuropsychological testing in multiple sclerosis. Neurology 1989 Feb;39(2 Pt 1):161-166. [doi: 10.1212/wnl.39.2.161] [Medline: 2915783]
11. Ruano L, Portaccio E, Goretti B, Niccolai C, Severo M, Patti F, et al. Age and disability drive cognitive impairment in multiple sclerosis across disease subtypes. Mult Scler 2017 Aug;23(9):1258-1267. [doi: 10.1177/1352458516674367] [Medline: 27738090]
12. Dusankova JB, Kalincik T, Havrdova E, Benedict RHB. Cross cultural validation of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS) and the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). Clin Neuropsychol 2012;26(7):1186-1200. [doi: 10.1080/13854046.2012.725101] [Medline: 23034066]
13. Sonder JM, Burggraaff J, Knol DL, Polman CH, Uitdehaag BM. Comparing long-term results of PASAT and SDMT scores in relation to neuropsychological testing in multiple sclerosis. Mult Scler 2014 Apr;20(4):481-488. [doi: 10.1177/1352458513501570] [Medline: 24019305]
14. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). Neurology 1983 Nov 01;33(11):1444-1444. [doi: 10.1212/wnl.33.11.1444]
15. López-Góngora M, Querol L, Escartín A. A one-year follow-up study of the Symbol Digit Modalities Test (SDMT) and the Paced Auditory Serial Addition Test (PASAT) in relapsing-remitting multiple sclerosis: an appraisal of comparative longitudinal sensitivity. BMC Neurol 2015 Mar 22;15:40 [FREE Full text] [doi: 10.1186/s12883-015-0296-2] [Medline: 25886168]
16. Silva AM, Moreira M, Samões R, Bennencourt A, Martins J, Duarte S, et al. Clinical validity of EDSS and SDMT in the detection of MS disease activity. 2017 Oct 25 Presented at: ECTRIMS Online Library, conference abstract P1404, 199425; 25-28 October 2017; Paris, France. [doi: 10.26226/morressier.59a3e8b4d462b8028d8945f6]
17. Rao S, Martin AL, Huelin R, Wissinger E, Khankhel Z, Kim E, et al. Correlations between MRI and Information Processing Speed in MS: A Meta-Analysis. Mult Scler Int 2014;2014:975803 [FREE Full text] [doi: 10.1155/2014/975803] [Medline: 24795824]
18. Akbar N, Honarmand K, Kou N, Feinstein A. Validity of a computerized version of the symbol digit modalities test in multiple sclerosis. J Neurol 2011 Mar 6;258(3):373-379. [doi: 10.1007/s00415-010-5760-8] [Medline: 20924594]
19. Benedict R, Duquin J, Jurgensen S, Rudick R, Feitcher J, Munschauer F, et al. Repeated assessment of neuropsychological deficits in multiple sclerosis using the Symbol Digit Modalities Test and the MS Neuropsychological Screening Questionnaire. Mult Scler 2008 Aug;14(7):940-946. [doi: 10.1177/1352458508090923] [Medline: 18573822]
20. Benedict RH, Smerbeck A, Parikh R, Rodgers J, Cadavid D, Erlanger D. Reliability and equivalence of alternate forms for the Symbol Digit Modalities Test: implications for multiple sclerosis clinical trials. Mult Scler 2012 Sep;18(9):1320-1325. [doi: 10.1177/1352458511435717] [Medline: 22277740]
21. Rao SM, Leo GJ, Ellington L, Nauertz T, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis. II. Impact on employment and social functioning. Neurology 1991 May;41(5):692-696. [doi: 10.1212/wnl.41.5.692] [Medline: 1823781]
22. Rao SM, Losinski G, Mourany L, Schindler D, Mamone B, Reece C, et al. Processing speed test: Validation of a self-administered, iPad-based tool for screening cognitive dysfunction in a clinic setting. Mult Scler 2017 Dec;23(14):1929-1937. [doi: 10.1177/1352458516688955] [Medline: 28080262]
23. Rudick RA, Miller D, Bethoux F, Rao SM, Lee J, Stough D, et al. The Multiple Sclerosis Performance Test (MSPT): an iPad-based disability assessment tool. J Vis Exp 2014 Jun 30(88):e51318 [FREE Full text] [doi: 10.3791/51318] [Medline: 25046650]
24. Maillart E, Labauge P, Cohen M, Maarouf A, Gallien P, Vukusic S, et al. Digital Self-Assessment for Patients Living with Multiple Sclerosis: Clinical Results of the Mscopilot Multi-Center Study. 2019 Feb 28 Presented at: the 4th annual forum of the Americas Committee for Treatment and Research in Multiple Sclerosis (ACTRIMS); February 28 - March 2 2019; Dallas, Texas.
25. Montalban X, Mulero P, Midaglia L, Graves J, Hauser SL, Julian L, et al. FLOODLIGHT: Smartphone-based self-monitoring is accepted by patients and provides meaningful, continuous digital outcomes augmenting conventional in-clinic multiple sclerosis measures. 2018 Oct 10 Presented at: 34th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS); October 10-12, 2018; Berlin, Germany. [doi: 10.1358/dof.2018.043.12.2927589]
26. Wendrich K, van Oirschot P, Martens MB, Heerings M, Jongen PJ, Krabbenborg L. Toward Digital Self-monitoring of Multiple Sclerosis: Investigating First Experiences, Needs, and Wishes of People with MS. Int J MS Care 2019;21(6):282-291 [FREE Full text] [doi: 10.7224/1537-2073.2018-083] [Medline: 31889935]
27. van Oirschot P, Heerings M, Wendrich K, Beirao P, van de Haar J, den Teuling B, et al. Self-monitoring of mobility, signal processing speed, physical activity and sleep by persons with multiple sclerosis. 2018 Oct 10 Presented at: the 34thCongress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS); 10-12 October 2018; Berlin, Germany.

XSL•FO
RenderX

28. MS sherpa personalized healthcare. URL: https://www.mssherpa.nl/ [accessed 2020-07-03]
29. Cronbach L. Coefficient alpha and the internal structure of tests. Psychometrika 1951 Sep;16(3):297-334. [doi: 10.1007/BF02310555]
30. Nunnaly JC. Psychometric theory. New York: McGraw-Hill; 1978.
31. Bushnell CD, Johnston DC, Goldstein LB. Retrospective assessment of initial stroke severity: comparison of the NIH Stroke Scale and the Canadian Neurological Scale. Stroke 2001 Mar;32(3):656-660. [doi: 10.1161/01.str.32.3.656] [Medline: 11239183]
32. Chien C, Hu M, Tang P, Sheu C, Hsieh C. A comparison of psychometric properties of the smart balance master system and the postural assessment scale for stroke in people who have had mild stroke. Arch Phys Med Rehabil 2007 Mar;88(3):374-380. [doi: 10.1016/j.apmr.2006.11.019] [Medline: 17321832]
33. Cohen J. Statistical power analysis for the behavioral sciences, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
34. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychological Methods 1996;1(1):30-46. [doi: 10.1037/1082-989x.1.1.30]
35. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012]
36. Kiely KM, Butterworth P, Watson N, Wooden M. The Symbol Digit Modalities Test: Normative data from a large nationally representative sample of Australians. Arch Clin Neuropsychol 2014 Dec;29(8):767-775. [doi: 10.1093/arclin/acu055] [Medline: 25352087]
37. Benedict RHB, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson S, et al. Brief International Cognitive Assessment for MS (BICAMS): international standards for validation. BMC Neurol 2012 Jul 16;12:55 [FREE Full text] [doi: 10.1186/1471-2377-12-55] [Medline: 22799620]
38. Filser M, Schreiber H, Pöttgen J, Ullrich S, Lang M, Penner IK. The Brief International Cognitive Assessment in Multiple Sclerosis (BICAMS): results from the German validation study. J Neurol 2018 Nov;265(11):2587-2593. [doi: 10.1007/s00415-018-9034-1] [Medline: 30171410]
39. Cores EV, Vanotti S, Eizaguirre B, Fiorentini L, Garcea O, Benedict RHB, et al. The Effect of Culture on Two Information-Processing Speed Tests. Appl Neuropsychol Adult 2015;22(4):241-245. [doi: 10.1080/23279095.2014.910214] [Medline: 25372116]
40. Wojcik CM, Beier M, Costello K, DeLuca J, Feinstein A, Goverover Y, National MS Society Cognition Work Team. Computerized neuropsychological assessment devices in multiple sclerosis: A systematic review. Mult Scler 2019 Dec;25(14):1848-1869 [FREE Full text] [doi: 10.1177/1352458519879094] [Medline: 31637963]
41. Rao SM, Losinski G, Mourany L, Schindler D, Mamone B, Reece C, et al. Processing speed test: Validation of a self-administered, iPad-based tool for screening cognitive dysfunction in a clinic setting. Mult Scler 2017 Dec;23(14):1929-1937. [doi: 10.1177/1352458516688955] [Medline: 28080262]
42. van Dongen L, Westerik B, van der Hiele K, Visser LH, Schoonheim MM, Douw L, et al. Introducing Multiple Screener: An unsupervised digital screening tool for cognitive deficits in MS. Mult Scler Relat Disord 2020 Feb;38:101479 [FREE Full text] [doi: 10.1016/j.msard.2019.101479] [Medline: 31760365]
43. Giunti G, Kool J, Rivera Romero O, Dorronzoro Zubiete E. Exploring the Specific Needs of Persons with Multiple Sclerosis for mHealth Solutions for Physical Activity: Mixed-Methods Study. JMIR Mhealth Uhealth 2018 Feb 09;6(2):e37 [FREE Full text] [doi: 10.2196/mhealth.8996] [Medline: 29426814]
44. Simblett SK, Evans J, Greer B, Curtis H, Matcham F, Radaelli M, RADAR-CNS consortium. Engaging across dimensions of diversity: A cross-national perspective on mHealth tools for managing relapsing remitting and progressive multiple sclerosis. Mult Scler Relat Disord 2019 Jul;32:123-132. [doi: 10.1016/j.msard.2019.04.020] [Medline: 31125754]

Abbreviations

2MWT: 2-minute walking test
BICAMS: Brief International Cognitive Assessment for Multiple Sclerosis
CPS: cognitive processing speed
EDSS: Expanded Disability Status Scale
HC: healthy control
ICC: intraclass correlation coefficient
MS: multiple sclerosis
MSFC: Multiple Sclerosis Functional Composite
NMSF: Nationaal MS Fonds
PASAT: Paced Auditory Serial Addition Test
s2MWT: smartphone 2-minute walking test
SDMT: Symbol Digit Modalities Test
sSDMT: smartphone variant of Symbol Digit Modalities Test

XSL•FO
RenderX

sWBT: smartphone walking balance test
TUG: Timed Up and Go
vSDMT: variant of the Symbol Digit Modalities Test

Edited by G Eysenbach; submitted 08.02.20; peer-reviewed by A Landmark, P Hall; comments to author 12.06.20; revised version received 16.07.20; accepted 26.07.20; published 05.10.20.

Please cite as:
van Oirschot P, Heerings M, Wendrich K, den Teuling B, Martens MB, Jongen PJ
Symbol Digit Modalities Test Variant in a Smartphone App for Persons With Multiple Sclerosis: Validation Study
JMIR Mhealth Uhealth 2020;8(10):e18160
URL: https://mhealth.jmir.org/2020/10/e18160
doi:10.2196/18160
PMID:33016886

©Pim van Oirschot, Marco Heerings, Karine Wendrich, Bram den Teuling, Marijn B Martens, Peter J Jongen. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 05.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Agreement Between Spatiotemporal Gait Parameters Measured by a Markerless Motion Capture System and Two Reference Systems—a Treadmill-Based Photoelectric Cell and High-Speed Video Analyses: Comparative Study

Felipe García-Pinillos[1,2*], PhD; Diego Jaén-Carrillo[3*], PhD; Victor Soto Hermoso[1*], PhD; Pedro Latorre Román[4*], PhD; Pedro Delgado[2*], PhD; Cristian Martinez[2], PhD; Antonio Carton[3*], MSc; Luis Roche Seruendo[3], PhD

[1]University of Granada, Granada, Spain
[2]Universidad de La Frontera, Temuco, Chile
[3]Universidad San Jorge, Zaragoza, Villanueva de Gallego, Spain
[4]Universidad de Jaén, Jaén, Spain
[*]these authors contributed equally

Corresponding Author:
Diego Jaén-Carrillo, PhD
Universidad San Jorge
Autovía Mudéjar
Zaragoza, Villanueva de Gallego
Spain
Phone: 34 678638967
Email: djaen@usj.es

Abstract

Background: Markerless systems to capture body motion require no markers to be attached to the body, thereby improving clinical feasibility and testing time. However, the lack of markers might affect the accuracy of measurements.

Objective: This study aimed to determine the absolute reliability and concurrent validity of the Kinect system with MotionMetrix software for spatiotemporal variables during running at a comfortable velocity, by comparing data between the combination system and two widely used systems—OptoGait and high-speed video analysis at 1000 Hz.

Methods: In total, 25 runners followed a running protocol on a treadmill at a speed of 12 km/h. The Kinect+MotionMetrix combination measured spatiotemporal parameters during running (ie, contact time, flight time, step frequency, and step length), which were compared to those obtained from two reference systems.

Results: Regardless of the system, flight time had the highest coefficients of variation (OptoGait: 16.4%; video analysis: 17.3%; Kinect+MotionMetrix: 23.2%). The rest of the coefficients of variation reported were lower than 8.1%. Correlation analysis showed very high correlations (r>0.8; P<.001) and almost perfect associations (intraclass correlation coefficient>0.81) between systems for all the spatiotemporal parameters except contact time, which had lower values. Bland-Altman plots revealed smaller systematic biases and random errors for step frequency and step length and larger systematic biases and random errors for temporal parameters with the Kinect+MotionMetrix system as compared to OptoGait (difference: contact time +3.0%, flight time −7.9%) and high-speed video analysis at 1000 Hz (difference: contact time +4.2%, flight time −11.3%). Accordingly, heteroscedasticity was found between systems for temporal parameters (r^2>0.1).

Conclusions: The results indicate that the Kinect+MotionMetrix combination slightly overestimates contact time and strongly underestimates flight time as compared to the OptoGait system and high-speed video analysis at 1000 Hz. However, it is a valid tool for measuring step frequency and step length when compared to reference systems. Future studies should determine the reliability of this system for determining temporal parameters.

(JMIR Mhealth Uhealth 2020;8(10):e19498) doi:10.2196/19498

KEYWORDS

OptoGait; runners; sport technology; validity; gait; motion capture; video; feasibility; accuracy; Kinect

XSL•FO
RenderX

Introduction

The use of marker-based motion capture technology has increased significantly in both research and diagnostics. This is evident in its prominent use in the field of biomechanics. However, inherent limitations in data collection can preclude its employment in settings such as patient homes, sports fields, and other public areas where implementing an array of cameras proves problematic. A potential solution that has been suggested is the use of a markerless motion capture system [1,2].

Markerless systems do not require attaching any markers or sensors to the body, which substantially improves clinical feasibility and testing time. However, the lack of markers might affect the accuracy of measurements. Therefore, studies focused on analyzing the validity of these systems under different circumstances are especially important. In this context, a markerless motion capture system (ie, the Kinect) has received much attention from clinicians, sports practitioners, and researchers [1,3-9]. The Kinect sensor was originally designed for using body movement to interact with video games on the Microsoft Xbox platform. The system projects an infrared laser speckle pattern onto the viewing area of the infrared camera. This infrared camera detects the pattern and enables the creation of a 3D map by measuring deformations in the reference speckle pattern.

Previous studies have evaluated the validity of the Kinect sensor for the assessment of gait characteristics [1,3-7]. Among these studies, different pieces of software, including different filters and calibrators, have been examined. Clark et al [4] assessed the validity of the Kinect system with a customized software created in LabVIEW 2009 for examining the spatiotemporal characteristics of gait in 21 healthy individuals. In contrast, Lamine et al [7] compared the validity of the Kinect for gait kinematics analysis with that of the Vicon system, and a Cartesian calibration was performed for both motion capture devices. Similarly, Pfister et al [6] compared the concurrent validity of the Kinect with Brekel Kinect software for sagittal plane gait kinematics analysis with that of the Vicon Nexus. Dolatabadi et al [5] determined the concurrent validity of the Microsoft Kinect for Windows for measuring gait spatiotemporal parameters. Schmitz et al [1] tested the validity of the Kinect system with the KinectFusion software for kinematic data evaluation. All of these aforementioned studies included the use of the MotionMetrix software, which might have implications for the accuracy of measurements.

Since the validity of a gait analysis system is essential to determine whether the results are due to changes in gait pattern or simply systematic measurement errors, this study aimed to evaluate the absolute reliability and concurrent validity of the Kinect system with MotionMetrix software for measuring spatiotemporal variables during running at comfortable velocity (ie, 12 km/h) by comparing the data with two widely used systems—the OptoGait system and high-speed video analysis at 1000 Hz. Based on our previous experience with the Kinect+MotionMetrix combination system, we hypothesized spatiotemporal parameters to be similar to those reported by the OptoGait and high-speed video analysis systems.

Methods

With the introduction of new systems, establishment of their reliability and validity is essential before practical use. In this study, the MotionMetrix system was compared to both high-speed video analysis system (1000 Hz) and the OptoGait system for measuring spatiotemporal parameters during a running protocol followed at a comfortable velocity.

Participants

A group of 25 amateur endurance runners (male: n=22, 88%; female: n=3, 12%; age: mean 24 years, SD 6 years; height: mean 1.75 m, SD 0.07 m; body mass: mean 71 kg, SD 7.4 kg) voluntarily participated in this study. All participants met the following inclusion criteria: (1) age, ≥18 years, (2) ability to run 10 km in less than 50 min, and (3) absence of any injury (points 2 and 3 are valid for the 6 months before data collection). After receiving detailed information on the objectives and procedures of the study, each participant signed an informed consent form in order to participate, which complied with the ethical standards of the World Medical Association's Declaration of Helsinki (2013). The participants were informed that they were free to leave the study at any time. The study was approved by the Ethics Committee of the University of La Frontera (Universidad de La Frontera, Temuco, Chile; Ref: 030_019).

Procedures

Participants were individually tested on one specific day. Prior to all testing, participants refrained from vigorous physical activity for at least 48 hours, and all tests were performed at least 3 hours after a meal. Tests were performed with the participants' usual training shoes to measure their typical performance.

Participants performed a running protocol on a motorized treadmill (Woodway Pro XL). The initial speed was set at 8 km/h, and the speed increased by 1 km/h each minute until a speed of 12 km/h was reached. Thereafter, in order to control the influence of running velocity on spatiotemporal parameters [10], the running velocity was fixed. Since previous studies [11,12] on human locomotion have shown that accommodation to running on a treadmill occurs at around 6 to 8 min, an 8-min accommodation program was performed at 12 km/h. Once the accommodation period was reached, recording started. The recording period lasted for 3 min and was performed at the same running velocity. Therefore, the entire running protocol lasted for 15 min. The slope was maintained at 0% over the entire protocol.

Materials and Testing

Overview

Anthropometric data were measured using a precision stadiometer and balance (SECA 222 and 634). Spatiotemporal parameters measured during running included contact time (seconds), defined as time from when the foot contacts the ground to when the toes lift off the ground; flight time (seconds), defined as time from toe-off to initial ground contact of consecutive footfalls (eg, right-left); step length (meters),

XSL•FO
RenderX

defined as length the treadmill belt moves from toe-off to initial ground contact in successive steps from forefoot to forefoot; and step frequency (steps per minute [SPM]), defined as number of ground contact events per minute. We used two different systems to measure these parameters—the OptoGait system versus high-speed video analysis at 1000 Hz. Both systems have the same temporal accuracy (±1 ms). Participants' right legs were analyzed for temporal parameters in order to control potential influencing factors (ie, asymmetry) [13]. Further information about the systems are given below.

Microsoft Kinect

The Microsoft Kinect sensor (version 1.0, Microsoft) was designed to interact with video games through body movement. This sensor can track 3D motions via a depth sensor. It is also capable of locating 20 body joints in a 3D space at 30 Hz. We set two Microsoft Kinect sensors on either side of a treadmill in a specific configuration (170 cm from the center of the treadmill in the forward direction and 190 cm in the perpendicular direction from this point, according to the manufacturer guidelines). These sensors were used with MotionMetrix software (MotionMetrix AB). If both sensors were able to track the same point simultaneously (according to brand information), the Microsoft Kinect sensors could reach 60 Hz. Manufacturer recommendations were taken into account (ie, dynamic calibration provided by the software, tight clothes, no shiny black fabric or reflexes, no moving shoelaces, no moving hair, no sunlight, and no treadmill parts blocking the view of the runner). We recorded 30 s of data with this system (between minutes 1:30 and 2:00 of the recording period for each participant).

OptoGait

The OptoGait system detects any interruptions and therefore measures both contact time and flight time with a precision of 1/1000 seconds. Previous studies have analyzed the validity and reliability of this system during walking [14-18] and running [19]. Two parallel bars were placed on the lengthwise side edges of the treadmill at the same level as the contact surface, and the default filter setting of 0_0 (Gait R.in filter: 0 and Gait R.out filter: 0) was accepted. This setting indicates that contact time begins when more than 0 light-emitting diodes (LEDs) are activated (ie, when at least 1 LED is activated) and finishes when the number of LEDs activated return to 0. This setup has been shown to provide the smallest bias for temporal parameters in racewalking [20]. Spatiotemporal parameters (ie, contact time, flight time, step length, and step frequency) were measured for every step during the 30-second recording interval (between minutes 1:30 and 2:00 of the recording period for each participant).

High-Speed Video Analysis

High-speed video analysis has been shown to be a reliable and valid method for measuring running kinematics [21-24]. In this study, one experienced rater was involved in video analysis. In order to determine the test-retest reliability of the measurements, 10 recording intervals were analyzed on two different days, 24 hours apart, and an almost perfect agreement was found for all the spatiotemporal parameters measured (intraclass correlation coefficients [ICCs]>0.94). In this study, 2D video data were simultaneously collected at 1000 Hz using a high-speed camera (Imaging Source DFK 33UX174, The Imaging Source Europe GmbH). The range of interest was adjusted to achieve 1000 fps (784×144 resolution). The camera was placed perpendicular to the treadmill from a posterior view at 2 m from the center of the treadmill and at a height of 0.80 m. We recorded 30-s videos between minutes 1:30 and 2:00 of the recording period for each participant. Subsequently, videos were analyzed using the open license Kinovea software (version 0.8.27, Kinovea Open Source Project), and spatiotemporal parameters were determined. The contact time and flight time were calculated by identifying both the initial contact and take-off frames and counting the frames in between. Step length and step frequency were calculated as follows:

Step time (seconds) = flight time (seconds) + contact time (seconds) **(1)**

Step frequency (steps/second) = 1/step time (seconds) **(2)**

Step frequency (steps/minute) = 60 × step frequency (steps/second) **(3)**

Step length (meters) = running velocity (meter/minute)/step frequency (steps/min) **(4)**

Statistical Analysis

Descriptive statistics are represented as mean and standard deviation. The normal distribution of the data and the homogeneity of variances were confirmed through the Shapiro-Wilk and Levene tests, respectively. Coefficients of variation (CVs, %) were calculated as a measure of absolute reliability [25,26]. To determine concurrent validity, a Pearson correlation analysis was performed between spatiotemporal parameters obtained from MotionMetrix and those obtained from the OptoGait system and video analysis. The following criteria were adopted to interpret the magnitude of correlations between measurement variables: <0.1 (trivial), 0.1-0.3 (small), 0.3-0.5 (moderate), 0.5-0.7 (large), 0.7-0.9 (very large), and 0.9-1.0 (almost perfect) [27]. The ICCs were also calculated between systems (MotionMetrix vs OptoGait and MotionMetrix vs video analysis) for spatiotemporal parameters during running. Based on the characteristics of this experimental design and the guidelines reported by Koo and Li [28], we decided to use a two-way random-effects model (ICC [2,k]), mean of measurements, and absolute definition for the ICC measurement. The interpretation of the ICC was based on the benchmarks reported by a previous study [29]: ICC<0 (poor), ICC 0-0.20 (slight), ICC 0.21-0.40 (fair), ICC 0.41-0.60 (moderate), ICC 0.61-0.80 (substantial), and ICC>0.81 (almost perfect). Additionally, the 95% CI of the ICC value was provided. Finally, Bland-Altman plots (ie, limits of agreement method; mean difference [1.96 SD]) [30] were constructed to examine the presence of systematic and proportional bias between video analysis at 1000 Hz and the OptoGait system and estimated values (ie, Kinect with MotionMetrix system) of spatiotemporal parameters during running. Heteroscedasticity of error was defined as r^2>0.1 [25]. The level of significance used was P<.05. Data analysis was performed using SPSS (version 23).

XSL•FO
RenderX

Results

Table 1 shows descriptive values of spatiotemporal parameters acquired from three different systems and the differences between systems (in absolute and relative values). As a measure of absolute reliability, Table 2 shows the CVs of spatiotemporal parameters obtained from the three different systems. The Kinect+MotionMetrix system obtained higher CVs than the reference systems for all the spatiotemporal parameters. Regardless of the system, flight time had the highest CVs (OptoGait: 16.4%; video analysis: 17.3%; Kinect+MotionMetrix*:* 23.2%). The rest of CVs reported were lower than 8.1%.

Table 1. Descriptive data of spatiotemporal parameters obtained from different systems (ie, Kinect with the MotionMetrix software, OptoGait, and high-speed video analysis at 1000 Hz).

Variable	OG[a], mean (SD)	VA[b], mean (SD)	MM[c], mean (SD)	MM-OG difference, mean (%)[d]	MM-VA difference, mean (%)
Contact tine (s)	0.265 (0.015)	0.262 (0.013)	0.273 (0.021)	0.008 (3.0)	0.011 (4.2)
Flight time (s)	0.089 (0.018)	0.097 (0.017)	0.082 (0.019)	−0.007 (−7.9)	−0.011 (−11.3)
Step frequency (spm[e])	166.60 (6.80)	165.76 (6.90)	167.23 (7.28)	0.63 (0.4)	1.47 (0.9)
Step length (cm)	115.53 (6.80)	116.58 (5.92)	115.44 (5.26)	−0.09 (−0.1)	−1.14 (−1.0)

[a]OG: OptoGait system.

[b]VA: High-speed video analysis.

[c]MM: Kinect system with the MotionMetrix software.

[d]%: Refers to values reported by the reference system.

[e]spm: Steps per minute.

Table 2. Coefficients of variation (%) of spatiotemporal parameters during running at 12 km/h obtained from three different systems.

Variable	OG[a]	VA[b]	MM[c]
Contact time, CV[d]	5.7	5.2	8.1
Flight time, CV	16.4	17.3	23.2
Step frequency, CV	4.1	4.2	4.4
Step length, CV	4.6	4.6	5.02

[a]OG: OptoGait system.

[b]VA: High-speed video analysis.

[c]MM: Kinect system with the MotionMetrix software.

[d]CV: Coefficient of variation.

A Pearson correlation analysis was conducted, and ICCs between systems, as well as their CIs, were calculated (Table 3). In the comparison between the OptoGait and Kinect+MotionMetrix systems, a significant substantial correlation (r>0.645; P<.001; ICC=0.712) was obtained for contact time, whereas significant, almost perfect correlations were found for flight time, step frequency, and step length (r>0.901; P<.001; ICCs>0.89). Similarly, the video analysis versus Kinect+MotionMetrix comparison showed a significant substantial correlation (r>0.664, P<.001; ICC=0.667) for contact time and significant, almost perfect correlations for the rest of variables (r>0.928; P<.001; ICCs>0.84).

XSL•FO
RenderX

Table 3. Pearson correlation analysis and intraclass correlation coefficients between spatiotemporal parameters obtained from the Kinect system with the MotionMetrix software and those obtained from two different systems (ie, OptoGait and high-speed video analysis) during running at a comfortable speed.

Variables	MM[a] versus OG[b]			MM versus VA[c]		
	r	*P* value	ICC[d] (95% CI)	*r*	*P* value	ICC (95% CI)
Contact time	0.645	<.001	0.712 (0.333-0.874)	0.664	<.001	0.667 (0.153-0.862)
Flight time	0.901	<.001	0.894 (0.420-0.967)	0.928	<.001	0.838 (0.159-0.961)
Step frequency	0.962	<.001	0.978 (0.951-0.990)	0.940	<.001	0.959 (0.873-0.984)
Step length	0.971	<.001	0.986 (0.968-0.994)	0.954	<.001	0.964 (0.881-0.986)

[a]MM: Kinect system with the MotionMetrix software.

[b]OG: OptoGait system.

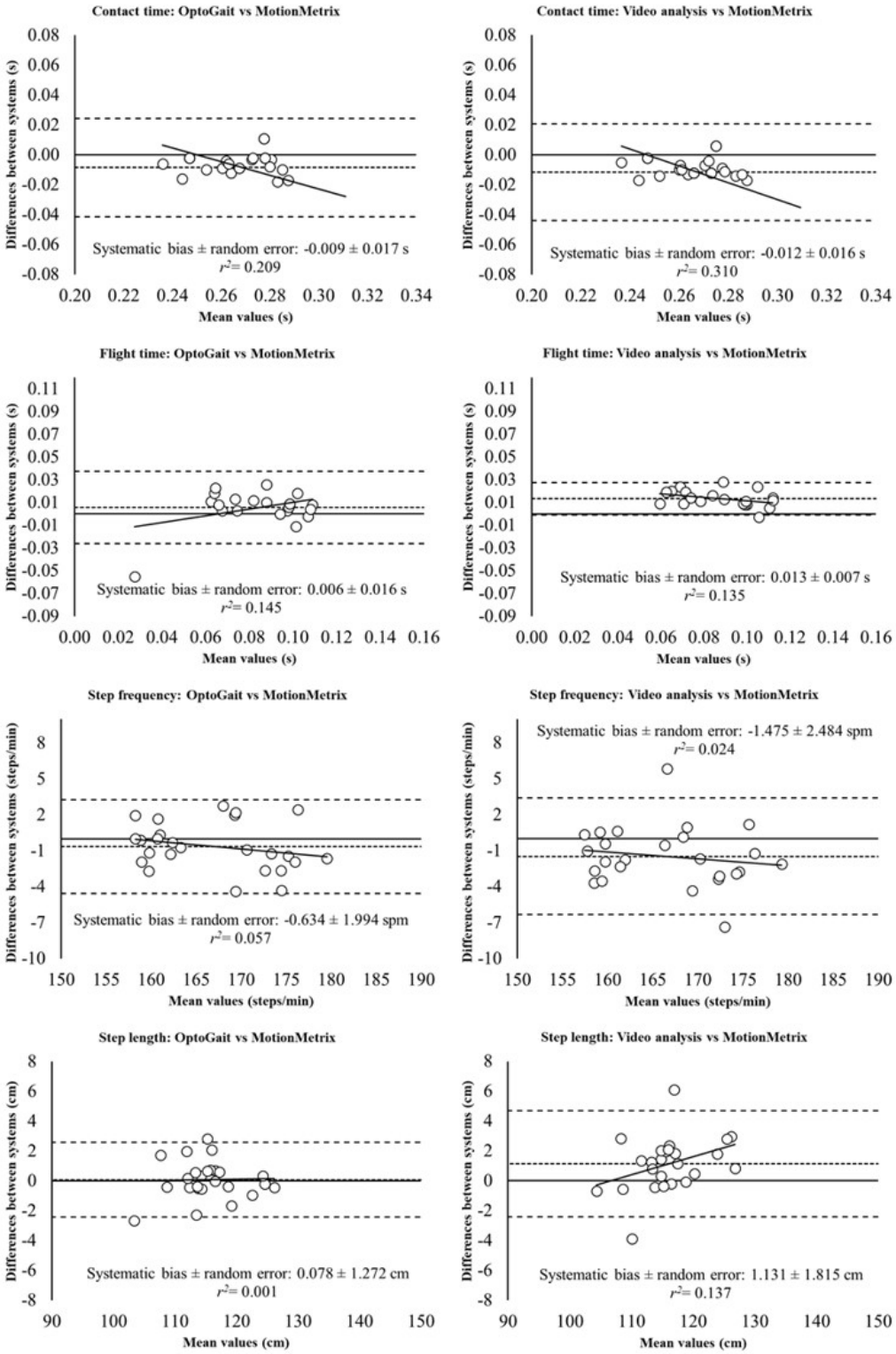
[c]VA: High-speed video analysis.

[d]ICC: intraclass correlation coefficient.

Through Bland-Altman plots, the differences (systematic bias and random error) and the degree of agreement between systems (95% limits of agreement) were determined (Figure 1). The plots of OptoGait and Kinect+MotionMetrix revealed small systematic biases and random errors for step frequency and step length (step frequency difference: mean −0.634 steps per minute, SD 1.994 SPM; step length difference: mean 0.078 cm, SD 1.272 cm), but greater systematic biases and random errors for contact time and fight time (contact time difference: mean −0.009 s, SD 0.017 s; flight time difference: mean 0.006 s, SD 0.016 s) during running at a comfortable velocity. Accordingly, heteroscedasticity was found in temporal parameters (contact time and flight time, r^2>0.1), whereas homoscedasticity was found in step frequency and step length (r^2<0.1). Similarly, when data were compared between video analysis and the Kinect+MotionMetrix system, small systematic biases and random errors for step frequency and step length were found (step frequency difference: mean −1.475 SPM, SD 2.484 SPM; step length difference: mean 1.131 cm, SD 1.815 cm), whereas greater biases and errors were found for contact time and flight time (contact time difference: mean −0.012 s, SD 0.016 s; flight time difference: mean 0.013 s, SD 0.007 s). In this case, homoscedasticity was found only in step frequency (r^2=0.024), whereas the rest of parameters revealed heteroscedasticity (r^2>0.1).

XSL•FO
RenderX

Figure 1. Bland-Altman plots for the measurement of spatiotemporal parameters (ie, from up to bottom: contact time, flight time, step frequency, and step length) during running at a comfortable speed obtained from Kinect with MotionMetrix software and two different systems (OptoGait and high-speed video analysis at 1000 Hz). The plot includes the mean difference (dotted line) and 95% limits of agreement (dashed line), along with the regression line (solid line).



Discussion

This study aimed to evaluate the absolute reliability and concurrent validity of the Kinect with MotionMetrix software for measuring spatiotemporal variables during running at a comfortable velocity by comparing data from the combination system to data obtained from two widely used systems—the OptoGait system and high-speed video analysis at 1000 Hz. Our results showed that the Kinect+MotionMetrix combination reported higher CVs than the reference systems, even though values higher than 10% only were found for flight time, which is similar to the other systems tested in this study (ie, OptoGait and high-speed video analysis). Additionally, regarding the

XSL•FO
RenderX

concurrent validity, an almost perfect level of agreement between systems was found for all the spatiotemporal parameters except contact time, which reported a moderate agreement.

There is limited research on the reliability and validity of markerless motion capture systems for measuring biomechanical parameters during walking or running on a treadmill. Some studies have examined the validity of the Kinect sensor for the assessment of gait characteristics [1,3-7], even though methodological inconsistencies (eg, software used, filters and calibrations applied, testing protocol used, and conditions present, etc) make comparisons difficult.

In that context, findings that have been reported about the validity of the Kinect system for measuring spatiotemporal parameters are controversial. A previous study [5] concluded that the Kinect for Windows is a valid tool for measuring the spatiotemporal parameters of gait during walking. Other studies [3,4] have reported important differences between spatiotemporal parameters measured from a 3D motion capture system and those measured from the Kinect system. Clark et al [4] indicated that the Kinect system reported lower values (ie, −16% step time, −19% stride time, and −1.7% step length) than a 3D system during walking. Similarly, Xu et al [3] indicated that the Kinect system reported valid step time and stride time values, but shorter stance times (ie, −9%) than a 3D system during walking. Therefore, it seems that the validity of the Kinect system for measuring spatiotemporal parameters is highly dependent on variables, such as the software and filter setting used, the gold standard or reference system tested, the protocol performed, and the target variables measured.

Since it has been demonstrated that running on a treadmill carries some biomechanical differences compared to above-ground running [31], caution must be taken when interpreting the results. Some of these studies, which were focused on determining the validity of the Kinect system, were conducted above the ground [4,5], while only 3 studies were conducted on a treadmill [3,6,7].

To the best of our knowledge, only one study [6] has examined the validity of the system during running. This is an important point because validity and reliability data during walking should not be extrapolated to running conditions, since the magnitude of the parameters change and different phases appear (ie, fight time does not exist during walking and there is no double-support time during running). The aforementioned study [6] examined sagittal plane gait kinematics with no mention of spatiotemporal parameters at different walking and jogging velocities (ie, walking velocity of 4.8 km/h increasing to a jogging velocity of 8.8 km/h) being lower than the velocity in this study (ie, 12 km/h), and the authors concluded that the measurement accuracy of the Kinect system was not acceptable for clinical measurement analysis (ie, the system did not provide consistent hip or knee measurements compared to a 3D system). Notably, Pfister et al [6] used old software combined with the Kinect system (ie, Brekel Kinect software), which might explain the differences between the Pfister et al [6] study and our study. The Brekel software worked at 30 Hz, while the software used in this study (ie, MotionMetrix) can reach 60 Hz. Therefore, a higher accuracy is expected with our study. Ours is the first study to examine the validity of the Kinect with MotionMetrix software for measuring spatiotemporal variables while participants ran on a treadmill at a comfortable velocity.

Ultimately, the main limitation of this study was the differences in the precision of the systems. Both the OptoGait system and high-speed video analysis captured data at 1000 Hz, whereas the Kinect with MotionMetrix software works at 60 Hz. This means that this system has a precision of 0.017 s, while the reference systems have a precision of 0.001 s. This point might explain the differences found in the CVs. Therefore, the differences between systems might be explained by the limitations of the markerless Kinect with MotionMetrix software.

In summary, our results indicate that the Kinect system with the MotionMetrix software slightly overestimates contact time and strongly underestimates flight time compared to both the OptoGait system (difference: contact time +3.0%, flight time: −7.9%) and high-speed video analysis at 1000 Hz (difference: contact time +4.2%, flight time: −11.3%). However, it is a valid tool for measuring step frequency and step length when compared to these reference systems (differences lower than 1%). Future studies should determine the reliability of this system for determining flight time with a CV of around 23%.

From a practical perspective, spatiotemporal gait characteristics are readily assessable by such software (ie, MotionMetrix) in conjunction with two Kinect sensors attached to a treadmill after a simple 30-second calibration, even though users must be aware of the characteristic of measurements. Further clinical implications of this system include low cost, time efficient, and wide availability.

Conflicts of Interest

None declared.

References

1. Schmitz A, Ye M, Shapiro R, Yang R, Noehren B. Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system. J Biomech 2014 Jan 22;47(2):587-591. [doi: 10.1016/j.jbiomech.2013.11.031] [Medline: 24315287]
2. van der Kruk E, Reijne MM. Accuracy of human motion capture systems for sport applications; state-of-the-art review. Eur J Sport Sci 2018 Jul 09;18(6):806-819. [doi: 10.1080/17461391.2018.1463397] [Medline: 29741985]

XSL•FO
RenderX

3. Xu X, McGorry RW, Chou L, Lin J, Chang C. Accuracy of the Microsoft Kinect for measuring gait parameters during treadmill walking. Gait Posture 2015 Jul;42(2):145-151 [FREE Full text] [doi: 10.1016/j.gaitpost.2015.05.002] [Medline: 26002604]
4. Clark RA, Bower KJ, Mentiplay BF, Paterson K, Pua Y. Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. J Biomech 2013 Oct 18;46(15):2722-2725. [doi: 10.1016/j.jbiomech.2013.08.011] [Medline: 24016679]
5. Dolatabadi E, Taati B, Mihailidis A. Concurrent validity of the Microsoft Kinect for Windows v2 for measuring spatiotemporal gait parameters. Med Eng Phys 2016 Sep;38(9):952-958. [doi: 10.1016/j.medengphy.2016.06.015] [Medline: 27387901]
6. Pfister A, West AM, Bronner S, Noah JA. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. J Med Eng Technol 2014 Jul 30;38(5):274-280. [doi: 10.3109/03091902.2014.909540] [Medline: 24878252]
7. Lamine H, Bennour S, Laribi M, Romdhane L, Zaghloul S. Evaluation of Calibrated Kinect Gait Kinematics Using a Vicon Motion Capture System. Comput Methods Biomech Biomed Engin 2017 Oct 27;20(sup1):111-112. [doi: 10.1080/10255842.2017.1382886] [Medline: 29088586]
8. Auvinet E, Multon F, Manning V, Meunier J, Cobb J. Validity and sensitivity of the longitudinal asymmetry index to detect gait asymmetry using Microsoft Kinect data. Gait Posture 2017 Jan;51:162-168. [doi: 10.1016/j.gaitpost.2016.08.022] [Medline: 27776270]
9. Chung H, Chung Y, Liang C. Design and Implementation of a Novel System for Correcting Posture Through the Use of a Wearable Necklace Sensor. JMIR Mhealth Uhealth 2019 May 28;7(5):e12293 [FREE Full text] [doi: 10.2196/12293] [Medline: 31140439]
10. García-Pinillos F, García-Ramos A, Ramírez-Campillo R, Latorre-Román PÁ, Roche-Seruendo L. How Do Spatiotemporal Parameters and Lower-Body Stiffness Change with Increased Running Velocity? A Comparison Between Novice and Elite Level Runners. J Hum Kinet 2019 Nov;70:25-38 [FREE Full text] [doi: 10.2478/hukin-2019-0036] [Medline: 31915473]
11. Schieb DA. Kinematic Accommodation of Novice Treadmill Runners. Res Q Exerc Sport 1986 Mar;57(1):1-7. [doi: 10.1080/02701367.1986.10605381]
12. Lavcanska V, Taylor NF, Schache AG. Familiarization to treadmill running in young unimpaired adults. Hum Mov Sci 2005 Aug;24(4):544-557. [doi: 10.1016/j.humov.2005.08.001] [Medline: 16176843]
13. Radzak KN, Putnam AM, Tamura K, Hetzler RK, Stickley CD. Asymmetry between lower limbs during rested and fatigued state running gait in healthy individuals. Gait Posture 2017 Jan;51:268-274. [doi: 10.1016/j.gaitpost.2016.11.005] [Medline: 27842295]
14. Healy A, Linyard-Tough K, Chockalingam N. Agreement Between the Spatiotemporal Gait Parameters of Healthy Adults From the OptoGait System and a Traditional Three-Dimensional Motion Capture System. J Biomech Eng 2019 Jan 01;141(1):1-4. [doi: 10.1115/1.4041619] [Medline: 30285079]
15. Gomez Bernal A, Becerro-de-Bengoa-Vallejo R, Losa-Iglesias ME. Reliability of the OptoGait portable photoelectric cell system for the quantification of spatial-temporal parameters of gait in young adults. Gait Posture 2016 Oct;50:196-200. [doi: 10.1016/j.gaitpost.2016.08.035] [Medline: 27644096]
16. Lienhard K, Schneider D, Maffiuletti NA. Validity of the Optogait photoelectric system for the assessment of spatiotemporal gait parameters. Med Eng Phys 2013 Apr;35(4):500-504. [doi: 10.1016/j.medengphy.2012.06.015] [Medline: 22818403]
17. Lee MM, Song CH, Lee KJ, Jung SW, Shin DC, Shin SH. Concurrent Validity and Test-retest Reliability of the OPTOGait Photoelectric Cell System for the Assessment of Spatio-temporal Parameters of the Gait of Young Adults. J Phys Ther Sci 2014 Jan;26(1):81-85 [FREE Full text] [doi: 10.1589/jpts.26.81] [Medline: 24567681]
18. Lee M, Song C, Lee K, Shin D, Shin S. Agreement between the spatio-temporal gait parameters from treadmill-based photoelectric cell and the instrumented treadmill system in healthy young adults and stroke patients. Med Sci Monit 2014 Jul 14;20:1210-1219 [FREE Full text] [doi: 10.12659/MSM.890658] [Medline: 25017613]
19. Jaén-Carrillo D, García-Pinillos F, Cartón-Llorente A, Almenar-Arasanz AJ, Bustillo-Pelayo JA, Roche-Seruendo LE. Test–retest reliability of the OptoGait system for the analysis of spatiotemporal running gait parameters and lower body stiffness in healthy adults. Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology 2020 Feb 03;234(2):154-161. [doi: 10.1177/1754337119898353]
20. Hanley B, Tucker CB. Reliability of the OptoJump Next System for Measuring Temporal Values in Elite Racewalking. J Strength Cond Res 2019 Dec;33(12):3438-3443. [doi: 10.1519/JSC.0000000000003008] [Medline: 30640307]
21. Balsalobre-Fernández A, Agopyan H, Morin JB. The Validity and Reliability of an iPhone App for Measuring Running Mechanics. J Appl Biomech 2017 Jul;33(3):222-226. [doi: 10.1123/jab.2016-0104] [Medline: 27918692]
22. Esculier JF, Silvini T, Bouyer LJ, Roy J. Video-based assessment of foot strike pattern and step rate is valid and reliable in runners with patellofemoral pain. Phys Ther Sport 2018 Jan;29:108-112. [doi: 10.1016/j.ptsp.2016.11.003] [Medline: 28666810]
23. Dingenen B, Barton C, Janssen T, Benoit A, Malliaras P. Test-retest reliability of two-dimensional video analysis during running. Phys Ther Sport 2018 Sep;33:40-47. [doi: 10.1016/j.ptsp.2018.06.009] [Medline: 30005426]
24. Pipkin A, Kotecki K, Hetzel S, Heiderscheit B. Reliability of a Qualitative Video Analysis for Running. J Orthop Sports Phys Ther 2016 Jul;46(7):556-561. [doi: 10.2519/jospt.2016.6280] [Medline: 27266886]

XSL•FO
RenderX

25. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998 Oct;26(4):217-238. [doi: 10.2165/00007256-199826040-00002] [Medline: 9820922]
26. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000 Jul;30(1):1-15. [doi: 10.2165/00007256-200030010-00001] [Medline: 10907753]
27. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc 2009 Jan;41(1):3-13. [doi: 10.1249/MSS.0b013e31818cb278] [Medline: 19092709]
28. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
29. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977 Mar;33(1):159. [doi: 10.2307/2529310]
30. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. Lancet 1995 Oct;346(8982):1085-1087. [doi: 10.1016/s0140-6736(95)91748-9]
31. Van Hooren B, Fuller JT, Buckley JD, Miller JR, Sewell K, Rao G, et al. Is Motorized Treadmill Running Biomechanically Comparable to Overground Running? A Systematic Review and Meta-Analysis of Cross-Over Studies. Sports Med 2020 Apr 4;50(4):785-813 [FREE Full text] [doi: 10.1007/s40279-019-01237-z] [Medline: 31802395]

Abbreviations

CV: coefficient of variation
ICC: intraclass correlation coefficient
LED: light-emitting diode
SPM: steps per minute

Edited by G Eysenbach; submitted 20.04.20; peer-reviewed by M Narasaki Jara, V Shah; comments to author 21.08.20; revised version received 25.08.20; accepted 13.09.20; published 23.10.20.

Please cite as:
García-Pinillos F, Jaén-Carrillo D, Soto Hermoso V, Latorre Román P, Delgado P, Martinez C, Carton A, Roche Seruendo L
Agreement Between Spatiotemporal Gait Parameters Measured by a Markerless Motion Capture System and Two Reference Systems—a Treadmill-Based Photoelectric Cell and High-Speed Video Analyses: Comparative Study
JMIR Mhealth Uhealth 2020;8(10):e19498
URL: http://mhealth.jmir.org/2020/10/e19498/
doi:10.2196/19498
PMID:33095181

©Felipe García-Pinillos, Diego Jaén-Carrillo, Victor Soto Hermoso, Pedro Latorre Román, Pedro Delgado, Cristian Martinez, Antonio Carton, Luis Roche Seruendo. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 23.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

African American Adolescents and Young Adults, New Media, and Sexual Health: Scoping Review

Sierra Teadt[1], BA; Jade C Burns[2], PhD; Tiffany M Montgomery[3], PhD; Lynae Darbes[2], PhD

[1]School of Public Health, University of Michigan, Ann Arbor, MI, United States

[2]Department of Health Behavior and Biological Sciences, School of Nursing, University of Michigan, Ann Arbor, MI, United States

[3]College of Nursing & Health Professions, Drexel University, Philadelphia, PA, United States

Corresponding Author:
Sierra Teadt, BA
School of Public Health
University of Michigan
1415 Washington Heights
Ann Arbor, MI, 48109
United States
Phone: 1 2692547062
Email: steadt@umich.edu

Abstract

Background: Rates of sexually transmitted infections and unintended pregnancies are disproportionately high among African American adolescents and young adults (AYA). New media platforms such as social networking sites, microblogs, online video sites, and mobile phone applications may be a promising approach in promoting safe sex and preventing sexually transmitted infections.

Objective: The purpose of this scoping review was to address promising approaches in new media that may serve as valuable tools in health promotion, prevention, education, and intervention development aimed at African American AYA.

Methods: An electronic search was conducted using Google Scholar, Scopus, Cumulative Index to Nursing and Allied Health (CINHAL), and PubMed online databases. Concept blocks and MeSH terminology were used to identify articles around African American youth and new media.

Results: The search yielded 1169 articles, and 16 publications met the criteria. Studies from the review found themes in new media that included feasibility, changing attitudes, and improving knowledge related to sexual health behavior among youth of color.

Conclusions: New media is a promising and feasible platform for improving the sexual health of African American AYA. Further research is suggested to better understand the benefits of new media as a sexual health promotion tool among this specific population.

(JMIR Mhealth Uhealth 2020;8(10):e19459) doi:10.2196/19459

KEYWORDS

African American; adolescent; young adult; technology; safe sex; sexually transmitted infections; sexual behavior; new media; social media; internet

Introduction

Rates of sexually transmitted infections (STIs) and unintended pregnancies are disproportionately high among African American adolescents and young adults [1]. Multiple personal and social factors contribute to this increased risk, including (1) having more sex partners than youth of other ethnicities, (2) early sexual debut, (3) inconsistent condom use, and (4) limited access to sexual health promotion resources [2-5]. STI prevalence and other risk factors among this population are particularly acute in minority communities of low socioeconomic status [6].

While leading public health and medical organizations support comprehensive sex education for adolescents and young adults, there are several limitations of traditional or "formal" sex education platforms that often take place in structured settings (eg, schools, youth centers) [7]. According to the Centers for Disease Control and Prevention (CDC), sexual health education

XSL•FO
RenderX

is more commonly required in high school than middle or elementary school [7]. As a result, this information is provided too late for many African American adolescents, as it is estimated that 20% of youth have reported having had sex by the age of 15 years, and it is known that African American youth report age at first sexual intercourse earlier than their peers [8]. Additionally, most schools fail to provide instruction covering all 16 sexual health topics that the CDC considers essential; less than half of high schools and only 20% of middle schools cover all 16 topics [7]. Furthermore, only 35% of high school students and only 10% of middle school students were taught how to correctly use a condom in 2014 [7]. Lastly, 88% of schools in the United States allowed parents to exempt their children from sexual health education in 2014 [7].

New platforms for sexual health promotion and prevention aimed at African American adolescents and young adults are especially important, as this population faces multiple barriers to active engagement in health care [9]. Such barriers include the inability to pay for health care services, lack of transportation, long waiting times, conflict with work or school schedules, confidentiality concerns, and embarrassment that is attached to sexual health services [10,11]. Additional barriers include actual or perceived fear and distrust of health care institutions, discrimination, and provider bias [11]. Additionally, minority youth in urban settings report prioritizing basic needs, such as housing, food, and transportation, over HIV risk reduction or prevention [1]. These individuals often experience difficulty when navigating and affording quality sexual health services [12]. Similarly, minority youth have higher rates of medical poverty compared to their white counterparts [1]. The financial, cultural, and institutional barriers to health care among minority youth often result in a lack of access to comprehensive adolescent health services, including sexual health services [1].

One promising approach to improving adolescents' and young adults' high-risk sexual behavior and filling critical gaps in knowledge due to formal sex education platforms may be to promote safe sex and STI prevention via new media [13]. These platforms include social networking sites, microblogs, online video sites, and mobile phone applications (see Table 1) [13]. African American adolescents not only use cell phones and the internet at high rates but they are also open to seeking and receiving sexual health information via new media due to its accessibility and ability to provide a wide range of information [14]. While African American adolescents and young adults are actively using new media on a day-to-day basis, it is essential to assess the ways these platforms have been — and could be — leveraged to promote engagement in health care and reduce risk-related behaviors [15].

Table 1. Classification of new media.

Types of new media[a]	Primary purpose	Examples
Social networking sites	Peer networking	Facebook, Instagram, MySpace
Collaborative websites	Information sharing, discussion	Wikipedia, AskFM, answer.com
Blogs (and microblogs)	Opinion sharing, discussion	Twitter, Tumblr, Blogger
Content communities	Entertainment, information sharing	YouTube, Snapchat, Reddit
Virtual reality/online gaming	Simulate experiences, entertainment	It's Your Game: Keep it Real, PlayForward
Communication/messengers	Discussion	WhatsApp, Facebook Messenger, GroupMe

[a]Not inclusive of all new media platforms that exist today. Adapted from [16].

Smartphone use and engagement with social media are high among African American adolescents and young adults [17]. They are already being used by this age group to search for general health information and sensitive health topics [17]. New media platforms provide a venue for greater anonymity and client sensitivity, which are critical to an adolescent's self-expression [18,19]. As the use of new media continues to increase among adolescents and young adults, these platforms may serve as valuable tools in health promotion, prevention, education, and intervention development aimed at this population.

Many evidence-based sexual risk-reduction interventions target African American youth. For instance, the "Compendium of Evidence-Based Interventions and Best Practices for HIV Prevention" is a collection of good and best-evidence sexual risk-reduction interventions compiled by the CDC [20]. This collection includes 59 evidence-based interventions, 10 of which specifically target African Americans and 3 of which target African American youth. None of these interventions include new media as a method for intervention delivery. So, while the extent of the sexual health risk among African American adolescents is evident, it is not clear which health promotion strategies (eg, peer-to-peer vs online) will be most effective in reducing risky sexual behaviors among this population [21].

Accordingly, the goal of this review was to investigate the various forms of new media, the current state of evidence on the platforms used, and how it can improve health care engagement and sexual health outcomes among African American youth**.** The following research question is addressed: How is new media useful in the delivery of STI prevention and risk-reduction interventions among African American youth?

Methods

Search Strategy

A scoping review methodology was selected as it (1) helped to identify review parameters, (2) identified a process of mapping the existing literature, and (3) explored a research gap [22,23]. We used the framework by Arksey and O'Malley [22] along with the Preferred Reporting Items for Systematic Reviews and

XSL•FO
RenderX

Meta-Analyses (PRISMA) guidelines [23,24] for this review. An informationist was consulted to provide expertise in creating tailored search strategies on this topic. The literature search occurred between March 2018 and December 2018 using the following electronic search engines and databases: Google Scholar, Scopus, Cumulative Index to Nursing and Allied Health (CINHAL), and PubMed. Concept blocks were used in each search engine to combine keywords in the title and abstract and used with MeSH terminology. A manual review of systematic reviews was also employed. See Table 2 for a list of standard terms used. A name search was then conducted on articles from technology and media experts along with a manual review of relevant articles and a CDC list of evidence-based interventions around youth and sexual health. Finally, a table of evidence was constructed to organize the articles by the level of evidence, sample size, type of new media used, results, and any limitations within the study.

Table 2. Search strategy concept blocks.

Concept block	MeSH terms	Title and abstract terms	Additional terms
Adolescent	Adolescent, young adult	Adolescent, youth, young people, young adult	Teenager, teen
New media	Social media, cell phone, internet, telemedicine, text messaging, multimedia, mobile applications, smartphone	Social media, new digital media, internet, mobile phones, text messaging, Facebook, instant messaging, multimedia, online social networks, computer, technology, mobile health, smartphone, Web 2.0, eHealth, mHealth[a], SMS	Apps, SNS[b], social networking sites
Sexual health	Risk reduction behavior, safe sex, sexually transmitted diseases, condoms, HIV infections, sex education, sexual behavior	Risk reduction, sexual health, HIV/STI[c] risk, HIV prevention, sexually transmitted diseases, sexual practices, sexually transmitted infection	N/A[d]
African American	African American	African American	Black, minority group

[a]mHealth: mobile health.

[b]SNS: social networking site.

[c]STI: sexually transmitted infection.

[d]N/A: not applicable.

Selection and Assessment of Articles

An initial search trail document was created to detail all review findings. Duplicates were screened for independently and then by team review. Next, abstracts were reviewed for relevance, followed by a full-text review. Discrepancies related to the retention or removal of articles were discussed until consensus was achieved. Finally, themes were created from the existing domains. Inclusion criteria included (1) African American adolescents and young adult participants aged 13-24 years, (2) sexual health, (3) new media use as defined in our background section, (4) publication after 2009, and (5) publication within the United States. Study samples were required to reflect the US African American population (13%) [25]. Exclusion criteria included (1) non-English language articles; (2) letters to the editor, opinions, commentaries, and narrative reviews; (3) studies used for recruitment only; and (4) text messaging, which has been well-covered within the literature for the adolescent and young adult population.

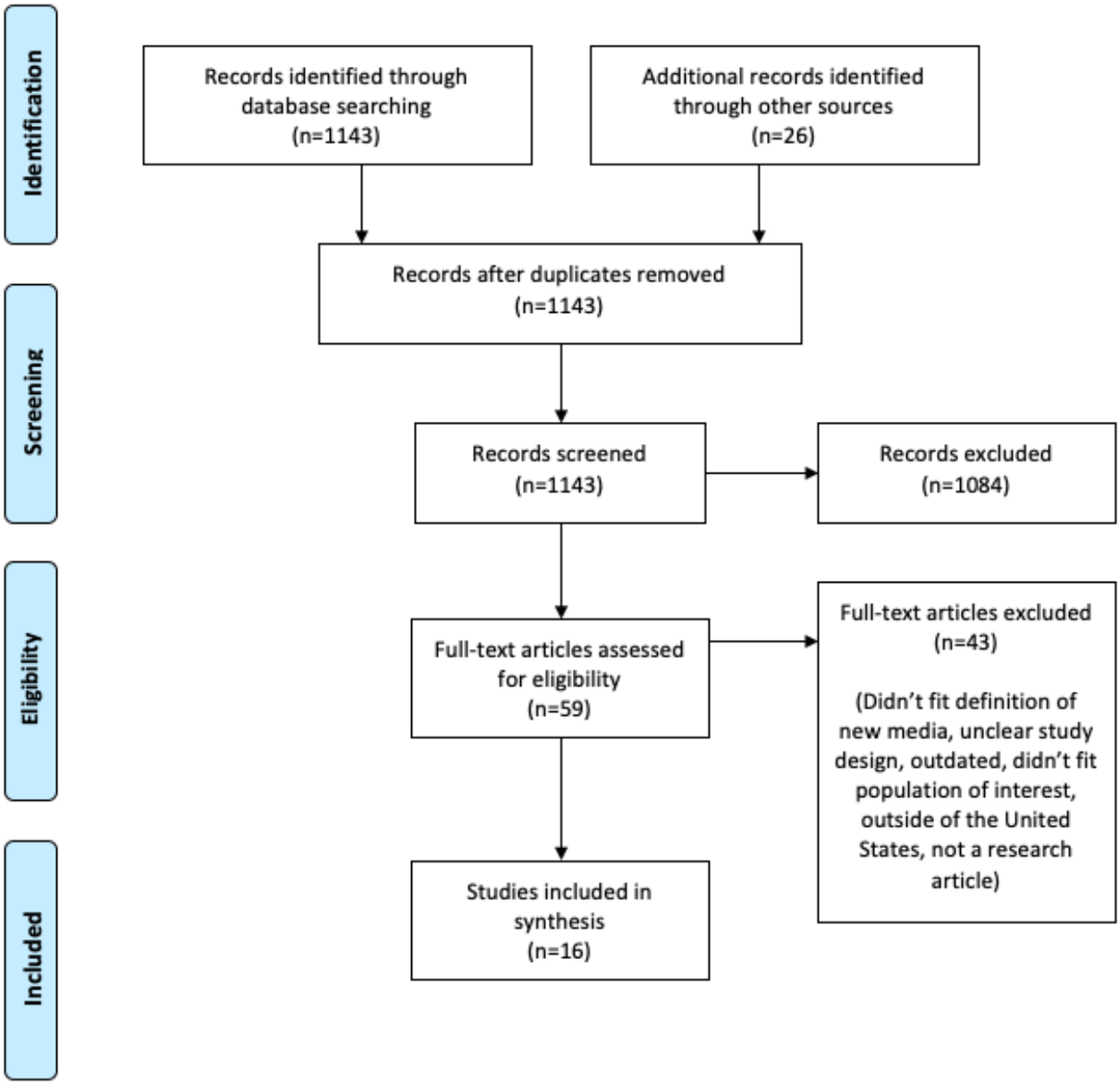
Results

Descriptive Characteristics of Reviewed Articles

In total, 16 selected studies [15,26-40] met the inclusion criteria (see PRISMA diagram; Figure 1). Of these, 10 studies used quantitative methods [27,28,30,33,35-40], 5 studies used qualitative methods [15,26,29,31,32], and 1 study used mixed methods [34]. Reviews included in the synthesis were categorized as quantitative, qualitative, or mixed methods based on the numerical or observational nature of the studies included. The 16 included studies were summarized by the study method, type of new media platform, and sample. The most common forms of new media utilized within the included studies were social media (eg, Facebook, Twitter) [15,26-29,33-35,37,40], internet-based interventions (eg, *It's Your Game*-Tech, *Keep It Up!*) [28,30,32,33,36], mobile applications [29,31,33,39], and interactive video games [15,38]. Studies reported various reasons for utilizing new media, including improving contraception or condom use, communicating credible information regarding HIV and STIs, reducing the transmission of HIV and STIs, improving attitudes around sexual health, and promoting STI testing–related behaviors. Around half of the studies (7/16, 44%) indicated utilizing new media as an effective sexual health promotion tool due to its ease of use and wide accessibility among adolescents and young adults [15,26,30-32,34,39].

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram.



Synthesis of Results Table

Table 3 shows the resulting articles by the level of evidence, sample size, type of new media used, results, and any limitations within the study.

XSL•FO
RenderX

Table 3. Synthesis of results.

First author, year	Study method	New media platform	Sample
Bull, 2012 [27]	Cluster randomized controlled trial (RCT)	Facebook	1578 youth, 16-25 years old; 35% AA[a]
Condran, 2017 [35]	Scoping review	Social media (eg, Facebook, Twitter, Instagram)	24 articles, some AA youth
Cordova, 2015 [31]	Qualitative interviews	mHealth[b] app	29 adolescents, 13-18 years old; 65% AA
Cordova, 2017 [36]	Systematic review	Internet	3 articles; >75% AA or Hispanic youth, 13-25 years old
Dolcini, 2015 [32]	Qualitative investigation	Internet	81 AA youth, 15-17 years old
Eason, 2017 [37]	Cross-sectional study	Facebook	112 AA adults, 18-49 years old
Fiellin, 2017 [38]	RCT	Interactive video game	333 youth, 11-14 years old; 88.6% racial/ethnic minorities
Guilamo-Ramos, 2014 [14]	Focus groups	Social networking sites, games, mobile phones	106 youth, 12-19 years old; 53% AA
Guse, 2012 [28]	Systematic review	Web-based, social networking sites	10 articles; youth 13-24 years old, some AA
Jemmott, 2017 [39]	Pilot study	iPad app	4 AA men, 18-24 years old
Muessig, 2015 [33]	Systematic review	Web-based, social media, smartphone apps	61 articles; some AA youth, 15-25 years old
Shegog, 2014 [30]	Feasibility study	Web-based	33 youth, 12-14 years old; 70% AA
Stevens, 2017 [40]	Cross-sectional study	Social media (eg, Facebook, Instagram, Twitter)	249 AA and Latino youth, 13-24 years old
Taggart, 2015 [34]	Systematic review	Social media	35 articles; 18-40 years old; some AA
Veinot, 2011 [26]	Cohort study	Social networking sites	94 youth, 14-24 years old; 80% AA
Veinot, 2013 [29]	Quasiexperimental study	Social media apps	75 AA youth, 14-24 years old

[a]AA: African American.

[b]mHealth: mobile health.

Feasibility of New Media for Reaching Youth of Color and At-Risk Youth

Several studies cited the ability of new media to reach minority and other at-risk youth regarding sexual health [27,28,34,37]. In one study, retention of participants declined in the long term, but a significant number (1195/1578 participants, 75%) returned to complete a follow-up in the short term [27]. Still, this intervention was successful in recruiting 1578 youth, 773 of whom were minority youth [27]. Additionally, the intervention reached large numbers of youth with STI-related and HIV-related information via Facebook, with most participants viewing the content that they were intended to view; during the study period, there were 277 posts by visitors to the Just/Us Facebook page, and 93 individuals, which represents 10% of those enrolled in the intervention, were identified as “loyal” visitors to the page [27]. Another study revealed targeted Facebook messaging to be effective in reaching young African Americans living in the southeastern United States; 149 of the 176 individuals (85%) who responded to the invitation to participate in the study did so after viewing the Facebook message [37]. In a systematic review analyzing studies that used a wide range of new media (eg, social networking sites and internet-based interventions) for sexual health education, several at-risk populations were recruited, including low-income urban youth, HIV-positive youth, and minority youth [28]. Of the 10 articles included in the review, 6 articles succeeded in recruiting and retaining African American youth to a number of new media–based interventions (eg, +CLICK and MySpace) [28]. Another systematic review illustrated the capacity of new media to engage users spanning various geographic locations, ages, genders, races, and socioeconomic status when communicating about HIV; the capacity was largely due to the ability of users to remain anonymous [34]. The common feature of social media platforms that allows for the anonymity of users may encourage marginalized groups to feel more comfortable when it comes to engaging in HIV communication via such platforms [34]. Users in 6 studies reported that social media anonymity allowed for a decrease in stigma, fear, and discrimination surrounding HIV and therefore allowed participants to engage more in a discussion through social media than they would through in-person interactions [34]. This review also highlighted the usefulness of utilizing new media as a strategy for increasing access to HIV care or prevention for those who may typically face barriers to achieving such access in person, such as marginalized and at-risk groups [34].

XSL•FO
RenderX

Ability of New Media to Change Sexual Health–Related Attitudes and Behaviors

Approximately one-third (5/16, 31%) of the studies explored the effectiveness of new media in changing attitudes and behaviors related to sexual health [30,35,37,38,40]. When comparing current practices to pre-social media service utilization, one scoping review found an increase in the utilization of services, such as an increase in the number of referrals and testing rates when social media was used as a way of reaching populations [35]. In another study, 60 of 112 participants (54%) cited viewing HIV and STI prevention messages on Facebook in the past year as the most critical factor in their decision to change their high-risk sexual behaviors [37]. In yet another study, participants improved their sexual health attitudes and knowledge 12 months after exposure to an interactive video game intervention [38]. Middle-school students who utilized an interactive, internet-based sexual health curriculum demonstrated increased perceptions of friends' positive beliefs about delaying sex, more significant reasons for not having sex, increased self-efficacy for condom use, and greater intentions to abstain from sex until marriage [30]. The final study revealed that youth exposed to sexual health messages through social media were nearly 2.5 times more likely than youth who were not exposed to sexual health messages through social media to have used contraception or a condom at last intercourse. In contrast, parents, schools, and traditional media as information sources were not significantly related to contraception and condom use [40].

Role of New Media in Filling Gaps in Knowledge and Information

Just under half (7/16, 44%) of the selected articles discussed the value of new media in providing information and filling critical gaps in knowledge related to sexual and reproductive health (eg, STI testing and disease prevention and management) [15,26,29,31,34,38,40]. Participants in one study shared the importance of utilizing the app as a tool for disseminating culturally specific HIV and STI information (eg, symptoms and condom use) [36]. In another study, youth who participated in an interactive video game demonstrated an average increase of 1.13 points in sexual health knowledge scores compared to the scores of youth in the control group [38]. An additional study revealed that adolescents are motivated to seek sexual health information through new media due to its accessibility and widespread use, while acknowledging in-person interactions as a frequently used resource regarding specific and reliable information [15]. To address the potential loss of trustworthiness experienced with in-person health education, new media–based interventions can utilize components of in-person interactions (eg, interactivity and specificity) [15]. Another article revealed that social media was the fourth most commonly used source of sexual risk–reduction information among African American and Latino youth, following television and movies, school, and parents [40]. However, youth in this study received information via online spaces at similar levels to information via friends and parents [40]. In one systematic review, the most common benefit of utilizing social media as a tool for HIV communication was the ability to both share and receive information [34]. Social media users also cited additional benefits of this platform: the ability to receive information regarding disease management and the ability to access nontraditional sources of information regarding HIV prevention and testing [34]. Youth who participated in focus groups suggested the use of mobile platforms and applications to address the need for credible information regarding sexual health [26]. However, participants in another study reported inconsistencies in accessing credible HIV-related and STI-related information through information technologies (eg, the internet and mobile phones), resulting in a critical gap in knowledge related to STIs [29]. Youth revealed that their most desired feature of a new media–based intervention was credible information, such as articles or question-and-answer services [29].

Discussion

Principal Findings

This scoping review yielded 16 distinct new media interventions to improve condom use and attitudes around sexual health, communicate credible information regarding HIV and STIs, and promote STI testing behaviors among African American youth. These interventions are distinct in their methods of new media delivery, the outcomes assessed, and their regional locations. However, they hold a shared desire to increase the overall sexual health of African American youth — a doubly vulnerable population at risk for STIs.

Several themes were identified in this review. First, new media has been shown to be a feasible method of delivery for future sexual health interventions. New media interventions received upwards of 75% participation by the African American youth who were targeted for inclusion in the studies. This is higher than the 10%-20% recruitment rate of minorities that is often the reality for the bulk of existing research [41]. One of the reasons cited for high use of new media was anonymity. The desire to conceal their identity may be of particular importance to adolescents who may not want their parents or friends to know that they are engaged in a sexual health education program. Anonymity has been found to increase participation in new media sexual health interventions among other at-risk groups, including adolescents and emerging adults [42] and adolescent women of color [43]. The decreased chances of undesired participant identification, in addition to convenience (ie, no travel requirements or synchronous login requirements) may make new media a better delivery mechanism for sexual health interventions than traditional face-to-face intervention delivery methods for some at-risk adolescents.

Next, several new media interventions identified in this scoping review were effective in changing negative attitudes and behaviors related to sexual health to more positive ones. Attitudes are an important aspect of behavior, as highlighted by the Theory of Reasoned Action (TRA) [44] and Theory of Planned Behavior (TPB) [45]. These theories purport that beliefs influence attitudes; attitudes, in turn, influence intentions; and finally, intentions influence behaviors. Understanding the link between these constructs is an important step in acknowledging the importance of attitude change as a key precursor to behavior change. TRA and TPB have been commonly used as the theoretical underpinnings for behavior change interventions.

XSL•FO
RenderX

Two of the 13 interventions for African Americans that are endorsed in the "Compendium of Evidence-Based Interventions and Best Practices for HIV Prevention" use this theory in their theoretical frameworks [20].

Several new media interventions were also effective in increasing knowledge related to sexual risk reduction and helped to fill information gaps. While it is widely accepted that an increase in knowledge alone does not directly lead to risk reduction or behavior change, this construct is still often utilized by interventionists and health care researchers. It is a major component of Social Cognitive Theory (SCT) and can be applied specifically to health promotion interventions [46]. Increased knowledge may lead to better decision making and healthier behaviors. Thus, the increased knowledge reported by participants of the selected sexual health new media interventions is a positive outcome that may lead to positive changes in their sexual health behaviors. SCT is even more widely used in sexual health interventions than TRA and TPB. Eight of the 13 "Compendium of Evidence-Based Interventions and Best Practices for HIV Prevention" interventions tailored to African American youth and emerging adults utilize SCT as a theoretical underpinning [20].

While new media interventions may decrease risky sexual behavior and improve important aspects of future behavior change, comparison of new media efficacy is difficult. Any attempt to make this type of comparison may be difficult due to the wide variation of core components utilized by each intervention. Studies comparing one sexual health intervention that was implemented using various new media and traditional delivery mechanisms found no difference in outcomes [47,48]. This suggests that mode of delivery does not affect intervention efficacy, but rather, it is a mechanism used to reach a target population. It is the content of an intervention that determines its effectiveness and not its method of delivery. Other researchers have also found that the content of sexual health interventions is more indicative of effectiveness than mode of delivery [19]. Further research on the efficacy of intervention mode of delivery is warranted.

The findings of this scoping review support the use of new media platforms such as social networking sites, microblogs, online video sites, and mobile phone applications to promote safer sex behaviors and STI prevention. New media platforms should be leveraged to promote engagement in health care and reduce risk-related behaviors among African American youth. This population already utilizes new media at high rates, an indication that these platforms are highly accepted modes of communication and information delivery within this group. Moreover, African American youth have reported increased utilization of health care resources after seeing information about these services online. This suggests that new media is not a substitution for in-person services but that it can be used as a conduit for increased utilization of traditionally delivered services.

It is to the advantage of sexual health educators and interventionists to become familiar with and employ platforms that are already in heavy use by their target populations. Health care promotion and disease prevention organizations have encouraged the use of new media for the past decade [49,50]. One important factor in the creation of new media interventions is the use of adaptation methods. Instead of creating entirely new interventions, it is highly suggested to adapt evidence-based interventions for use among new populations and via new delivery methods using models such as Intervention Mapping [51], the CDC's Map of Adaptation Process [52], and the ADAPT-ITT model [53].

Limitations

Several limitations exist for this research. The primary limitation is the use of a broad range of platforms and measures among the studies selected for review. Furthermore, some studies utilized social media platforms, while others utilized web-based gaming platforms. Some studies measured attitudes, while others measured different sexual behaviors such as delayed onset of first sexual encounter or consistent condom use during sexual intercourse. Just as online platforms are similar yet can vary greatly in their functionality and usability, behavioral outcomes may be alike or associated in some ways without being proxies for one another; thus, they should not be compared. For this reason, it is not appropriate to attempt head-to-head comparisons of the new media interventions identified in this study. The varying factors and outcomes assessed in each intervention make it difficult to draw definitive conclusions regarding the relationship between the new media interventions and individual sexual health outcomes or behaviors. Furthermore, only a limited number of new media studies have focused solely on African American youth and young adults, and few of these interventions were gender-related. While African Americans accounted for at least 13% of each study's participants, 11 studies in this review included other populations. Only one study was gender-specific. Therefore, while being culturally appropriate and gender-specific are highly recommended aspects of intervention creation [54-56], they may not have been important factors in many of the interventions included in this review. Finally, it was not in the scope of this study to compare the effectiveness of sexual risk-reduction interventions utilizing traditional delivery methods and those using new media. While it is not highly anticipated that intervention mode of delivery affects behavioral outcomes, this potential impact should be studied further in the presence of all intervention core components.

Conclusions

While research in this area is limited, the results of this scoping review indicate that new media is a promising sexual health promotion tool for African American adolescents and young adults. A range of new media platforms was shown to be effective in reaching African American youth, improving sexual health–related attitudes and behaviors and filling gaps in sexual health–related knowledge and information. The encouraging results of this review suggest more research should be devoted to further exploring the benefits of utilizing new media as a tool for improving the sexual health of African American adolescents and young adults. Additionally, this review supports the value of tailoring both new and existing new media–based interventions towards this specific population.

XSL•FO
RenderX

Acknowledgments

We acknowledge Judith E. Smith, MSLIS, University of Michigan Taubman Health Sciences informationist, for her expertise in the creation of our initial search strategy.

Conflicts of Interest

None declared.

References

1. Augustine J. Youth of Color—At Disproportionate Risk of Negative Sexual Health Outcomes. Advocates for Youth. 2010. URL: https://www.advocatesforyouth.org/wp-content/uploads/storage//advfy/documents/youth_of_color.pdf [accessed 2020-09-21]
2. George Dalmida S, Aduloju-Ajijola N, Clayton-Jones D, Thomas TL, Erazo Toscano RJ, Lewis R, et al. Sexual Risk Behaviors of African American Adolescent Females: The Role of Cognitive and Religious Factors. J Transcult Nurs 2018 Jan 30;29(1):74-83. [doi: 10.1177/1043659616678660] [Medline: 27909236]
3. Adolescent health. Office of Disease Prevention and Health Promotion. 2018. URL: https://www.healthypeople.gov/2020/topics-objectives/topic/Adolescent-Health [accessed 2020-09-21]
4. Kann L, McManus T, Harris WA, Shanklin SL, Flint KH, Queen B, et al. Youth Risk Behavior Surveillance - United States, 2017. MMWR Surveill Summ 2018 Jun 15;67(8):1-114 [FREE Full text] [doi: 10.15585/mmwr.ss6708a1] [Medline: 29902162]
5. Lindberg LD, Maddow-Zimet I, Marcell AV. Prevalence of Sexual Initiation Before Age 13 Years Among Male Adolescents and Young Adults in the United States. JAMA Pediatr 2019 Jun 01;173(6):553-560 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.0458] [Medline: 30958512]
6. Centers for Disease Control and Prevention. STDs in Racial and Ethnic Minorities - 2017 Sexually Transmitted Diseases Surveillance. 2018. URL: https://www.cdc.gov/std/stats17/minorities.htm [accessed 2020-04-03]
7. American Adolescents' Sources of Sexual Health Information. Guttmacher Institute. 2017. URL: https://www.guttmacher.org/fact-sheet/facts-american-teens-sources-information-about-sex# [accessed 2020-06-27]
8. Adolescent Sexual and Reproductive Health in the United States. Guttmacher Institute. 2019. URL: https://www.guttmacher.org/fact-sheet/american-teens-sexual-and-reproductive-health [accessed 2020-06-27]
9. Wong CA, Merchant RM, Moreno MA. Using social media to engage adolescents and young adults with their health. Healthc (Amst) 2014 Dec;2(4):220-224 [FREE Full text] [doi: 10.1016/j.hjdsi.2014.10.005] [Medline: 25984444]
10. Buzi RS, Smith PB. Access to Sexual and Reproductive Health Care Services: Young Men's Perspectives. 2013 Jul 02;40(2):149-157. [doi: 10.1080/0092623x.2012.736923]
11. Sexually transmitted disease surveillance 2017. Centers for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/std/stats17/toc.htm [accessed 2020-09-23]
12. STDs in racial and ethnic minorities. Centers for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/std/stats16/minorities.htm [accessed 2020-09-23]
13. Whiteley L, Brown L, Swenson R. African American Adolescents and New Media: Associations with HIV/STI risk behavior and psychosocial variables. Ethn Dis 2011;21(2):A.
14. Guilamo-Ramos V, Lee JJ, Kantor LM, Levine DS, Baum S, Johnsen J. Potential for using online and mobile education with parents and adolescents to impact sexual and reproductive health. Prev Sci 2015 Jan 13;16(1):53-60. [doi: 10.1007/s11121-014-0469-z] [Medline: 24522898]
15. Trends in teen pregnancy and childbearing. Office of Adolescent Health. 2016. URL: https://www.hhs.gov/ash/oah/adolescent-development/reproductive-health-and-teen-pregnancy/teen-pregnancy-and-childbearing/trends/index.html [accessed 2020-09-22]
16. Burns J. Adolescent males, risk behaviors and new media. In: Quallich S, Lajiness M, Mitchell K, editors. A Manual for Men's Health: A Practice Guide for APRNs and PAs. New York, NY: Springer Publishing Company; 2018:54-77.
17. Kachur R, Mesnick J, Liddon N. Adolescents, Technology, and Reducing Risk for HIV, STDs, and Pregnancy. Centers for Disease Control and Prevention. 2013. URL: https://www.cdc.gov/std/life-stages-populations/adolescents-white-paper.pdf [accessed 2020-09-22]
18. Burns JC, Arnault DS, Spencer G. Examining attitudes, norms, and perceived control: Young African American males' views of social media as a mode for condom use education. Cogent Social Sciences 2019 Mar 4;5(1). [doi: 10.1080/23311886.2019.1588840]
19. Wadham E, Green C, Debattista J, Somerset S, Sav A. New digital media interventions for sexual health promotion among young people: a systematic review. Sex Health 2019 Apr;16(2):101-123. [doi: 10.1071/SH18127] [Medline: 30819326]
20. Compendium of Evidence-Based Interventions and Practices for HIV Prevention. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/hiv/research/interventionresearch/compendium/index.html [accessed 2020-09-22]

XSL•FO
RenderX

21. Hswen Y, Bickham DS. Engaging African American teens in co-creating and disseminating social media based HIV prevention messages. J Hosp Manag Health Policy 2018 Aug;2:41-41. [doi: 10.21037/jhmhp.2018.07.06]
22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
23. Armstrong R, Hall BJ, Doyle J, Waters E. Cochrane Update. 'Scoping the scope' of a cochrane review. J Public Health (Oxf) 2011 Mar;33(1):147-150. [doi: 10.1093/pubmed/fdr015] [Medline: 21345890]
24. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 01;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
25. QuickFacts: United States. U.S. Census Bureau. 2018. URL: https://www.census.gov/quickfacts/fact/table/US/PST045218 [accessed 2020-09-23]
26. Veinot T, Campbell T, Kruger D, Grodzinski A, Franzen S. Drama and danger: the opportunities and challenges of promoting youth sexual health through online social networks. AMIA Annu Symp Proc 2011;2011:1436-1445 [FREE Full text] [Medline: 22195207]
27. Bull SS, Levine DK, Black SR, Schmiege SJ, Santelli J. Social media-delivered sexual health intervention: a cluster randomized controlled trial. Am J Prev Med 2012 Nov;43(5):467-474 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.022] [Medline: 23079168]
28. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, et al. Interventions using new digital media to improve adolescent sexual health: a systematic review. J Adolesc Health 2012 Dec;51(6):535-543. [doi: 10.1016/j.jadohealth.2012.03.014] [Medline: 23174462]
29. Veinot TC, Campbell TR, Kruger DJ, Grodzinski A. A question of trust: user-centered design requirements for an informatics intervention to promote the sexual health of African-American youth. J Am Med Inform Assoc 2013 Jul;20(4):758-765 [FREE Full text] [doi: 10.1136/amiajnl-2012-001361] [Medline: 23512830]
30. Shegog R, Peskin MF, Markham C, Thiel M, Karny E, Addy RC, et al. It's Your Game-Tech: Toward Sexual Health in the Digital Age. Creat Educ 2014 Aug;5(15):1428-1447 [FREE Full text] [doi: 10.4236/ce.2014.515161] [Medline: 25705561]
31. Cordova D, Bauermeister JA, Fessler K, Delva J, Nelson A, Nurenberg R, et al. A Community-Engaged Approach to Developing an mHealth HIV/STI and Drug Abuse Preventive Intervention for Primary Care: A Qualitative Study. JMIR Mhealth Uhealth 2015 Dec;3(4):e106 [FREE Full text] [doi: 10.2196/mhealth.4620] [Medline: 26685288]
32. Dolcini MM, Warren J, Towner SL, Catania JA, Harper GW. Information Age: Do Urban African American Youth Find Sexual Health Information Online? Sex Res Social Policy 2015 Jun 01;12(2):110-114 [FREE Full text] [doi: 10.1007/s13178-014-0174-5] [Medline: 26052264]
33. Muessig KE, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman LB. A systematic review of recent smartphone, Internet and Web 2.0 interventions to address the HIV continuum of care. Curr HIV/AIDS Rep 2015 Mar;12(1):173-190 [FREE Full text] [doi: 10.1007/s11904-014-0239-3] [Medline: 25626718]
34. Taggart T, Grewe ME, Conserve DF, Gliwa C, Roman IM. Social Media and HIV: A Systematic Review of Uses of Social Media in HIV Communication. J Med Internet Res 2015 Nov;17(11):e248 [FREE Full text] [doi: 10.2196/jmir.4387] [Medline: 26525289]
35. Condran B, Gahagan J, Isfeld-Kiely H. A scoping review of social media as a platform for multi-level sexual health promotion interventions. The Canadian Journal of Human Sexuality 2017 Apr;26(1):26-37. [doi: 10.3138/cjhs.261-a1]
36. Córdova D, Mendoza Lua F, Ovadje L, Hong E, Castillo B, Salas-Wright CP. Randomized Controlled Trials of Technology-Based HIV/STI and Drug Abuse Preventive Interventions for African American and Hispanic Youth: Systematic Review. JMIR Public Health Surveill 2017 Dec 13;3(4):e96 [FREE Full text] [doi: 10.2196/publichealth.7129] [Medline: 29237577]
37. Eason S. Exposure to HIV Prevention Messages on Social Media and Behavior Change Intent. Walden University ScholarWorks. 2017. URL: http://scholarworks.waldenu.edu/dissertations [accessed 2020-09-23]
38. Fiellin LE, Hieftje KD, Pendergrass TM, Kyriakides TC, Duncan LR, Dziura JD, et al. Video Game Intervention for Sexual Risk Reduction in Minority Adolescents: Randomized Controlled Trial. J Med Internet Res 2017 Sep 18;19(9):e314. [doi: 10.2196/jmir.8148]
39. Jemmott LS, Jemmott JB, Lanier Y, Thompson C, Baker JL. Development of a Barbershop-Based HIV/STI Risk Reduction Intervention for Young Heterosexual African American Men. Health Promotion Practice 2016 Aug 20;18(1):110-118. [doi: 10.1177/1524839916662601]
40. Stevens R, Gilliard-Matthews S, Dunaev J, Todhunter-Reid A, Brawner B, Stewart J. Social Media Use and Sexual Risk Reduction Behavior Among Minority Youth. Nursing Research 2017;66(5):368-377. [doi: 10.1097/nnr.0000000000000237]
41. Graham LF, Scott L, Lopeyok E, Douglas H, Gubrium A, Buchanan D. Outreach Strategies to Recruit Low-Income African American Men to Participate in Health Promotion Programs and Research: Lessons From the Men of Color Health Awareness (MOCHA) Project. Am J Mens Health 2018 Apr 26;12(5):1307-1316. [doi: 10.1177/1557988318768602]
42. Jones K, Williams J, Sipsma H, Patil C. Adolescent and emerging adults' evaluation of a Facebook site providing sexual health education. Public Health Nurs 2018 Oct 01;36(1):11-17. [doi: 10.1111/phn.12555]

XSL•FO
RenderX

43. Brayboy LM, McCoy K, Thamotharan S, Zhu E, Gil G, Houck C. The use of technology in the sexual health education especially among minority adolescent girls in the United States. Current Opinion in Obstetrics and Gynecology 2018;30(5):305-309. [doi: 10.1097/gco.0000000000000485]
44. Fishbein M, Ajzen I. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Reading, MA: Addison-Wesley Publishing Company; 1975.
45. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
46. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
47. Levitz N, Wood E, Kantor L. The Influence of Technology Delivery Mode on Intervention Outcomes: Analysis of a Theory-Based Sexual Health Program. J Med Internet Res 2018 Aug 29;20(8):e10398. [doi: 10.2196/10398]
48. Lan C, Lightfoot A, Gere D, Taboada A, Meyer K, Harwood J, et al. Live or Virtual? Comparing Two Versions of AMP!, A Theater-Based Sexual Health Intervention for Adolescents. American Journal of Sexuality Education 2019 Apr 09;14(3):292-314. [doi: 10.1080/15546128.2019.1586270]
49. Social Media Guidelines and Best Practices. Centers for Disease Control and Prevention. 2010. URL: https://www.cdc.gov/socialmedia/tools/guidelines/pdf/textmessages.pdf [accessed 2020-09-22]
50. The Health Communicator's Social Media Toolkit. Centers for Disease Control and Prevention. 2011. URL: https://www.cdc.gov/socialmedia/tools/guidelines/pdf/socialmediatoolkit_bm.pdf [accessed 2020-09-25]
51. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory- and evidence-based health education programs. Health Educ Behav 1998 Oct;25(5):545-563. [doi: 10.1177/109019819802500502] [Medline: 9768376]
52. McKleroy VS, Galbraith JS, Cummings B, Jones P, Harshbarger C, Collins C, ADAPT Team. Adapting evidence-based behavioral interventions for new settings and target populations. AIDS Educ Prev 2006 Aug;18(4 Suppl A):59-73. [doi: 10.1521/aeap.2006.18.supp.59] [Medline: 16987089]
53. Wingood GM, DiClemente RJ. The ADAPT-ITT model: a novel method of adapting evidence-based HIV Interventions. J Acquir Immune Defic Syndr 2008 Mar 01;47 Suppl 1:S40-S46. [doi: 10.1097/QAI.0b013e3181605df1] [Medline: 18301133]
54. Alleyne B, Gaston G. Gender Disparity and HIV Risk Among Young Black Women in College: A Literature Review. Affilia 2010 Apr 29;25(2):135-145. [doi: 10.1177/0886109910364348]
55. HIV risk reduction (RR) efficacy review: efficacy criteria. Centers for Disease Control and Prevention. 2014. URL: https://www.cdc.gov/hiv/dhap/prb/prs/efficacy/rr/criteria/index.html [accessed 2020-09-22]
56. Melnyk B, Morrison-Beedy D. Intervention Research: Designing, Conducting, Analyzing, and Funding. New York, NY: Springer Publishing Company; 2012.

Abbreviations

AA: African American
AYA: adolescents and young adults
CDC: Centers for Disease Control and Prevention
mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SCT: Social Cognitive Theory
SNS: social networking site
STI: sexually transmitted disease
TPB: Theory of Planned Behavior
TRA: Theory of Reasoned Action

Edited by G Eysenbach; submitted 18.04.20; peer-reviewed by A Yee, KA Nguyen; comments to author 12.06.20; revised version received 20.07.20; accepted 03.08.20; published 05.10.20.

Please cite as:
Teadt S, Burns JC, Montgomery TM, Darbes L
African American Adolescents and Young Adults, New Media, and Sexual Health: Scoping Review
JMIR Mhealth Uhealth 2020;8(10):e19459
URL: https://mhealth.jmir.org/2020/10/e19459
doi:10.2196/19459
PMID:33016890

XSL•FO
RenderX

©Sierra Teadt, Jade C Burns, Tiffany M Montgomery, Lynae Darbes. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 05.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

The Effect of Smartphone App–Based Interventions for Patients With Hypertension: Systematic Review and Meta-Analysis

Hongxuan Xu[1], MD; Huanyu Long[1], MD

The Key Laboratory of Geriatrics, Beijing Institute of Geriatrics, Beijing Hospital, National Center of Gerontology, National Health Commission, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

Corresponding Author:
Hongxuan Xu, MD
The Key Laboratory of Geriatrics, Beijing Institute of Geriatrics, Beijing Hospital
National Center of Gerontology, National Health Commission
Institute of Geriatric Medicine, Chinese Academy of Medical Sciences
No 1 DaHua Road, Dong Dan
Beijing, 100730
China
Phone: 86 13524791942
Email: xueramber@gmail.com

Abstract

Background: Hypertension is a major cause of cardiovascular disease, which is the leading cause of premature death. People with hypertension who do not comply with recommended treatment strategies have a higher risk of heart attacks and strokes, leading to hospitalization and consequently greater health care costs. The smartphone, which is now ubiquitous, offers a convenient tool to aid in the treatment of hypertension through the use of apps targeting lifestyle management, and such app-based interventions have shown promising results. In particular, recent evidence has shown the feasibility, acceptability, and success of digital interventions in changing the behavior of people with chronic conditions.

Objective: The aim of this study was to systematically compile available evidence to determine the overall effect of smartphone apps on blood pressure control, medication adherence, and lifestyle changes for people with hypertension.

Methods: This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement guidelines. Databases were searched to identify randomized controlled trials related to the influence of an app-based intervention in people with hypertension. Data extracted from the included studies were subjected to a meta-analysis to compare the effects of the smartphone app intervention to a control.

Results: Eight studies with a total of 1657 participants fulfilled the inclusion criteria. Pooled analysis of 6 studies assessing systolic blood pressure showed a significant overall effect in favor of the smartphone intervention (weighted mean difference –2.28, 95% CI –3.90-0.66). Pooled analysis of studies assessing medication adherence demonstrated a significant effect (P<.001) in favor of the intervention group (standard mean difference 0.38, 95% CI 0.26-0.50) with low heterogeneity (I^2=0%). No difference between groups was demonstrated with respect to physical activity.

Conclusions: A smartphone intervention leads to a reduction in blood pressure and an increase in medication adherence for people with hypertension. Future research should focus on the effect of behavior coaching apps on medication adherence, lifestyle change, and blood pressure reduction.

(JMIR Mhealth Uhealth 2020;8(10):e21759) doi:10.2196/21759

KEYWORDS

hypertension; smartphone; blood pressure; mobile; lifestyle; adherence; smartphone app; medication adherence

Introduction

Hypertension, or high blood pressure, is generally defined according to a systolic blood pressure (SBP) reading above 130/140 mmHg and/or diastolic blood pressure (DBP) above 80/90 mmHg. Hypertension is a major cause of cardiovascular disease, which is the leading cause of premature death [1]. Hypertension affects approximately 244.5 million adults in China, only 15.3% of whom have their condition under control [2]. People with hypertension who do not comply with recommended treatment strategies have a higher risk of heart

XSL•FO
RenderX

attacks and strokes, leading them to be hospitalized and left with greater health care costs [3]. Although faithfully taking prescribed medication and following suggested lifestyle changes can lead to a dramatic improvement in blood pressure, few people actually follow their doctor's advice, and thus fail to control their hypertension, leading to high rates of mortality and disability from heart conditions and other vascular diseases [4].

Self-measured blood pressure (SMBP) is believed to improve medication adherence, and is now a common intervention for hypertension management [5,6]. Lifestyle changes such as dietary sodium restriction, weight loss, and aerobic exercise can substantially decrease blood pressure [7]. In addition, the smartphone, which is found everywhere, offers a convenient tool to aid in the treatment of hypertension through the use of apps targeting lifestyle management, which have been showing promising results [8,9]. Recent evidence demonstrates the feasibility, acceptability, and success of digital interventions in changing the behavior of people with chronic conditions [10-17]. Getting one's hypertension under control may involve many lifestyle and behavioral changes. We hypothesized that smartphone apps combined with regular blood pressure monitoring and digital behavior change interventions may be more effective than currently employed hypertension management strategies.

Functions of such apps include a reminder to take one's medication, tracking a biometric result, education and motivation, and individualized coaching based on measured values and nonpharmaceutical behaviors. A large number of apps for medication adherence have become available in the last few years [18].

The primary objective of this systematic review and meta-analysis was to analyze the literature to determine the effect of smartphone apps on blood pressure control, medication adherence, and lifestyle changes. Only studies using stand-alone smartphone apps were included in the meta-analysis. Smartphone interventions that are not based on an app were excluded due to not conforming to our primary objective.

Methods

Search Strategy

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement guidelines [19].

We carried out a keyword search using the terms "smartphone," "hypertension," and "randomized controlled trials." The Ovid MEDLINE, EMBASE, PubMed, and Cochrane Library databases were searched from the start date of May 14, 2020. These databases were searched using a combination of subject headings (such as Medical Subject Headings) and filters (such as "RCT") when available. We also reviewed the references of included studies to identify additional pertinent studies. We imposed no language or time restriction.

Inclusion and Exclusion Criteria

Two reviewers independently assessed the records identified from the search for eligibility. Any discrepancies were resolved by consensus. We included any randomized controlled trials comparing smartphone apps–based hypertension management versus usual care or SMBP in adult primary hypertension patients. The target population was adults (aged 18 years and above) with hypertension (as defined by the authors). The outcome had to be objectively measured blood pressure changes. We accepted any duration of intervention.

We excluded studies of patients with confounding chronic conditions such as chronic kidney disease or diabetes mellitus, and those with missing key data. A protocol was developed prior to commencing this review on PROSPERO (CRD42020140926).

Study Quality

Study quality was assessed by the two authors based on the seven domains defined by the Cochrane Collaboration tool for assessing risk of bias [20]: (1) random sequence

generation; (2) allocation concealment; (3) blinding of participants and personnel; (4) blinding of outcome assessment; (5) incomplete outcome data; (6) selective reporting; and (7) other biases, including baseline imbalance, early stopping, and bias due to vested financial interest or academic bias.

Potential publication bias across studies was assessed using a funnel plot.

Data Extraction

One author (XX) extracted all data and both authors (XX and LY) reviewed the data for accuracy. The following data were collected: (1) country, duration of the trial, date of publication; (2) numbers of individuals included, inclusion criteria, exclusion criteria; (3) intervention, concomitant intervention; and (4) systolic/diastolic blood pressure change and behavior change (physical activity and medication adherence were the only two behaviors consistently reported in the existing literature).

Data Synthesis

Meta-Analyses

Meta-analysis was performed with Revman 5.3. We used a random-effects model and calculated the weighted mean difference (WMD) to generate pooled estimates of SBP and DBP changes; a random-effects model and standard mean difference (SMD) were used to calculate the intervention effects of medication adherence and physical activity across studies. We calculated the standard deviation using an assumption of a 0.5 correlation for studies that did not report the standard deviation of the mean of change, following the Cochrane Handbook for Systematic Reviews of Interventions [20]. The I^2 statistic was used to assess the degree of statistical heterogeneity.

Blood pressure changes were divided into two subgroups based on the type of intervention.

XSL•FO
RenderX

Trial Sequential Analysis

Trial sequential analysis is a methodology that considers how much information is needed to anticipate a specific required information size [21]. We used the TSA program version 0.9.5.10 Beta (Copenhagen Trial Unit) to adjust the CIs due to sparse data and repetitive testing of cumulative data, and to calculate the required information size. If the cumulative Z-curve crosses a trial sequential monitoring boundary or enters the futility area, it can be concluded that a sufficient level of evidence may have been reached. Conversely, the conclusion evidence is insufficient if the Z-curve does not cross any boundaries. The required information size was calculated based on autogenerated empirical data per input data. We performed the trial sequential analysis at the level of an overall 5% risk of type I error and a power of 20%.

Sensitivity Analysis

We conducted a posthoc sensitivity analysis to assess the impact of the potential reporting bias of trials with small sample sizes (N<60).

Results

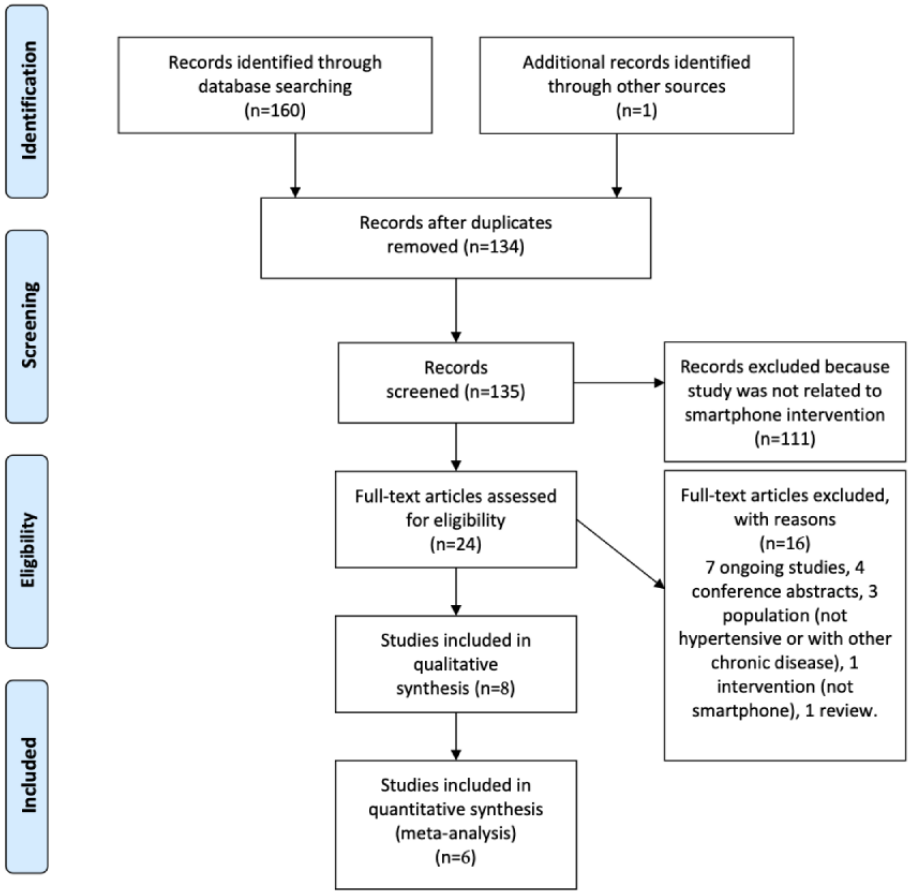
Included Studies

Eight studies [22-29] with a total of 1657 participants fulfilled the inclusion criteria; based on the results of sensitivity analysis, two studies [22,23] were excluded from the meta-analyses (Figure 1). SBP was assessed in 7 studies, DBP was assessed in 6 studies, medication adherence was assessed in 6 studies, and lifestyle changes were assessed in 2 studies. Characteristics of the included studies are summarized in Table 1.

The most common functions of mobile health apps were recording blood pressure, medication reminder, and abnormal values warning. Patient education or health recommendations were reported in 3 studies [23,25,29].

The follow-up period ranged from 6 weeks to 18 months, with a median of 6 months. Over 90% of participants were available for outcome assessment.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart of study selection.



XSL•FO
RenderX

Table 1. Characteristics of included studies.

Reference	Intervention	Control	Duration	Intervention protocol	Outcomes measured
Persell et al [29]	Hypertension Coaching App, SMBP[a] (n=144)	BP[b] tracking app, SMBP (n=152).	6 months	Record (automatically sync) BP daily for the first week, then weekly thereafter, along with reminder, encouragement, and education	BP
Ghezeljeh et al [23]	Social network self-management education (n=25)	Regular routine education (n=25)	6 weeks	Online education weekly	MA[c]
Morawski et al [26]	Medisafe app, SMBP (n=209)	None (n=202)	12 weeks	Record BP, reminder	BP
Logan et al [24]	Telemonitoring Self-Care Support System (n=55)	SMBP (n=55)	12 months	Record BP twice a week and twice in the evening	BP
Gong et al [28]	Yan Fu app, SMBP (n=225)	SMBP, record BP on paper (n=218)	6 months	Record BP at least once daily, reminder	BP, MA
Chandler et al [22]	SMASH app, SMBP (n=28)	enhanced standard care (n=26)	9 months	Record (automatically sync) BP every 3 days in the morning and evening, feedback	BP, MA
Kim et al [25]	app, SMBP, online disease management program (n=52)	usual care, online disease management program (n=43)	6 months	Record BP 3 times a week, 2 measurements per day, health recommendations, reminder	BP, MA
Márquez Contreras et al [27]	ALERHTA app (n=73)	usual care (n=75)	6/18 months	Record BP, reminder	BP, MA

[a]SMBP: self-measured blood pressure.

[b]BP: blood pressure.

[c]MA: medication adherence.

Risk of Bias

One study was judged to have a low risk of bias [24]. One study was judged to have a high risk of bias at one domain [29]. Results were unclear for the remaining 6 studies, mainly due to lack of detail of performance bias and selection bias (Multimedia Appendix 1).

Blood Pressure

Pooled outcomes of SBP (Figure 2) and DBP (Multimedia Appendix 1) were similar. SBP (WMD –2.28, 95%CI –3.90-0.66; 6 studies) with moderate heterogeneity (I^2=40%) and DBP (WMD –1.84, 95%CI –3.49 to –0.19; 5 studies) with moderate heterogeneity (I^2=54%) both showed a significant effect in favor of the intervention (*P*=.006 and *P*=.03, respectively). Blood pressure was significantly reduced in 4 studies (–2.78 mmHg), but was not significantly reduced in 2 studies that included education in the intervention (–0.33 mmHg).

Trial sequential analysis showed that the required information size of 20% power had been reached. The certainty of the evidence was high (Figure 3).

XSL•FO
RenderX

Figure 2. Meta-analysis results and forest plot for the effect of app-based interventions on improvement in systolic blood pressure.

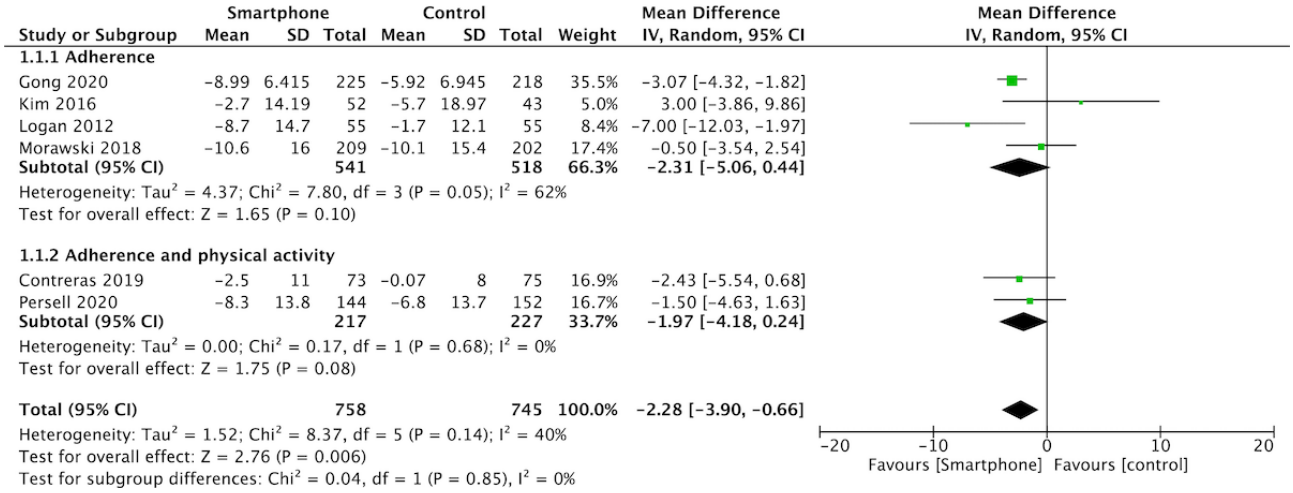
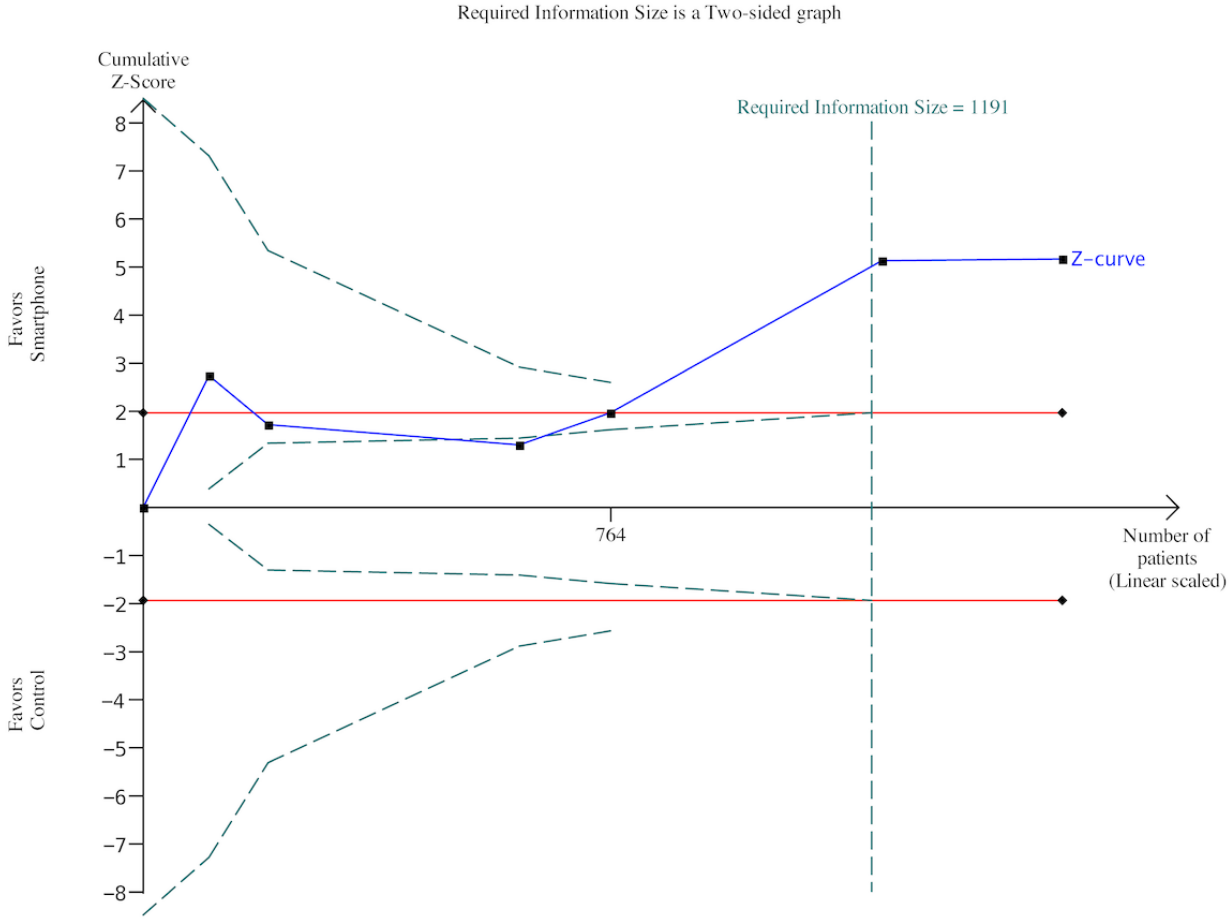


Figure 3. Trial sequential analysis of systolic blood pressure.



Medication Adherence

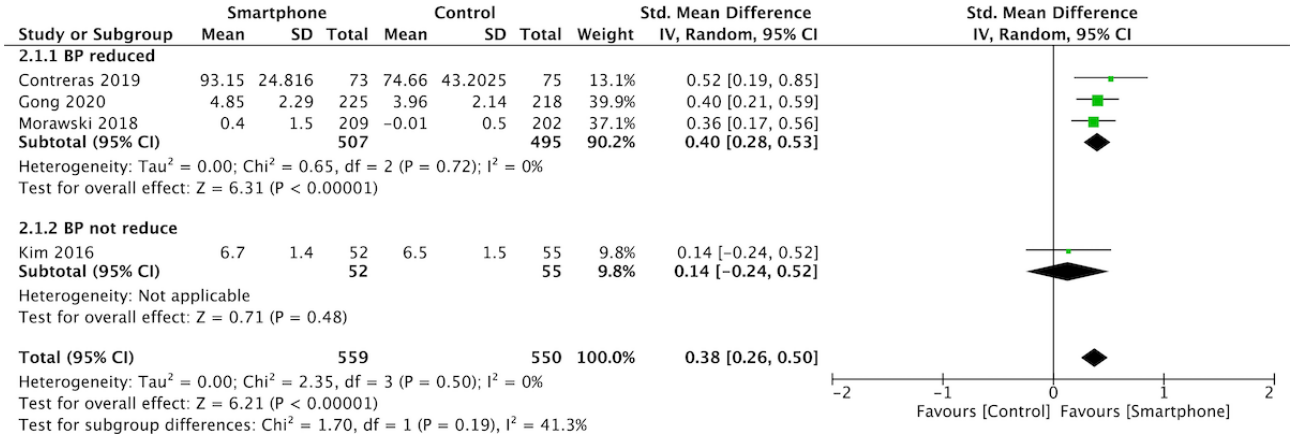
Four studies assessed medication adherence according to the Morisky Medication Adherence Scale-8 item (Figure 4) [22,25,26,28], one study determined adherence based on the pill count [25], and the other used the Hypertension SM Behavior Questionnaire [23].

Pooled analysis of medication adherence demonstrated a significant effect (P<0.0001) in favor of the intervention (SMD 0.38, 95%CI 0.26-0.50) with low heterogeneity (I^2=0%).

A test for subgroup differences showed an insignificant effect when the studies were grouped (χ^2=1.70, df=1, I^2=41.3%).

XSL•FO
RenderX

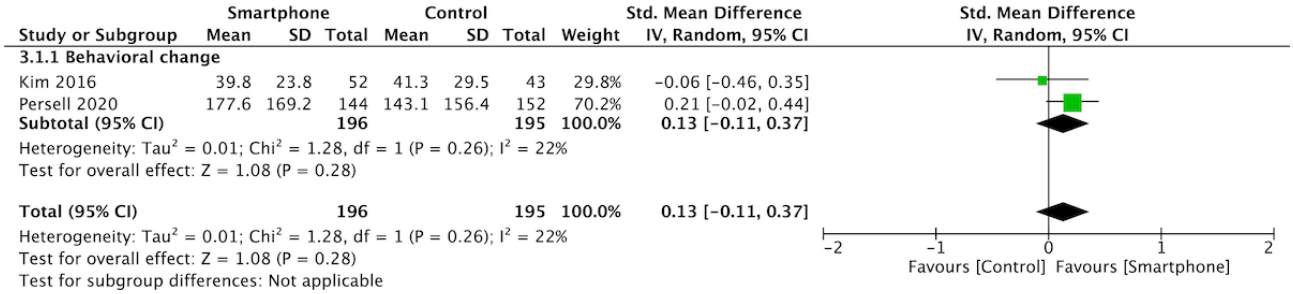
Figure 4. Meta-analysis and forest plot of the effect of the app-based intervention on medication adherence, assessed using the Morisky Medication Adherence Scale-8 item (4 studies).



Physical Activity

Two studies reported physical activity [25,29]. Pooled analysis did not show a statistically significant effect of the intervention (SMD 0.12, 95%CI –0.13-0.37, *P*=.33; Figure 5). One study showed a significant effect of reducing smoking [25] and one study showed a significant effect of confidence in controlling blood pressure [29].

Figure 5. Meta-analysis and forest plot of the effect of the app-based intervention on physical activity (2 studies).



Discussion

Principal Findings

We carried out this review to determine the effect of smartphone apps on the management of hypertension, and a pooled analysis of blood pressure and medication adherence of all studies demonstrated a significant effect in favor of the intervention. However, apps with digital behavior change interventions such as those targeting physical activity demonstrated little effect. The effect size of apps with behavioral instruction functions on blood pressure control was 9 times smaller than that of apps without these functions. Evidence of the impact of smartphone apps on physical activity was insufficient. Pooled analysis of two studies with insignificant blood pressure reduction via promoting physical activity showed a small but insignificant effect in favor of the intervention.

A previous review showed that SMBP alone is not associated with reducing blood pressure, whereas SMBP in conjunction with co-interventions significantly reduced blood pressure [6]. Another review showed that SMBP can lower blood pressure regardless of the number of hypertension-related comorbidities; however, for individuals with conditions such as obesity or stroke, SMBP should be combined with high-intensity co-interventions to effectively reduce blood pressure [30]. Moreover, individuals with chronic conditions using self-management digital interventions feel well cared for and tend to adopt a more active role in their health management [31]. Smartphone apps are able to integrate several co-interventions such as medication adherence, education, and lifestyle recommendations in one device. Our review indicated that SMBP in conjunction with a smartphone app improves medication adherence and reduces blood pressure. A reduction in SBP of 3 mmHg, as observed in intervention groups, would be expected to be associated with an 8% reduction in stroke mortality and a 5% reduction in mortality from coronary heart disease [1]. The observed magnitude of blood pressure reduction would have a significant impact on clinical practice considering the huge population worldwide suffering from hypertension.

A small minority of the apps included in our review incorporated educational functions, and the effect of these functions on blood pressure was insignificant. Available data on medication adherence of these studies was even more rare. The only study that reported medication adherence demonstrated an insignificant result. Yeung et al [32] found that in a low health literacy patient population, the use of educational tools such as flash cards and online videos significantly improved medication adherence in diabetes, hypertension, and heart failure patients. This study demonstrated that the successful education of patients regarding their medication use can significantly improve medication adherence. Medication nonadherence is estimated to cost the health care system US $5824 annually per person among patients with hypertension [33]. Low adherence to antihypertension therapy among hospitalized patients was

XSL•FO
RenderX

associated with increased costs of approximately US $3574 (95% CI US $2897-$4249) per person within a 3-year period [3]. This imposes a huge burden on the health care system and patients.

The lack of economic data such as the cost-effectiveness of smartphone-based interventions in improving medication adherence in patients with chronic health conditions highlights the need for further research to understand their role in cost savings while simultaneously improving medication adherence and health outcomes [34,35].

We have reason to believe that smartphone apps are feasible in general practice. Health care professionals should regard smartphone apps as potential tools for patients with hypertension to optimize management [36]. To better utilize these apps, several barriers need to be overcome, such as the generational difference in the propensity to use digital devices, lack of knowledge of available apps, ease of use of apps by the elderly, and privacy and data safety issues [36].

In our review, physical activity yielded an insignificant result. It is critical to discover how to usefully conceptualize and operationalize engagement with digital behavior change interventions. For instance, standardized terminology and measurement techniques will ensure more rapid advances in understanding engagement with digital behavior change interventions and developing methods to improve them [37]. The choice of app may be influenced by its immediate look and feel, "social proof," titles that appear realistic, and multicomponent features. Design features should focus on enhancing motivation, autonomy, personal relevance, and credibility [38-40].

It is vital that health care professionals and patients join together to form an effective and integrative relationship in clinical practice with mobile health apps.

Limitations

There were relatively few studies included in this meta-analysis. Therefore, the conclusions may be influenced by publication bias and should be regarded as preliminary. We did not include interventions that involved or targeted pediatric patients. The included trials were mainly conducted in North America and East Asia; thus, geographic unevenness may lead to underlying bias. Choice of the population such as excluding patients with mildly elevated blood pressure and different control interventions may also have yielded bias. With the various terms used to describe a digital intervention and its related health care, some studies may have been overlooked. The statistical findings should be interpreted cautiously for potential underlying heterogeneity.

Conclusion

A smartphone intervention leads to a reduction in blood pressure and an increase in medication adherence. Future research should discover the effect of behavior coaching apps on medication adherence, lifestyle change, and blood pressure reduction.

Acknowledgments

The authors thank AiMi Academic Services for English language editing and review services.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Risk of bias assessment for the included studies and meta-analysis for the effect of a smartphone app intervention on diastolic blood pressure (DBP) improvement.
[DOCX File , 1261 KB - mhealth_v8i10e21759_app1.docx]

References

1. Collins R, Peto R, MacMahon S, Hebert P, Fiebach NH, Eberlein KA, et al. Blood pressure, stroke, and coronary heart disease. Part 2, Short-term reductions in blood pressure: overview of randomised drug trials in their epidemiological context. Lancet 1990 Apr 07;335(8693):827-838. [doi: 10.1016/0140-6736(90)90944-z] [Medline: 1969567]
2. Wang Z, Chen Z, Zhang L, Wang X, Hao G, Zhang Z, China Hypertension Survey Investigators. Status of Hypertension in China: Results From the China Hypertension Survey, 2012-2015. Circulation 2018 May 29;137(22):2344-2356. [doi: 10.1161/CIRCULATIONAHA.117.032380] [Medline: 29449338]
3. Dragomir A, Côté R, Roy L, Blais L, Lalonde L, Bérard A, et al. Impact of adherence to antihypertensive agents on clinical outcomes and hospitalization costs. Med Care 2010 May;48(5):418-425. [doi: 10.1097/MLR.0b013e3181d567bd] [Medline: 20393367]
4. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. Circulation 2019 Mar 05;139(10):e56-e528. [doi: 10.1161/CIR.0000000000000659] [Medline: 30700139]

XSL•FO
RenderX

5. Fletcher BR, Hartmann-Boyce J, Hinton L, McManus RJ. The Effect of Self-Monitoring of Blood Pressure on Medication Adherence and Lifestyle Factors: A Systematic Review and Meta-Analysis. Am J Hypertens 2015 Oct;28(10):1209-1221. [doi: 10.1093/ajh/hpv008] [Medline: 25725092]
6. Tucker KL, Sheppard JP, Stevens R, Bosworth HB, Bove A, Bray EP, et al. Self-monitoring of blood pressure in hypertension: A systematic review and individual patient data meta-analysis. PLoS Med 2017 Sep;14(9):e1002389 [FREE Full text] [doi: 10.1371/journal.pmed.1002389] [Medline: 28926573]
7. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension 2018 Jun;71(6):e13-e115. [doi: 10.1161/HYP.0000000000000065] [Medline: 29133356]
8. Marko KI, Ganju N, Krapf JM, Gaba ND, Brown JA, Benham JJ, et al. A Mobile Prenatal Care App to Reduce In-Person Visits: Prospective Controlled Trial. JMIR Mhealth Uhealth 2019 May 01;7(5):e10520 [FREE Full text] [doi: 10.2196/10520] [Medline: 31042154]
9. Ek A, Alexandrou C, Söderström E, Bergman P, Delisle Nyström C, Direito A, et al. Effectiveness of a 3-Month Mobile Phone-Based Behavior Change Program on Active Transportation and Physical Activity in Adults: Randomized Controlled Trial. JMIR Mhealth Uhealth 2020 Jun 08;8(6):e18531 [FREE Full text] [doi: 10.2196/18531] [Medline: 32510462]
10. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
11. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile Telephone Text Messaging for Medication Adherence in Chronic Disease: A Meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
12. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
13. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text Messaging and Mobile Phone Apps as Interventions to Improve Adherence in Adolescents With Chronic Health Conditions: A Systematic Review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]
14. Badawy SM, Kuhns LM. Texting and Mobile Phone App Interventions for Improving Adherence to Preventive Behavior in Adolescents: A Systematic Review. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e50 [FREE Full text] [doi: 10.2196/mhealth.6837] [Medline: 28428157]
15. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-Centered eHealth Interventions for Children, Adolescents, and Adults With Sickle Cell Disease: Systematic Review. J Med Internet Res 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]
16. Radovic A, Badawy SM. Technology Use for Adolescent Health and Wellness. Pediatrics 2020 May;145(Suppl 2):S186-S194. [doi: 10.1542/peds.2019-2056G] [Medline: 32358210]
17. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: A systematic review of interventions across the cancer continuum. Psychooncology 2020 Jan;29(1):17-37. [doi: 10.1002/pon.5280] [Medline: 31692183]
18. Dayer LE, Shilling R, Van Valkenburg M, Martin BC, Gubbins PO, Hadden K, et al. Assessing the Medication Adherence App Marketplace From the Health Professional and Consumer Vantage Points. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e45 [FREE Full text] [doi: 10.2196/mhealth.6582] [Medline: 28428169]
19. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
20. Boutron I, Page MJ, Higgins JPT, Altman DG, Lundh A, Hróbjartsson A. Considering bias and conflicts of interest among the included studies. In: Higgins J, Green S, editors. Cochrane Handbook For Systematic Reviews Of Interventions. Hoboken, NJ: Wiley-Blackwell; 2019.
21. Thorlund K, Anema A, Mills E. Interpreting meta-analysis according to the adequacy of sample size. An example using isoniazid chemoprophylaxis for tuberculosis in purified protein derivative negative HIV-infected individuals. Clin Epidemiol 2010 Aug 09;2:57-66 [FREE Full text] [doi: 10.2147/clep.s9242] [Medline: 20865104]
22. Chandler J, Sox L, Kellam K, Feder L, Nemeth L, Treiber F. Impact of a Culturally Tailored mHealth Medication Regimen Self-Management Program upon Blood Pressure among Hypertensive Hispanic Adults. Int J Environ Res Public Health 2019 Apr 06;16(7):1226 [FREE Full text] [doi: 10.3390/ijerph16071226] [Medline: 30959858]
23. Najafi Ghezeljeh T, Sharifian S, Nasr Isfahani M, Haghani H. Comparing the effects of education using telephone follow-up and smartphone-based social networking follow-up on self-management behaviors among patients with hypertension. Contemp Nurse 2018;54(4-5):362-373. [doi: 10.1080/10376178.2018.1441730] [Medline: 29451091]

XSL•FO
RenderX

24. Logan AG, Irvine MJ, McIsaac WJ, Tisler A, Rossos PG, Easty A, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012 Jul;60(1):51-57. [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]
25. Kim JY, Wineinger NE, Steinhubl SR. The Influence of Wireless Self-Monitoring Program on the Relationship Between Patient Activation and Health Behaviors, Medication Adherence, and Blood Pressure Levels in Hypertensive Patients: A Substudy of a Randomized Controlled Trial. J Med Internet Res 2016 Jun 22;18(6):e116 [FREE Full text] [doi: 10.2196/jmir.5429] [Medline: 27334418]
26. Morawski K, Ghazinouri R, Krumme A, Lauffenburger J, Lu Z, Durfee E, et al. Association of a Smartphone Application With Medication Adherence and Blood Pressure Control: The MedISAFE-BP Randomized Clinical Trial. JAMA Intern Med 2018 Jun 01;178(6):802-809 [FREE Full text] [doi: 10.1001/jamainternmed.2018.0447] [Medline: 29710289]
27. Márquez Contreras E, Márquez Rivero S, Rodríguez García E, López-García-Ramos L, Carlos Pastoriza Vilas J, Baldonedo Suárez A, Compliance Group of Spanish Society of Hypertension (SEH-LELHA). Specific hypertension smartphone application to improve medication adherence in hypertension: a cluster-randomized trial. Curr Med Res Opin 2019 Jan;35(1):167-173. [doi: 10.1080/03007995.2018.1549026] [Medline: 30431384]
28. Gong K, Yan Y, Li Y, Du J, Wang J, Han Y, et al. Mobile health applications for the management of primary hypertension: A multicenter, randomized, controlled trial. Medicine (Baltimore) 2020 Apr;99(16):e19715. [doi: 10.1097/MD.0000000000019715] [Medline: 32311957]
29. Persell SD, Peprah YA, Lipiszko D, Lee JY, Li JJ, Ciolino JD, et al. Effect of Home Blood Pressure Monitoring via a Smartphone Hypertension Coaching Application or Tracking Application on Adults With Uncontrolled Hypertension: A Randomized Clinical Trial. JAMA Netw Open 2020 Mar 02;3(3):e200255 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.0255] [Medline: 32119093]
30. Sheppard JP, Tucker KL, Davison WJ, Stevens R, Aekplakorn W, Bosworth HB, et al. Self-monitoring of Blood Pressure in Patients With Hypertension-Related Multi-morbidity: Systematic Review and Individual Patient Data Meta-analysis. Am J Hypertens 2020 Mar 13;33(3):243-251 [FREE Full text] [doi: 10.1093/ajh/hpz182] [Medline: 31730171]
31. Morton K, Dennison L, May C, Murray E, Little P, McManus RJ, et al. Using digital interventions for self-management of chronic physical health conditions: A meta-ethnography review of published studies. Patient Educ Couns 2017 Apr;100(4):616-635 [FREE Full text] [doi: 10.1016/j.pec.2016.10.019] [Medline: 28029572]
32. Yeung DL, Alvarez KS, Quinones ME, Clark CA, Oliver GH, Alvarez CA, et al. J Am Pharm Assoc 2017 Jan;57(1):30-37 [FREE Full text] [doi: 10.1016/j.japh.2016.08.012] [Medline: 27816544]
33. Lloyd JT, Maresh S, Powers CA, Shrank WH, Alley DE. How Much Does Medication Nonadherence Cost the Medicare Fee-for-Service Program? Med Care 2019 Mar;57(3):218-224. [doi: 10.1097/MLR.0000000000001067] [Medline: 30676355]
34. Badawy SM, Kuhns LM. Economic Evaluation of Text-Messaging and Smartphone-Based Interventions to Improve Medication Adherence in Adolescents with Chronic Health Conditions: A Systematic Review. JMIR Mhealth Uhealth 2016 Oct 25;4(4):e121 [FREE Full text] [doi: 10.2196/mhealth.6425] [Medline: 27780795]
35. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
36. Byambasuren O, Beller E, Hoffmann T, Glasziou P. Barriers to and Facilitators of the Prescription of mHealth Apps in Australian General Practice: Qualitative Study. JMIR Mhealth Uhealth 2020 Jul 30;8(7):e17447 [FREE Full text] [doi: 10.2196/17447] [Medline: 32729839]
37. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
38. Badawy SM, Thompson AA, Liem RI. Technology Access and Smartphone App Preferences for Medication Adherence in Adolescents and Young Adults With Sickle Cell Disease. Pediatr Blood Cancer 2016 May;63(5):848-852. [doi: 10.1002/pbc.25905] [Medline: 26844685]
39. Badawy SM, Thompson AA, Kuhns LM. Medication Adherence and Technology-Based Interventions for Adolescents With Chronic Health Conditions: A Few Key Considerations. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e202 [FREE Full text] [doi: 10.2196/mhealth.8310] [Medline: 29273573]
40. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]

Abbreviations

DBP: diastolic blood pressure
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
SBP: systolic blood pressure
SMBP: self-measuring blood pressure

XSL•FO
RenderX

SMD: standard mean difference
WMD: weighted mean difference

Edited by G Eysenbach; submitted 24.06.20; peer-reviewed by S Badawy, P Cook; comments to author 09.08.20; revised version received 07.09.20; accepted 12.09.20; published 19.10.20.

Please cite as:
Xu H, Long H
The Effect of Smartphone App–Based Interventions for Patients With Hypertension: Systematic Review and Meta-Analysis
JMIR Mhealth Uhealth 2020;8(10):e21759
URL: http://mhealth.jmir.org/2020/10/e21759/
doi:10.2196/21759
PMID:33074161

©Hongxuan Xu, Huanyu Long. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 19.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using WeChat, a Chinese Social Media App, for Early Detection of the COVID-19 Outbreak in December 2019: Retrospective Study

Wenjun Wang[1], MD; Yikai Wang[1], MD; Xin Zhang[1], MD; Xiaoli Jia[1], MD; Yaping Li[1], MD; Shuangsuo Dang[1], MD, PhD

Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

Corresponding Author:
Shuangsuo Dang, MD, PhD
Second Affiliated Hospital of Xi'an Jiaotong University
157 Xiwu Road
Xi'an, 710004
China
Phone: 86 02987679688
Email: dangshuangsuo123@xjtu.edu.cn

Abstract

Background: A novel coronavirus, SARS-CoV-2, was identified in December 2019, when the first cases were reported in Wuhan, China. The once-localized outbreak has since been declared a pandemic. As of April 24, 2020, there have been 2.7 million confirmed cases and nearly 200,000 deaths. Early warning systems using new technologies should be established to prevent or mitigate such events in the future.

Objective: This study aimed to explore the possibility of detecting the SARS-CoV-2 outbreak in 2019 using social media.

Methods: WeChat Index is a data service that shows how frequently a specific keyword appears in posts, subscriptions, and search over the last 90 days on WeChat, the most popular Chinese social media app. We plotted daily WeChat Index results for keywords related to SARS-CoV-2 from November 17, 2019, to February 14, 2020.

Results: WeChat Index hits for “Feidian” (which means severe acute respiratory syndrome in Chinese) stayed at low levels until 16 days ahead of the local authority’s outbreak announcement on December 31, 2019, when the index increased significantly. The WeChat Index values persisted at relatively high levels from December 15 to 29, 2019, and rose rapidly on December 30, 2019, the day before the announcement. The WeChat Index hits also spiked for the keywords “SARS,” “coronavirus,” “novel coronavirus,” “shortness of breath,” “dyspnea,” and “diarrhea,” but these terms were not as meaningful for the early detection of the outbreak as the term “Feidian”.

Conclusions: By using retrospective infoveillance data from the WeChat Index, the SARS-CoV-2 outbreak in December 2019 could have been detected about two weeks before the outbreak announcement. WeChat may offer a new approach for the early detection of disease outbreaks.

(JMIR Mhealth Uhealth 2020;8(10):e19589) doi:10.2196/19589

KEYWORDS

novel coronavirus; SARS; SARS-CoV-2; COVID-19; social media; WeChat; early detection; surveillance; infodemiology; infoveillance

Introduction

An outbreak of pneumonia of unknown cause in Wuhan, the capital of Hubei province, China, occurred in December 2019 [1]. Shortly, the cause was identified as a novel coronavirus [1] that resembles severe acute respiratory syndrome (SARS) and it was named SARS-CoV-2 [2,3]. The outbreak has become a pandemic, with 2.7 million confirmed cases and nearly 200,000 deaths globally as of April 24, 2020 [4]. Early warning systems should be established to prevent or mitigate future disease outbreaks.

Traditional surveillance systems typically rely on clinical, virological, and microbiological data submitted by physicians and laboratories. Due to time and resource constraints, a lack of operational knowledge of reporting systems, and regulations

XSL•FO
RenderX

associated with these systems, substantial lags between an outbreak event and its report are common [5].

With the popularization of the internet and smartphones, an increasing number of people use social media (eg, Twitter and Facebook) to share information. Details of an event may have been posted about on social media for several days or even months before it was reported through health institutions and official reporting structures. Internet-based search engines are an important source for health information for people from all walks of life. Analyzing data on search behaviors provides a new approach for the detection and monitoring of diseases and symptoms. Technologies using social media, search queries, and other internet resources offer novel and economic approaches for detecting and tracking emerging diseases and such approaches (called infodemiology and infoveillance) have been successfully used in the cases of SARS [6], influenza [7], and dengue [8]. Herein, we explored whether the SARS-CoV-2 outbreak in China could have been detected earlier through data available on WeChat, a popular Chinese social media app. Internet search queries from Hubei province were also investigated.

Methods

WeChat (called Weixin in China; Tencent Inc) is the most popular social media app in China with over 1 billion monthly active users. WeChat Index, accessed on the WeChat app, is a publicly available data service that shows how frequently a specific keyword has appeared in posts, subscriptions, and search on WeChat over the previous 90 days. Using WeChat Index, we obtained daily data from November 17, 2019, to February 14, 2020, for keywords related to SARS-CoV-2, such as “SARS,” “Feidian” (SARS in Chinese), “pneumonia,” “fever,” “cough,” “shortness of breath,” “dyspnea,” “fatigue,” “stuffy nose,” “runny nose,” “diarrhea,” “coronavirus,” “novel coronavirus,” and “infection” (raw data in Multimedia Appendix 1). The corresponding Chinese words were used for all keywords except for “SARS”.

Baidu is the dominant Chinese internet search engine. Baidu Index (Baidu Inc) [9] can display how frequently a keyword has been queried over a certain time period in a given region. The keywords mentioned above were also investigated through Baidu Index for Hubei province.

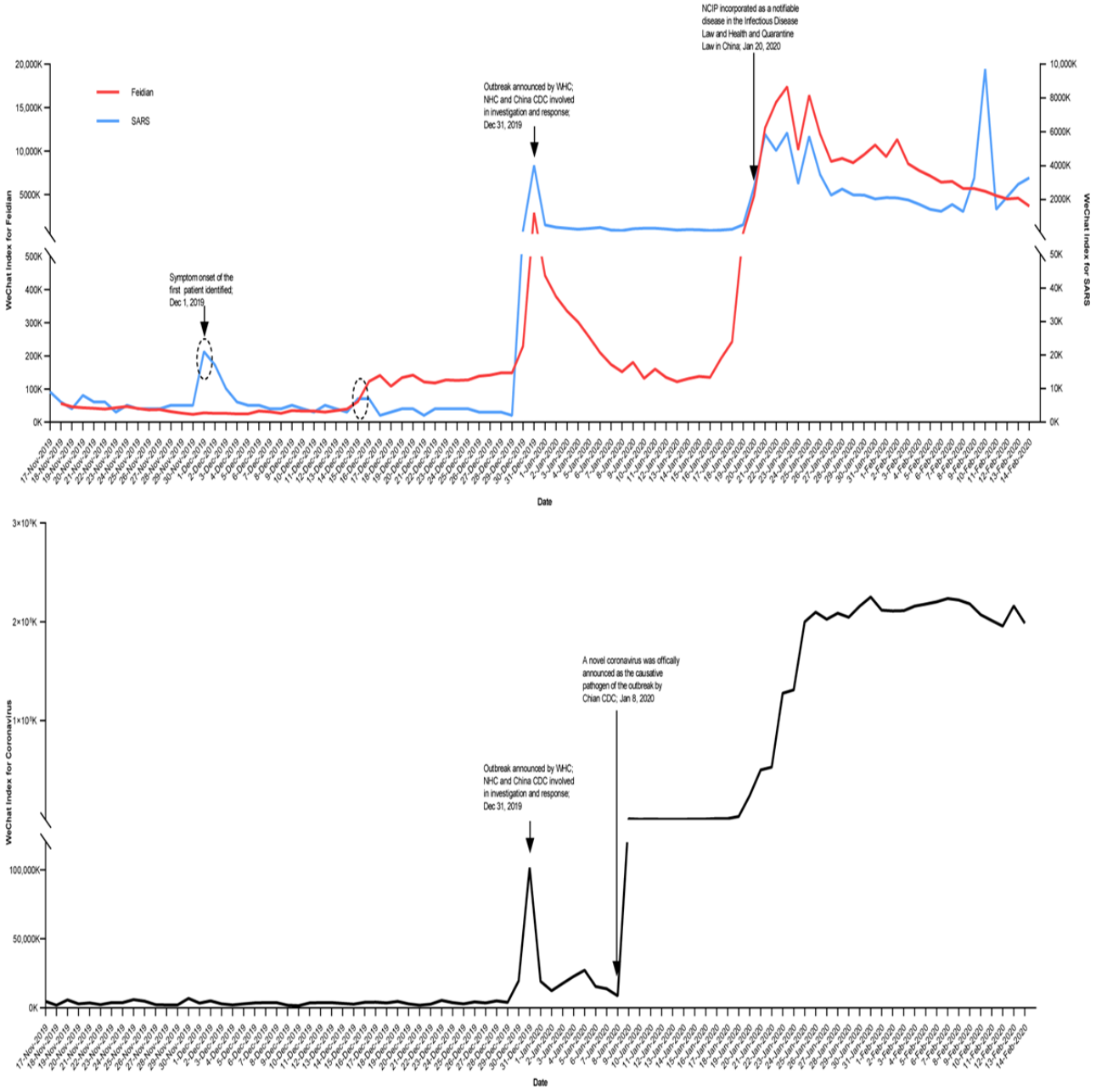
The daily data were plotted according to time for each of the keywords. As the outbreak is an isolated rather than recurrent event and the cutoff value to detect an outbreak based on social media and online search behavior is unknown, statistical analyses were not performed. The outbreak was announced by Wuhan Health Commission (WHC) on December 31, 2019; on this day, the Chinese Centers for Disease Control and Prevention (China CDC) became involved in the investigation and response [2]. If WeChat Index results for a keyword spiked or increased before the day of the outbreak announcement, the keyword was considered as a potential candidate outbreak sign [10].

Results

WeChat Index hits for “Feidian” stayed at low levels before December 15, 2019, after which they increased significantly. The WeChat Index results remained at relatively high levels until the day before the outbreak announcement, when the number of hits rose rapidly, reaching a peak on the day of the outbreak announcement (Figure 1). The WeChat Index results for “SARS” were stable, except for the first three days in December, with a peak on December 1, 2019 (Figure 1). The WeChat Index hits for “coronavirus” rose the day before the outbreak was announced, with a peak on the day of the announcement, followed by another peak after the novel coronavirus was officially announced as the causative pathogen of the outbreak by China CDC (Figure 1). From November 17, 2019, to December 30, 2019 (44 days), the WeChat Index results also spiked or increased for “novel coronavirus,” “shortness of breath,” “dyspnea,” and “diarrhea,” although these terms were not as meaningful for the early detection of the outbreak as “Feidian” (Multimedia Appendices 2 and 3).

XSL•FO
RenderX

Figure 1. WeChat Index results for the words Feidian, SARS, and coronavirus. The index results for "Feidian" began to rise on December 15, 2019 (dashed circle), persisted at relatively high levels until December 29, 2019, and rose rapidly on December 30, 2019, with a peak on December 31, 2019. The index results for "SARS" were atypical during the first three days of December, with a peak on December 1, 2019 (dashed circle). The index results for "coronavirus" began to rise on December 30, 2019, with a peak on December 31, 2019, followed by another increase on January 9, 2020. China CDC: Chinese Centers for Disease Control and Prevention; Feidian: Chinese abbreviation of severe acute respiratory syndrome; NCIP: novel coronavirus-infected pneumonia; NHC: National Health Commission of the People's Republic of China; SARS: severe acute respiratory syndrome.



The Baidu Index results for "Feidian," "SARS," "pneumonia," and "coronavirus" rose rapidly on December 30, 2019, the day before the outbreak announcement. According to Baidu Index results, no other keywords had an obvious increase from November 17, 2019, to December 30, 2019 (Multimedia Appendix 4).

Discussion

Principal Results

By exploring daily data from WeChat, a Chinese social media app, we found that the posting and search frequencies of several keywords related to SARS-CoV-2 deviated from typical frequencies ahead of the outbreak being announced in China in December 2019. Of these keywords, "Feidian" is especially worthy of attention. In 2003, the SARS outbreak caused mass panic among people in China and approximately half of the victims were health care workers [11]. Since then, Chinese physicians are on the alert for SARS as well as similar diseases [12]. If the clinical manifestations and chest images indicate viral pneumonia and several similar cases occur in a region in a short period, health care providers may think of SARS ("Feidian" in Chinese). When suspected cases are admitted to hospitals, the involved physicians may mention "Feidian" and communicate on WeChat using this word. This study found that the frequency of the word "Feidian" in WeChat began to rise

XSL•FO
RenderX

on December 15, 2019. According to publications regarding early cases of laboratory-confirmed SARS-CoV-2 infections, 5-11 patients had symptom onset by this day; the earliest onset was on December 1, 2019 [1,2]. Furthermore, the WeChat Index results for “Feidian” persisted at levels higher than those prior to December 15, 2019, and they reached a peak the day of the outbreak announcement. Altogether, the WeChat Index results for the word “Feidian” offered a strong warning sign of the developing SARS-CoV-2 outbreak. Using WeChat data in this way may enable the early detection of future outbreaks; for SARS-CoV-2, this data indicated an outbreak two weeks before the outbreak announcement.

The frequency of the term “SARS” in WeChat was unusually high from December 1 to 3, 2019, compared to the days before and after. According to Huang et al [13], the symptom onset date of the first patient identified was December 1, 2019. It is not clear whether this frequency abnormality is related to early cases. If it is, it indicates the existence of cases prior to the first reported one. The frequency of “novel coronavirus” in WeChat was abnormally high on December 11, 2019, with an index value of 400. However, its baseline value (0 or 50) was very low, so the index was sensitive to noise (Multimedia Appendix 3). The frequency of the word “coronavirus” in WeChat rose rapidly one day ahead of the outbreak announcement, so the role of this keyword was limited in the early detection of this outbreak. As for keywords related to symptoms, these symptoms are not specific to SARS-CoV-2 infection. Their increased frequency may be associated with the emergence of COVID-19, or it may be a coincidence. Although the other keywords explored in this study did not perform as well as “Feidian,” both these terms and keywords not explored in this study (eg, the names of drugs used to treat SARS-CoV-2 infection) may still prove valuable for future outbreak detection and monitoring. A previous investigation using Google Flu Trends showed that a combination of several keywords was better than a single keyword for making predictions [7].

"Infoveillance", which is the gathering and analyzing data from social media, internet search queries, and information from websites for infodemiology purposes, was proposed in 2004 by Eysenbach as a novel approach to early warning and detection of either disease outbreaks or infodemics. Infoveillance can be supplementary to traditional surveillance systems [5]. One such tool, the Global Public Health Intelligence Network (GPHIN), identified the SARS outbreak in China in 2003 more than two months earlier. In addition, they identified the outbreak of Middle East respiratory syndrome (MERS) in 2012 [6]. As far as we know, GPHIN and other established tools do not gather data from WeChat, the dominant Chinese social media app. This study shows that gathering and analyzing data from WeChat may be promising for the early detection of disease outbreaks. Considering WeChat has over 1 billion monthly active users in China, it has an advantage in detecting outbreaks within China. In addition, we found that WeChat data may provide better results than Baidu search query data because people may primarily communicate with others using WeChat [14].

Limitations

The main limitation of this study is its retrospective nature. The outbreak is a singular event. Using WeChat data for the early detection of outbreaks like this one should be further explored in the future. In addition, WeChat Index data earlier than 90 days ago is unavailable and the index calculation methodology is not public.

Conclusions

In summary, data from WeChat could have enabled the detection of the SARS-CoV-2 outbreak in 2019 about two weeks earlier than the outbreak announcement. Future studies can prospectively gather and analyze data from WeChat for the early detection of disease outbreaks in China. Tracking the source of keywords in WeChat that have atypical frequencies may become a promising approach for controlling a disease outbreak at its earliest stages.

Acknowledgments

This work was supported by the Science and Technology Research and Development Program of Shaanxi Province (2020ZDXM-SF-005).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Raw data of WeChat Index for keywords related to SARS-CoV-2.
[XLSX File (Microsoft Excel File), 20 KB - mhealth_v8i10e19589_app1.xlsx]

Multimedia Appendix 2
Keywords for which WeChat Index spiked or increased during the period from November 17 to December 30, 2019.
[DOCX File , 16 KB - mhealth_v8i10e19589_app2.docx]

Multimedia Appendix 3
WeChat Index curves for keywords related to SARS-CoV-2, other than “Feidian” and “SARS”.

XSL•FO
RenderX

[PDF File (Adobe PDF File), 2484 KB - mhealth_v8i10e19589_app3.pdf]

Multimedia Appendix 4
Baidu Index curves for keywords related to SARS-CoV-2.
[PDF File (Adobe PDF File), 2140 KB - mhealth_v8i10e19589_app4.pdf]

References

1. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 2020 Feb 22;395(10224):565-574. [doi: 10.1016/S0140-6736(20)30251-8] [Medline: 32007145]
2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med 2020 Feb 20;382(8):727-733. [doi: 10.1056/nejmoa2001017]
3. Gorbalenya AE, Baker SC, Baric RS. Severe acute respiratory syndrome-related coronavirus: The species and its viruses - a statement of the Coronavirus Study Group. BioRxiv 2020 Feb 11 [FREE Full text] [doi: 10.1101/2020.02.07.937862]
4. Coronavirus disease (COVID-2019) situation reports. World Health Orgnization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/ [accessed 2020-04-26]
5. Milinovich GJ, Williams GM, Clements ACA, Hu W. Internet-based surveillance systems for monitoring emerging infectious diseases. The Lancet Infectious Diseases 2014 Feb;14(2):160-168. [doi: 10.1016/S1473-3099(13)70244-5]
6. Dion M, AbdelMalik P, Mawudeku A. Big Data and the Global Public Health Intelligence Network (GPHIN). Can Commun Dis Rep 2015 Sep 03;41(9):209-214 [FREE Full text] [doi: 10.14745/ccdr.v41i09a02] [Medline: 29769954]
7. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
8. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. PLoS Negl Trop Dis 2011 May;5(5):e1206 [FREE Full text] [doi: 10.1371/journal.pntd.0001206] [Medline: 21647308]
9. Baidu Index. URL: https://index.baidu.com [accessed 2020-09-24]
10. Mohsin M, Hamdan A, Bakar A. Review on anomaly detection for outbreak detection. 2012 Presented at: International Conference on Information Science and Management (ICoCSIM); 2012; North Sumatra, Indonesia p. 22-28.
11. Wenzel RP, Bearman G, Edmond MB. Lessons from severe acute respiratory syndrome (SARS): implications for infection control. Arch Med Res 2005 Nov;36(6):610-616 [FREE Full text] [doi: 10.1016/j.arcmed.2005.03.040] [Medline: 16216641]
12. Zhong NS, Zeng GQ. Pandemic planning in China: applying lessons from severe acute respiratory syndrome. Respirology 2008 Mar;13 Suppl 1(s1):S33-S35. [doi: 10.1111/j.1440-1843.2008.01255.x] [Medline: 18366527]
13. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet 2020 Feb;395(10223):497-506. [doi: 10.1016/S0140-6736(20)30183-5]
14. Tu F. WeChat and civil society in China. Communication and the Public 2016 Sep 16;1(3):343-350. [doi: 10.1177/2057047316667518]

Abbreviations

China CDC: Chinese Centers for Disease Control and Prevention
GPHIN: Global Public Health Intelligence Network
MERS: Middle East respiratory syndrome
SARS: severe acute respiratory syndrome
WHC: Wuhan Health Commission

Edited by T Kool, G Eysenbach; submitted 25.04.20; peer-reviewed by MA Bahrami, E Bellei, I Idris; comments to author 08.05.20; revised version received 03.09.20; accepted 13.09.20; published 05.10.20.

Please cite as:
Wang W, Wang Y, Zhang X, Jia X, Li Y, Dang S
Using WeChat, a Chinese Social Media App, for Early Detection of the COVID-19 Outbreak in December 2019: Retrospective Study
JMIR Mhealth Uhealth 2020;8(10):e19589
URL: https://mhealth.jmir.org/2020/10/e19589
doi:10.2196/19589
PMID:32931439

XSL•FO
RenderX

©Wenjun Wang, Yikai Wang, Xin Zhang, Xiaoli Jia, Yaping Li, Shuangsuo Dang. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 27.09.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Measuring Mobility and Room Occupancy in Clinical Settings: System Development and Implementation

Gabriele Marini[1], MSc; Benjamin Tag[1], PhD; Jorge Goncalves[1], PhD; Eduardo Velloso[1], PhD; Raja Jurdak[2], PhD; Daniel Capurro[1], MD, PhD; Clare McCarthy[3], MHA; William Shearer[3], FANZCA; Vassilis Kostakos[1], PhD

[1]University of Melbourne, Melbourne, Australia

[2]Queensland University of Technology, Brisbane, Australia

[3]Northern Health, Melbourne, Australia

Corresponding Author:
Gabriele Marini, MSc
University of Melbourne
Grattan Street,
Melbourne, 3052
Australia
Phone: 61 390358966
Email: marinig@student.unimelb.edu.au

Abstract

Background: The use of location-based data in clinical settings is often limited to real-time monitoring. In this study, we aim to develop a proximity-based localization system and show how its longitudinal deployment can provide operational insights related to staff and patients' mobility and room occupancy in clinical settings. Such a streamlined data-driven approach can help in increasing the uptime of operating rooms and more broadly provide an improved understanding of facility utilization.

Objective: The aim of this study is to measure the accuracy of the system and algorithmically calculate measures of mobility and occupancy.

Methods: We developed a Bluetooth low energy, proximity-based localization system and deployed it in a hospital for 30 days. The system recorded the position of 75 people (17 patients and 55 staff) during this period. In addition, we collected ground-truth data and used them to validate system performance and accuracy. A number of analyses were conducted to estimate how people move in the hospital and where they spend their time.

Results: Using ground-truth data, we estimated the accuracy of our system to be 96%. Using mobility trace analysis, we generated occupancy rates for different rooms in the hospital occupied by both staff and patients. We were also able to measure how much time, on average, patients spend in different rooms of the hospital. Finally, using unsupervised hierarchical clustering, we showed that the system could differentiate between staff and patients without training.

Conclusions: Analysis of longitudinal, location-based data can offer rich operational insights into hospital efficiency. In particular, they allow quick and consistent assessment of new strategies and protocols and provide a quantitative way to measure their effectiveness.

(JMIR Mhealth Uhealth 2020;8(10):e19874) doi:10.2196/19874

KEYWORDS

localization; indoor; efficiency; Bluetooth; occupancy; mobility; metrics; smartphone; mobile phone

Introduction

Background

Hospitals and clinical contexts are spaces in which physical attributes are often closely linked to organizational procedures, processes, and protocols. Thus, the movement of people within a hospital can be thought of as the physical manifestation of a particular process. For instance, when a patient visits the hospital for surgery, a particular sequence is expected to be followed: admission, preparation, anesthesia, operation, and recovery. The patient progresses along this sequence by moving through the different rooms and spaces of the hospital. Similarly, staff movement is closely linked to organizational processes. Further, these rooms are highly specialized locations and serve specific clinical and operational functions.

XSL•FO
RenderX

In this paper, we demonstrate how to take advantage of this close link between physical movement and functional and organizational processes in hospitals to generate operational insights. Specifically, we show how capturing and studying the longitudinal movement of people in a hospital can provide insights into the operational characteristics and efficiency of the hospital. We show multiple analyses where long-term mobility patterns help us quantify a hospital's operational efficiency.

Hospitals are no strangers to localization technologies [1]. Indoor localization has seen significant technological improvements in recent years [2]—the relatively inexpensive Bluetooth low energy (BLE) beacons and Apple's iBeacon standard have brought indoor localization closer to mainstream use. Most of the applications in hospitals and clinical settings have focused on process mining [3] or real-time localization [4], that is, locating people or assets quickly and accurately. However, the data collected by such real-time systems are typically discarded and not accumulated for long periods. One reason for the lack of interest in longitudinal analyses is that we currently lack a movement-centered representation that is flexible enough to work with a variety of localization systems, which also allows researchers to study people's flow across indoor spaces and rooms.

Yet, outside clinical settings and health care, an increasing number of studies suggest that long-term localization data can provide meaningful insights in workplaces where efficiency is of the essence and where the flow of people and assets can be optimized by studying and analyzing their movements. For instance, construction sites aim to minimize the movement of heavy assets to reduce hazards and ensure a safer working environment, and researchers are currently working on similar projects [5]. Similarly, hospitals have an interest in reducing the downtime of operating rooms (operating rooms) as they can cost up to Aus $1500 (US $1063) per hour when in standby [6]. This means that hospitals are strongly motivated to develop policies and standards that reduce the downtime of expensive facilities such as operating rooms. However, it can be challenging to accurately describe why facilities have downtime or why patients have to wait for long periods.

Contribution

In this paper, we present a room- and movement-centered analysis of longitudinal, indoor localization data, focusing on the semantics of hospital rooms and their role in the workplace workflow. By analyzing this data, we were able to measure the occupancy levels of different rooms, understand how much time patients and staff spent in each phase of the treatment process, and visualize and quantify the movement of patients and staff over time. We present an overview of the system we developed and deployed, the challenges we faced, and the insights we generated by analyzing the longitudinal movement data of staff and patients. We tested and validated our approach in a long-term deployment in the ward of a public teaching hospital with 7 operating theaters. Previous works have looked at gathering insights from indoor localization systems in different settings: music festivals [7], museums and galleries [8], and clinical environments [9]. However, these works do not provide a long-term analysis, focusing instead on single-time events or short-term data collection.

From a technical standpoint, the contribution of our work is in the development of algorithms that can analyze longitudinal indoor localization data and generate high-level insights with minimal training and without an understanding of the specific application domain (eg, hospitals). Researchers in the field of geographic information sciences have tackled a similar challenge of abstracting indoor spaces [10] from a theoretical point of view, which has been shown to be useful in the modeling of movements in clinical settings [11]. To the best of our knowledge, this study is the first to invert the role of Bluetooth tags and anchor nodes in a longitudinal deployment. From a clinical perspective, our work demonstrates how localization systems can be used to generate operational and efficiency measures in clinical settings.

Methods

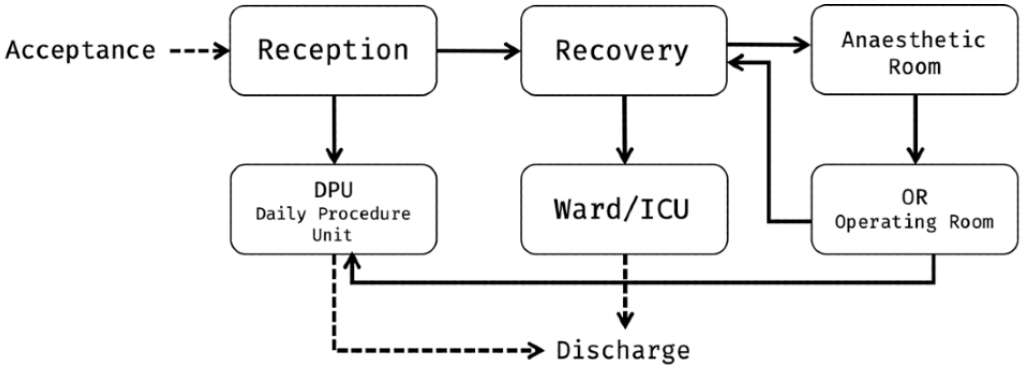
Overview

We developed a system that captures the movement of staff and patients through the operating ward of a hospital, which spans a single upper floor. Our technological implementation, which we describe next, considers a range of requirements in terms of range, precision, and power consumption. During our study, the staff members we observed, including nurses, surgeons, and technical staff, had a range of roles. We also captured patients' journeys through the various rooms of that ward, starting from the reception room to the preoperative rooms (daily process unity [DPU] for preparation) anesthetic rooms, operating rooms, and recovery rooms before discharge.

On the basis of discussions and interviews with staff members, we produced a directed graph, as presented in Figure 1, which shows the expected journey of a patient in this ward, from the moment they arrive at the reception to the time of discharge. As part of our aim to quantify the operational performance at this hospital, one of our objectives was to understand what the true patient journey looked like as opposed to the expected journey. This entailed looking at how much time each stage of the journey took and where staff members spent most of their time.

XSL•FO
RenderX

Figure 1. Graph representing the expected journey of a patient through the ward.



Indoor Localization System

Hospitals are logistically challenging environments where efficiency and patient care are of utmost importance. Therefore, tracking solutions involving bulky devices that require frequent maintenance and attention are not appropriate for such a setting. Thus, the system we deployed had to be unobtrusive, effortless to manage, and reliable. Given that rooms play an important operational role in hospitals, we were only interested in capturing location at the room level and not at a more granular scale, such as the specific coordinates of the people being tracked. Therefore, we were able to use a proximity-based indoor localization system. These systems have been around for many years [12] and have been used in a multitude of scenarios without deploying substantial infrastructure [13]. The following requirements had the most significant impact on our choice of the localization system:

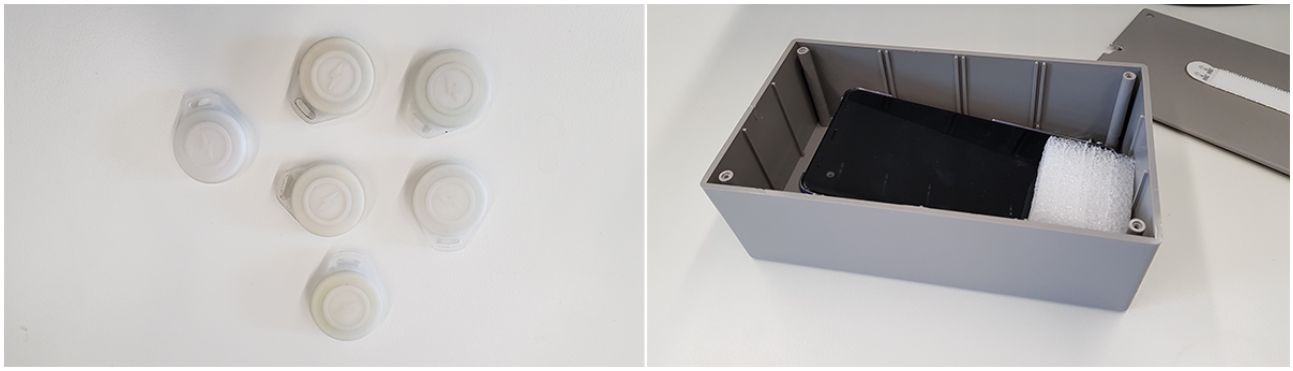
- Proximity-based localization systems require fixed hardware at a known location and mobile hardware that can infer its location when it comes close to fixed hardware. Due to the hospital's hygiene policies and their privacy requirements, nurses, surgeons, and patients may not be carrying their phones at all times. Further, patients who are about to have an operation are not able to carry any devices. Therefore, in our scenario, mobile hardware (or beacon) carried by people could not be their phone, and had to be a lightweight and noninvasive device that did not distract staff and patients.
- The equipment deployed throughout the theater needed to have a battery life that kept them running and scanning the surroundings for as long as possible without reducing their accuracy.
    - Because emergencies often take place, a power socket may suddenly become essential to staff members; therefore, equipment should not immediately turn off when unplugged.
    - Staff members often cycle through different departments in the hospital, making it hard to inform and educate them regarding our ongoing localization project. Staff members are instructed to keep the rooms safe for patients, and any new equipment may seem suspicious or new to the staff.
    - Similarly, patients might not be familiar with such devices, as they might unplug the anchor nodes or turn off their tracking device.
- We wanted to minimize the deployment of infrastructure, which can be disruptive and costly. A number of localization systems rely on Wi-Fi infrastructure, which is already present. However, this requires people to carry their own mobile devices, which is not always possible or desirable. Similarly, radio-frequency identification (RFID) systems have been used for localization, but they tend to have either a very short range of a few centimeters (passive RFID) or be heavy and noisy (active RFID) [14].

Considering the above requirements in terms of range, precision, and power consumption, we developed a proximity-based localization system that uses BLE [15]. The system consists of Android smartphones deployed as fixed anchor nodes throughout the hospital floor and RadBeacon Dot iBeacons [16] handed out to patients and staff members, as shown in Figure 2. The beacons are small and light enough to be attached to the staff badge or patient bracelet. This potentially increases adherence rates, as badges and bracelets are mandated to be worn at all times.

Typically, BLE-based localization systems rely on beacons as fixed anchor nodes and smartphones as carried devices [17]. Due to the constraints we have identified, we opted for an inverted set up where staff members and patients carry a BLE beacon, while smartphones act as fixed anchors strategically placed throughout the hospital. To the best of our knowledge, this is an approach that has not been tested before in a longitudinal study.

XSL•FO
RenderX

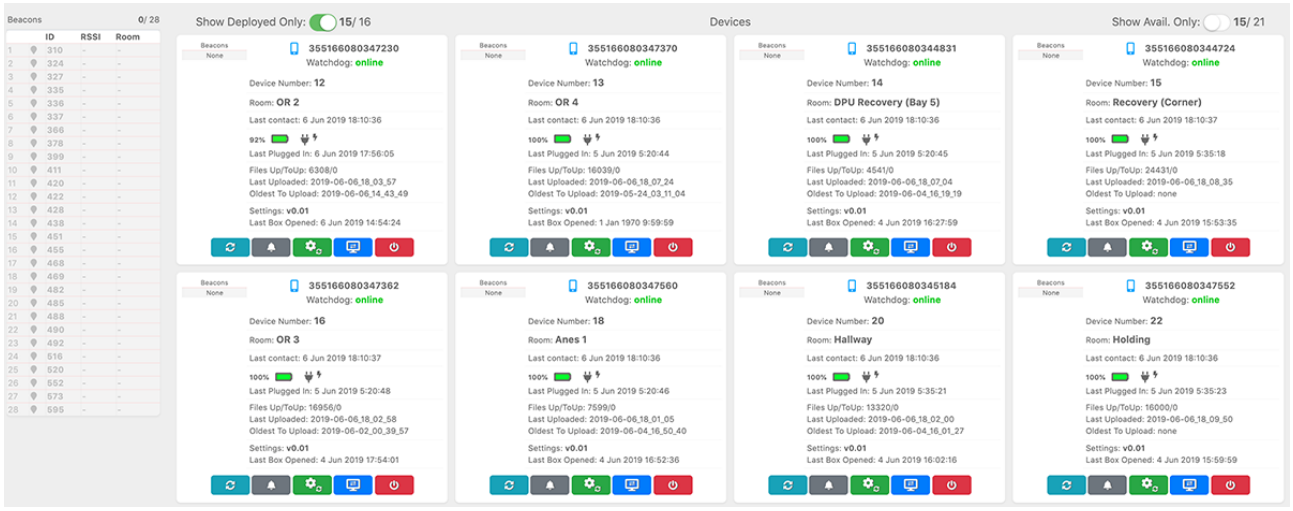
Figure 2. Beacons handed out to staff and patients; android devices that get mounted to the walls.



Software

We developed a custom Android app that continuously runs a low energy scan to detect all nearby beacons. The app continuously runs on the phone and locally stores the received signal strength indication (RSSI) readings of the received Bluetooth packets, along with the timestamp in epoch milliseconds, the minor identification number of the iBeacon emitting the packet, and the International Mobile Equipment Identity of the smartphone itself. On a daily basis, each smartphone uploaded a compressed data file containing the collected data to our server. A web-based dashboard allowed us to monitor our deployment, check the state of the smartphone network, and check for irregularities or mistakes in the smartphone and beacon labeling and configuration. Smartphones sent a brief update on their own state, including the last group of beacons they scanned, their battery level, and charging state, to the dashboard every few seconds. A screenshot of the dashboard is shown in Figure 3.

Figure 3. Snapshot of the web-based dashboard used to monitor the deployment and data collection. This was used to ensure the deployment runs as expected.



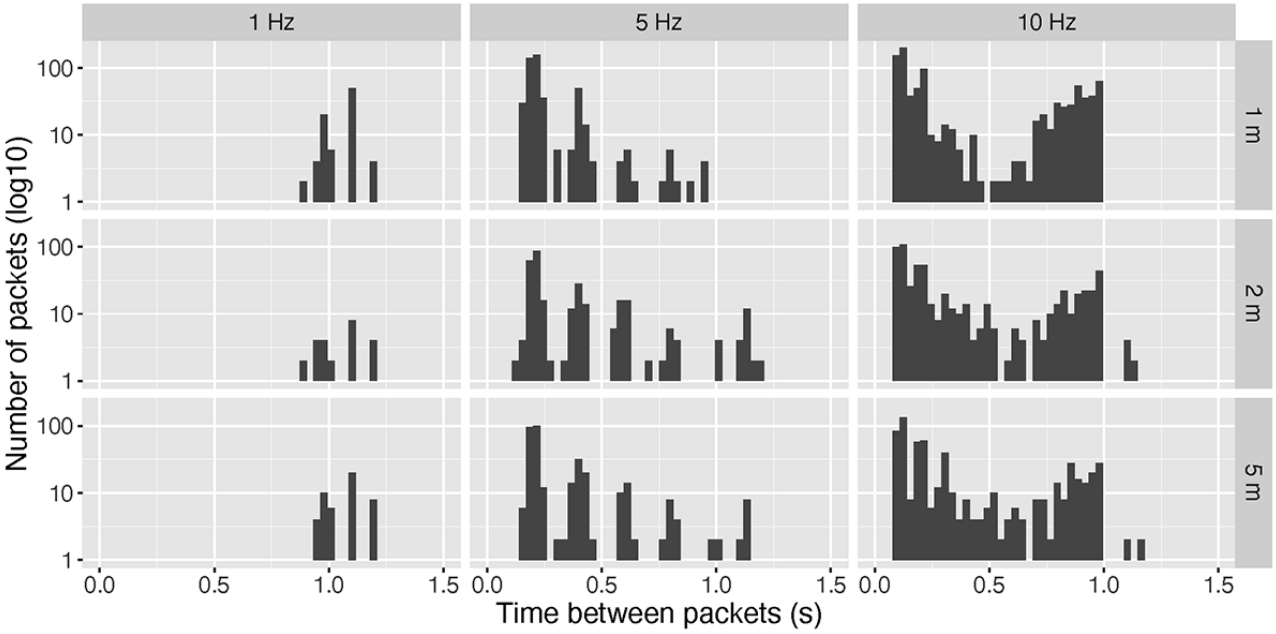
Beacon Configuration

According to the manufacturer's specifications, the beacon batteries can last between a month and more than a year, depending on the signal power and the advertisement rate to which the beacon is set. A lower power or advertisement rate allows for longer battery life, but the beacon emits its identity less frequently, meaning that in cases where a beacon is moving very fast, the anchor point can miss it. Conversely, a higher power and advertisement rate improves the chances of a beacon being detected at the cost of fairly short battery life.

Due to the way the hardware is built, the advertisement rate needs to be set before the experiment. For this reason, we conducted laboratory measurements to identify the optimal beacon settings for our scenario. We assumed beacons to be moving at walking speed and wanted them to be detectable at room-size distances, in our case, up to 5 m away. We configured beacons at different advertisement rates (1 Hz, 5 Hz, and 10 Hz) and placed them at varying distances from an anchor node (1 m, 2 m, and 5 m). In Figure 4, we show the effect of these 2 variables (advertisement rate and distance) on how the anchor node can detect those beacons (ie, the time between consecutive received packets). Effectively, this is a measurement of the number of packets lost. Given these experimental conditions, we identified that the optimum setting for our scenario, in terms of battery consumption and packet loss, is 5 Hz. At a frequency of 1 Hz, we observed an undesirable loss of packets; at a frequency of 10 Hz, we observed no noticeable performance improvement, whereas the drain on the battery doubled.

XSL•FO
RenderX

Figure 4. Laboratory measurements for different beacon advertisement rates (1 Hz, 5 Hz, and 10 Hz) and distance between the beacon and anchor node (1m, 2m, and 5m). We identified 5 Hz as the optimum setting, in terms of performance and battery consumption.



Deployment

The project was approved by the hospital's Health Human Research Ethics Committee. Each fixed anchor node comprises a plastic container, as shown in Figure 2 containing an Android HTC U11 smartphone. We deployed our anchor nodes throughout the ward (Figure 5) and plugged them into the nearest wall socket. Because our objective was to detect the presence of a beacon inside any given room, we placed 1 box in each room of interest after confirming with the staff that the box would not cause any inconvenience to them or the patients and that a nearby power source was available. In most of the rooms, the box was placed under a desk or mounted on a wall.

Figure 5. Unfiltered received signal strength indication readings and ground truth (gray line) for a single beacon moving across rooms. X-axis: time; y-axis: room identifier; color: RSSI strength (blue: weak; red: strong). RSSI: received signal strength indication.

We handed out RadBeacon iBeacons to staff members (nurses, theater technicians, and head theater technicians) who attached them to their badges, as shown in Figure 6. Staff members are mandated to carry their badges all the time, usually on their front pocket, or sometimes in their bottom or trouser pocket. When a beacon was handed out, we made a manual spreadsheet entry to link the beacon ID to the staff ID. This allowed us to link the beacon ID to staff roles during our analysis. Staff was instructed to keep the same beacon during the study and make a new data entry if they were issued a new beacon.

XSL•FO
RenderX

Figure 6. Examples of how beacons were handed out. Strapped to a staff badge or to a patient's bracelet.

Patients received their beacon along with their patient bracelets (Figure 6), as part of the hospital standard admission procedure. To identify patients, the hospital uses plastic bracelets, which strapped to their wrists during admission and cut off when they are discharged. When nurses gave a beacon to a patient, a spreadsheet entry was manually made to link the beacon ID to the patient ID. This allowed us to identify which beacon belonged to which patient during our analysis.

Ground-Truth Collection

Before we began our main deployment, we wanted to validate the data generated by the system. Ground-truth in localization refers to the true coordinates of an entity being localized. Ground-truth data may be collected by researchers independently through reliable and verifiable means, and is used to validate the correctness of the localization algorithm and fine tune its parameters. We systematically collected ground-truth data in multiple sessions. A researcher carried beacons with them and traversed the space while manually logging their precise location and exact time using a smartphone app. In total, we collected ground-truth data from 18 sessions, each lasting about 15 min and collected on a single day. All sessions were quite dynamic, meaning that the researcher moved continuously between rooms, rather than remaining static at a single location. During these sessions, we simulated realistic scenarios such as walking down the halls, entering an operating room, roaming around the surgery table, and eventually leaving and getting out of the ward and the tracked area.

Using this ground-truth data, we were able to use our system and test various filtering and analysis techniques until the trips captured by the system accurately reflected the ground truth. We present our results in the next section.

Results

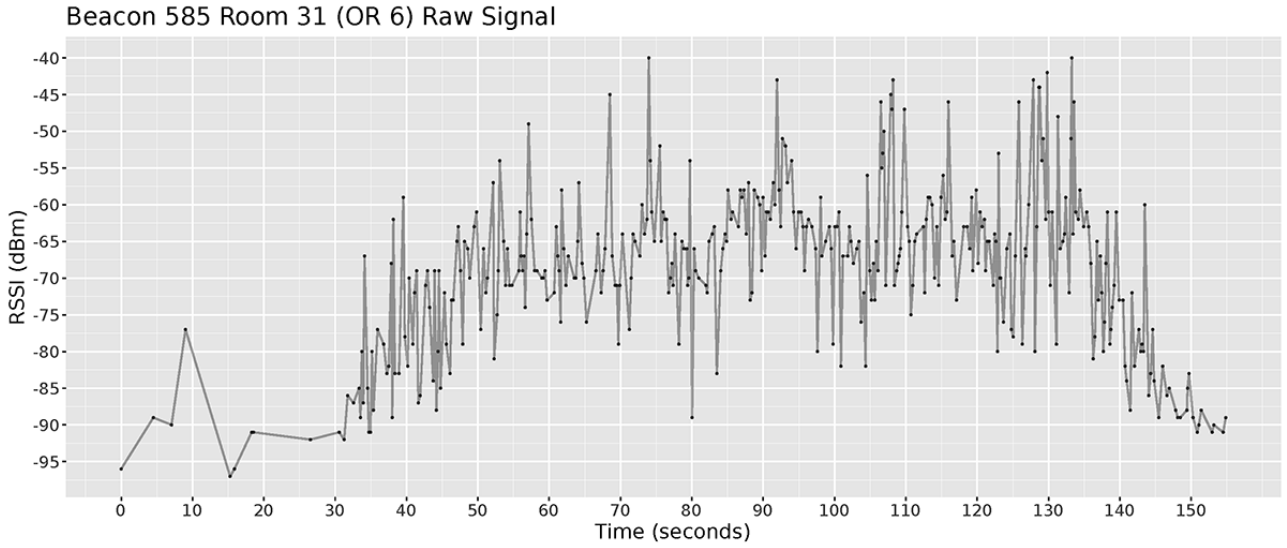
Our main deployment lasted for 30 days in September 2019. The deployment consisted of 20 anchor nodes installed in various rooms of the operating ward, including 7 operating rooms. We collected a total of 35 million packets emitted from 66 beacons handed out to 75 different people during this period. Some beacons were reused, whereas others were replaced. They were given to 17 patients by reception nurses (who also retrieved them when a patient was discharged), 15 nurses, and 7 theater tech staff.

Collected Raw Data

At the most basic level, our system captures raw signal strength data, also known as RSSI. For example, Figure 7 shows the RSSI for beacon 585 in the operating room.

XSL•FO
RenderX

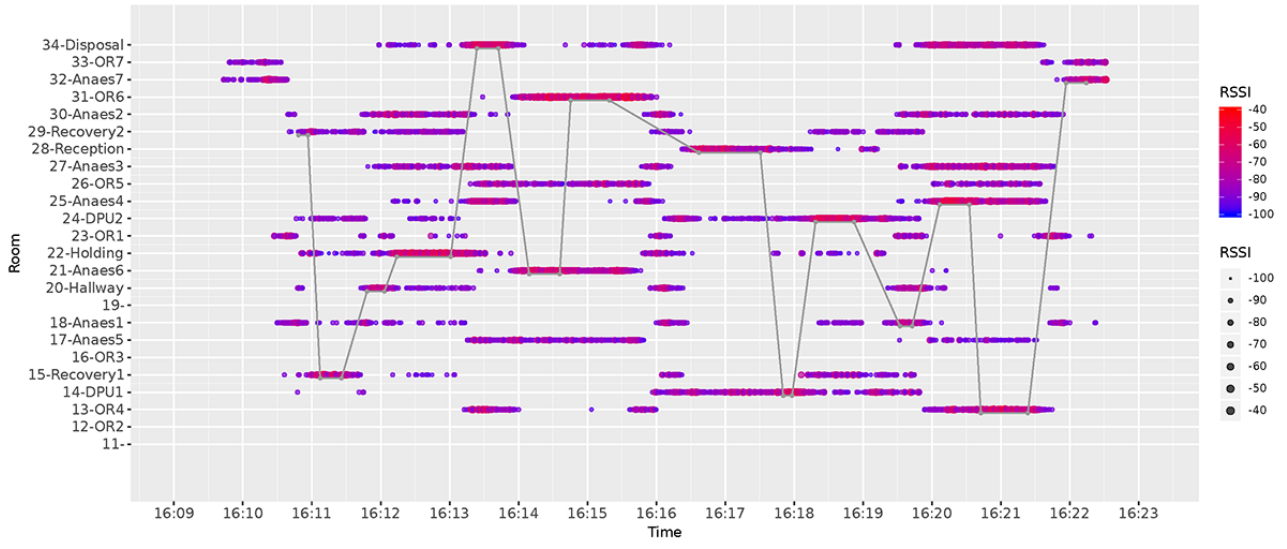
Figure 7. Received signal strength indication for a single beacon in a single room for a period of 3 min. A stronger signal suggests that the beacon was closer to the anchor node. RSSI: received signal strength indication; OR: operating room.



There is substantial literature on how to use Bluetooth for indoor localization purposes, and the reported applications range from a simple but less accurate triangulation [18] to more accurate but laborious finger-printing [19]. For our purposes, we were interested in making room-level inferences about individuals' presence because rooms are strongly linked to organizational processes. This means that we did not need to use triangulation (which requires the deployment of multiple anchor nodes), but could rely on single proximity measurements to infer the room within which an individual is. Bluetooth measurements are notoriously prone to unreliable RSSI readings [20] because of the physical nature of electromagnetic signals. The measurements can be affected by the mere presence of a human body or furniture, and especially walls and floors can attenuate the received signal thus altering the final RSSI reading. We actually used these limitations to our advantage by effectively trapping the signal within our areas of interest, and took advantage of the fact that signals traveling through walls are substantially reduced in strength.

Although walls and floors work as natural filters, a post hoc filter is still necessary because open doors or larger rooms can cause the signal to bounce off the walls and end up in an adjacent room. We visualize this phenomenon in Figure 8, where we show the raw data collected by all anchor nodes in a span of 13 min for beacon 585. In this graph, we show along the y-axis the room identifier where beacon 585 was actually detected. The colored circles indicate the detection and strength of the signal, with purple showing a weak RSSI and red indicating a strong RSSI. The graph shows a gray line corresponding to the true path (ground truth) that the beacon took through space during this 13 min. As the beacon moved between rooms (gray line), multiple anchor nodes were able to detect that beacon (colored circles). The graph also shows that while the beacon is in a particular room, the signal strength for the anchor node in that room is higher (bright red circles). Given the amount of noise in this data, we needed to apply filters to estimate the path of the beacon through the various rooms of the hospital.

Figure 8. Unfiltered RSSI readings and ground truth (grey line) for a single beacon moving across rooms. X-axis: time; y-axis: room identifier; colour: RSSI strength (blue: weak; red: strong).

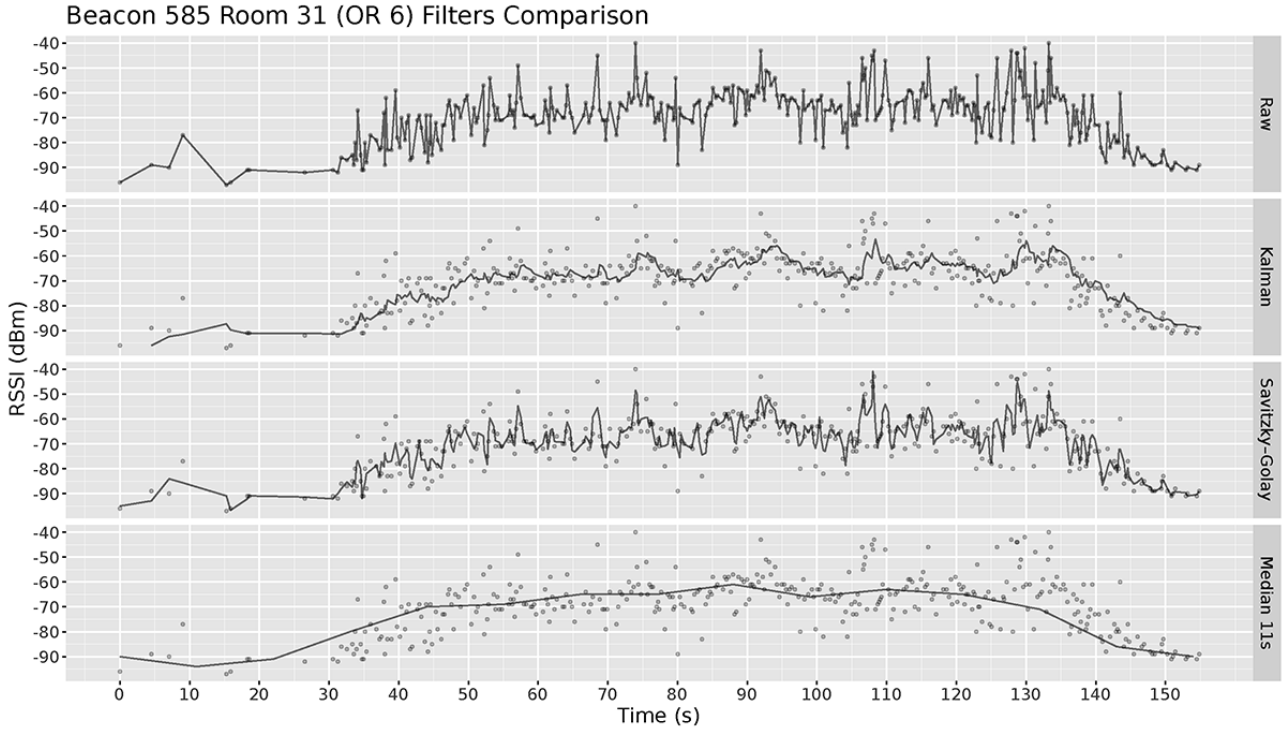


XSL•FO
RenderX

Localization Accuracy

Bluetooth-based positioning often uses Kalman filters to smooth the RSSI values and remove outlier readings. Previous work [21] shows that although it can improve accuracy, in certain cases, a median filter can work just as well, but it does not modify the data. The purpose of filtering is to remove noise so that the location of a person in a room can be confirmed at the correct time and with significant signal strength. We evaluated 3 different filters and compared their results with our collected ground-truth data. Figure 9 shows how the different filters smoothed the signal for the data presented in Figure 7. The Kalman and Savitzky–Golay filters (second and third row) were applied using their respective R package, whereas the median filter (fourth row) was manually implemented to maximize flexibility. As expected, the results in Figure 9 show that the Kalman and the Savitzky–Golay filters fit the data correctly but fail to smooth out the signal fluctuations enough to provide a clear pattern, whereas the median filter provides a smooth estimate that makes it more appropriate for comparison. In our case, we desired smoothed values because if the data from different rooms fluctuate substantially, then we were likely to have interference and infer that the beacon was moving back and forth between the 2 rooms.

Figure 9. Comparison of three approaches to filter received signal strength indication data. RSSI: received signal strength indication; OR: operating room.



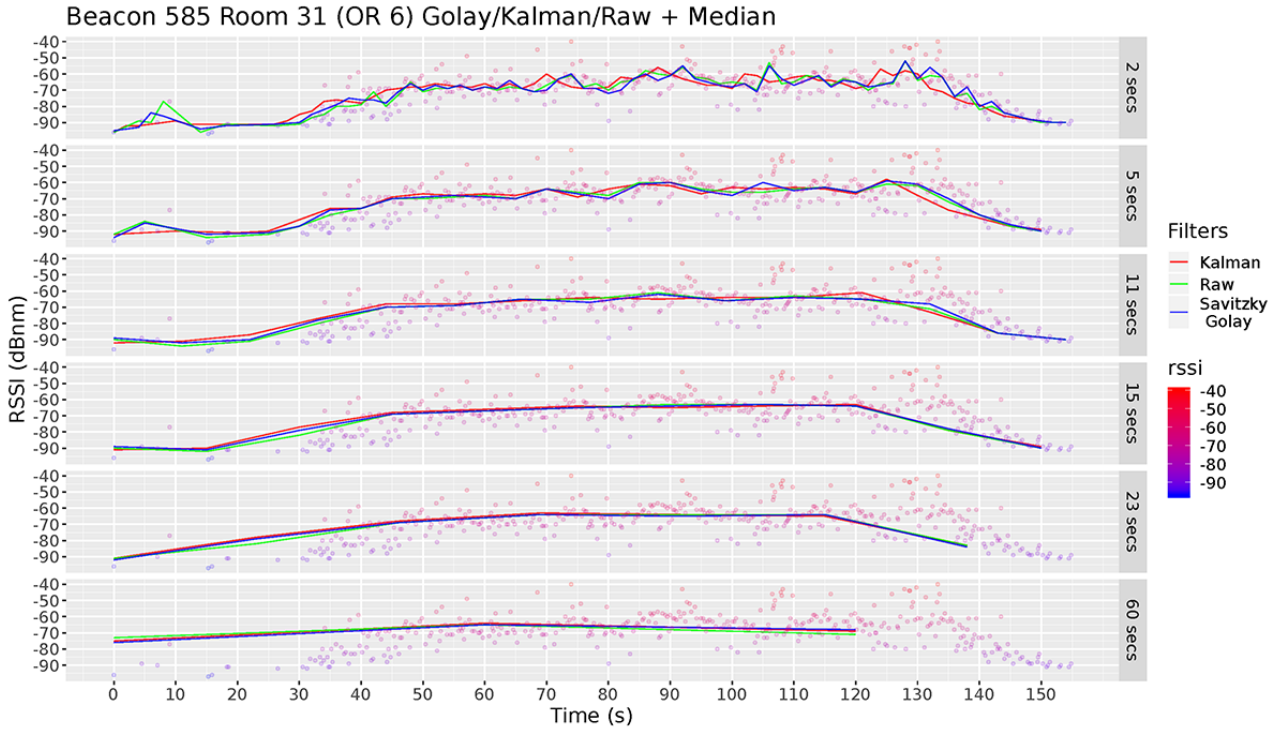
Next, we needed to determine (1) whether we should use median filtering on its own or combine it with another filter and (2) what time window our median filter should use. These decisions were made based on the analysis shown in Figure 10, where we show the performance of the median filter when applied to the raw data and prefiltered data. We also investigated the effect of different window sizes (between 2 and 60 seconds). Our results show that the addition of a Kalman or Savitzky–Golay to the median filter did not improve the accuracy. This means that our median filter was robust. We also observed that when using a smaller median time window, the resulting readings fluctuated considerably, which may lead to interference across adjacent rooms. On the other hand, a larger median time window could filter out most of the noise and short bursts of high RSSI packets, but in turn, it also lost the ability to detect granular movement; a 60-second window will inevitably miss smaller events of 10 or 20 seconds. For instance, a staff member visiting a particular room for 20 seconds, and then moving to another room may not be detected using a 60-second filtering window.

We used our ground-truth data to evaluate the accuracy of our system using a range of different window sizes. There is no universally accepted approach to measuring the accuracy of data such as ours. Therefore, we adopted an approach in which we penalized our system for inferring the wrong location at any given time, but we did not consider the magnitude of the error. To measure the system accuracy, we split the time series into intervals of 1 second and compared the ground-truth data with the reconstructed trace using a string comparison algorithm (Jaro–Winkler with a prefix scale of 0.25 and cosine distances) [22].

We observed that the optimum filtering window size was 15 seconds with the median filter, giving an accuracy of 96%. For smaller window sizes, the accuracy drops by up to 10%. We, therefore, used this window size in all our subsequent analyses, meaning that visits to a room that lasted less than 15 seconds may not be captured, but we will have stronger confidence in the visits that do indeed get captured.

XSL•FO
RenderX

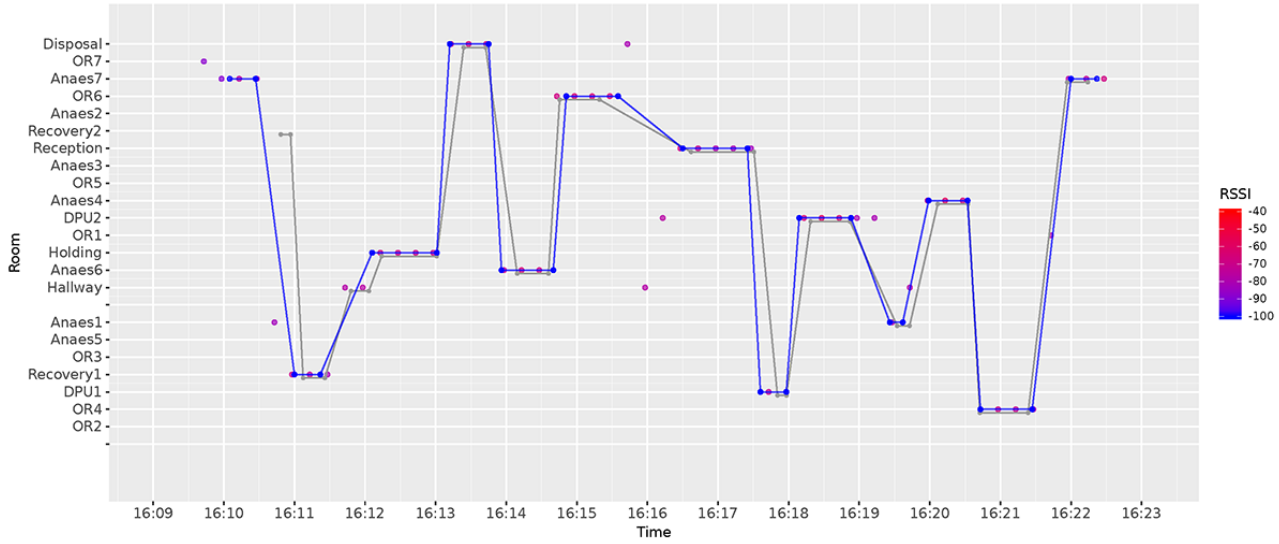
Figure 10. Comparison of using only the median filter (green line) versus using it in conjunction with other filters (red and blue). We test performance at different window sizes. RSSI: received signal strength indication.



Visualization of Movement

Figure 11 shows how our trace reconstruction process was compared with the ground truth. This figure uses the same dataset that we have used in all previous figures so far. We note that this graph shows data for 12 min during which time the position of the beacon was inferred to be in the wrong room for about 30 seconds. We also observe that our estimate sometimes appears to be slightly offset from the ground truth, suggesting that transitions between rooms were shifted by a few seconds rather than misclassified, meaning that our approach would not substantially affect the longitudinal analysis.

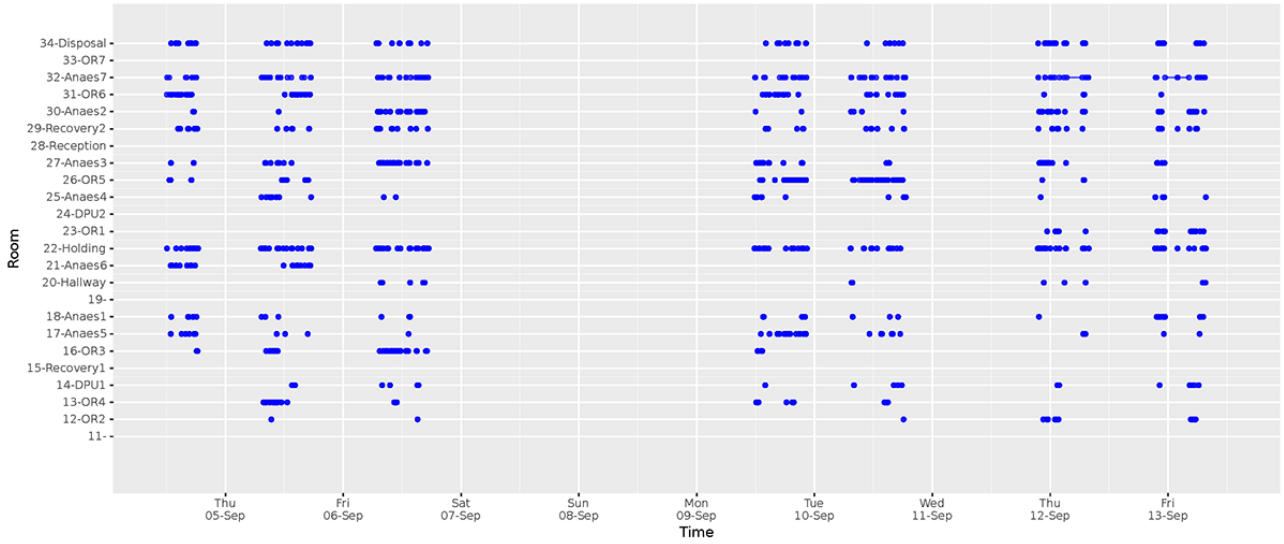
Figure 11. Trace reconstruction from our data, as compared to ground truth (grey line).



Using this trace reconstruction approach, we then turned to visualize and analyze data for longer periods. In Figure 12, we show the movement of a single nurse during a period of 10 days. From this graph, we can make a number of inferences. First, we can see that the nurse typically was at the operating ward from 9 am to 5 pm, and we can also identify the days when she was not there (eg, weekends). We can also see that she appears to spend more time in certain rooms (eg, 22-holding), whereas she rarely visited other rooms (eg, 12-OR2). We can also see differences in the patterns between days. For instance, on Monday and Tuesday, the nurse spent a lot of time in room OR5 (operating theater 5).

XSL•FO
RenderX

Figure 12. A reconstructed trace for a single nurse over a 10-day period. X-axis: date/time; y-axis: room.

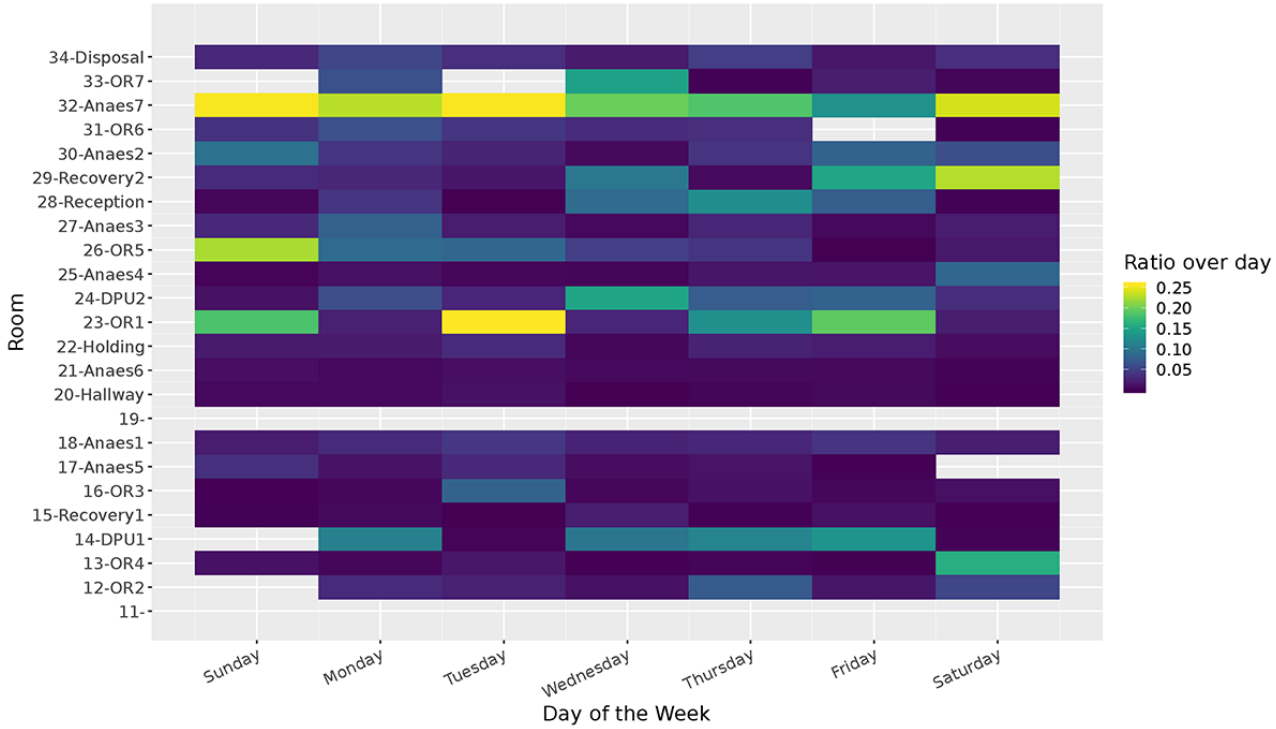


This graph demonstrates 2 things. First, it highlights the richness in the collected data and indicates the types of inferences that can be made from the data's temporal and spatial dimensions. Second, it demonstrates how visualizing movement traces does not scale, and indeed it becomes impractical to attempt to visualize all collected data during the whole month. Therefore, we must develop ways to summarize all our reconstructed trips and find meaningful ways to interpret data over long periods.

Operational Insights

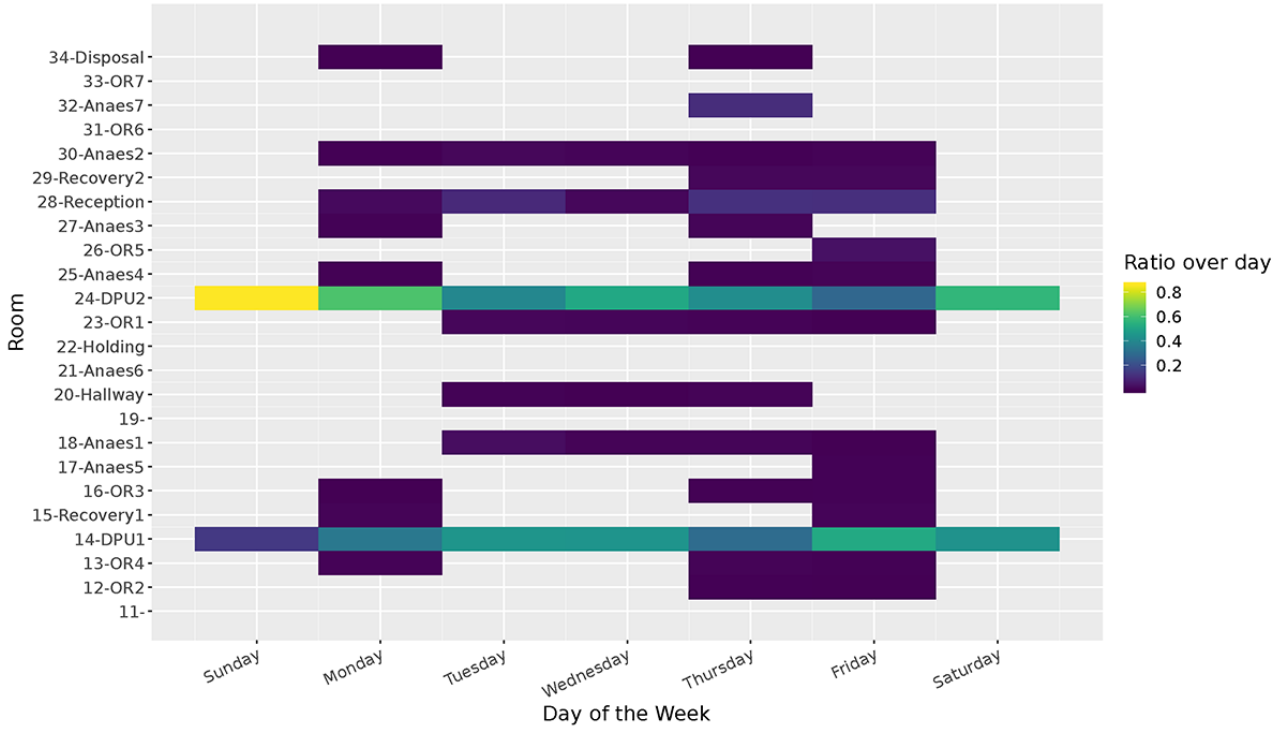
Representing the data as a time series of room-based events makes it easy to visualize patterns in the way people move through the spaces. Figure 12 shows the reconstructed traces of a single nurse for 10 days. In this figure, we can identify the daily and weekly patterns of the nurse and identify the rooms where most of the time was spent. We repeated this process, and in our analysis, we reconstructed all trips for all beacons and people across the entire duration of the study. Due to the richness of the data, it becomes inconvenient to visualize the data from all participants in this manner. Therefore, we sought alternative visualizations that could more clearly highlight patterns for the entire dataset. We achieved this by visualizing the occupancy rates of different rooms by different types of people. In Figures 13 and 14, we visualize the amount of time spent in each room by nurses and patients. This visualization confirms that there is a clear distinction between nurses and patients and the way they share their time across rooms during the week.

Figure 13. Room occupancy of nurses over a week.



XSL•FO
RenderX

Figure 14. Room occupancy of patients over a week.



All patients in this study were daily cases and therefore spent most of their time in the DPU room. Relatively speaking, only a small fraction of their time was spent in the operating room or the anesthetic room. In addition, this visualization indicates that some operating rooms were used more than others. For instance, OR1 was often used multiple times in the lapse of a week, whereas the same cannot be said of OR6.

It is important to highlight that our deployment did not recruit 100% of staff and patients. This means that it can be misleading to calculate the absolute occupancy rates of rooms. For this reason, in Table 1, we show how much time in total was spent in a given room by different types of people during our deployment. A person-day consisted of 86,400 seconds (ie, 24 hours) of presence by a single person. This is a more meaningful metric that can be used to compare across rooms. Here, we see that OR1 had the highest occupancy by staff, and the DPUs had the highest occupancy levels during our deployment. In addition, we see how much time, on average, patients spent in different rooms.

XSL•FO
RenderX

Table 1. Occupancy of different rooms, measured in person-days. Overall occupancy was made up of staff occupancy plus patient occupancy levels.

Room	Overall occupancy	Staff occupancy	Patient occupancy
28, reception	0.93	0.49	0.44
18, AR[a]1	0.53	0.47	0.06
30, AR2	0.57	0.54	0.03
27, AR3	0.42	0.40	0.02
25, AR4	0.23	0.22	0.01
17, AR5	0.23	0.19	0.03
21, AR6	0.09	0.09	0
23, OR[b]1	1.39	1.37	0.02
12, OR2	0.34	0.34	0
16, OR3	0.26	0.25	0.01
13, OR4	0.16	0.14	0.02
26, OR5	0.83	0.81	0.01
31, OR6	0.87	0.87	0
33, OR7	0.45	0.45	0
14, DPU[c]1	2.94	0.60	2.33
24, DPU2	3.43	0.65	2.78
34, disposal	0.46	0.46	0
22, holding	0.45	0.45	0
15, recovery 2	0.16	0.15	0.01
29, recovery 1	0.65	0.61	0.04
20, hallway	0.11	0.10	0.01

[a]AR: anesthetic room.

[b]OR: operating room.

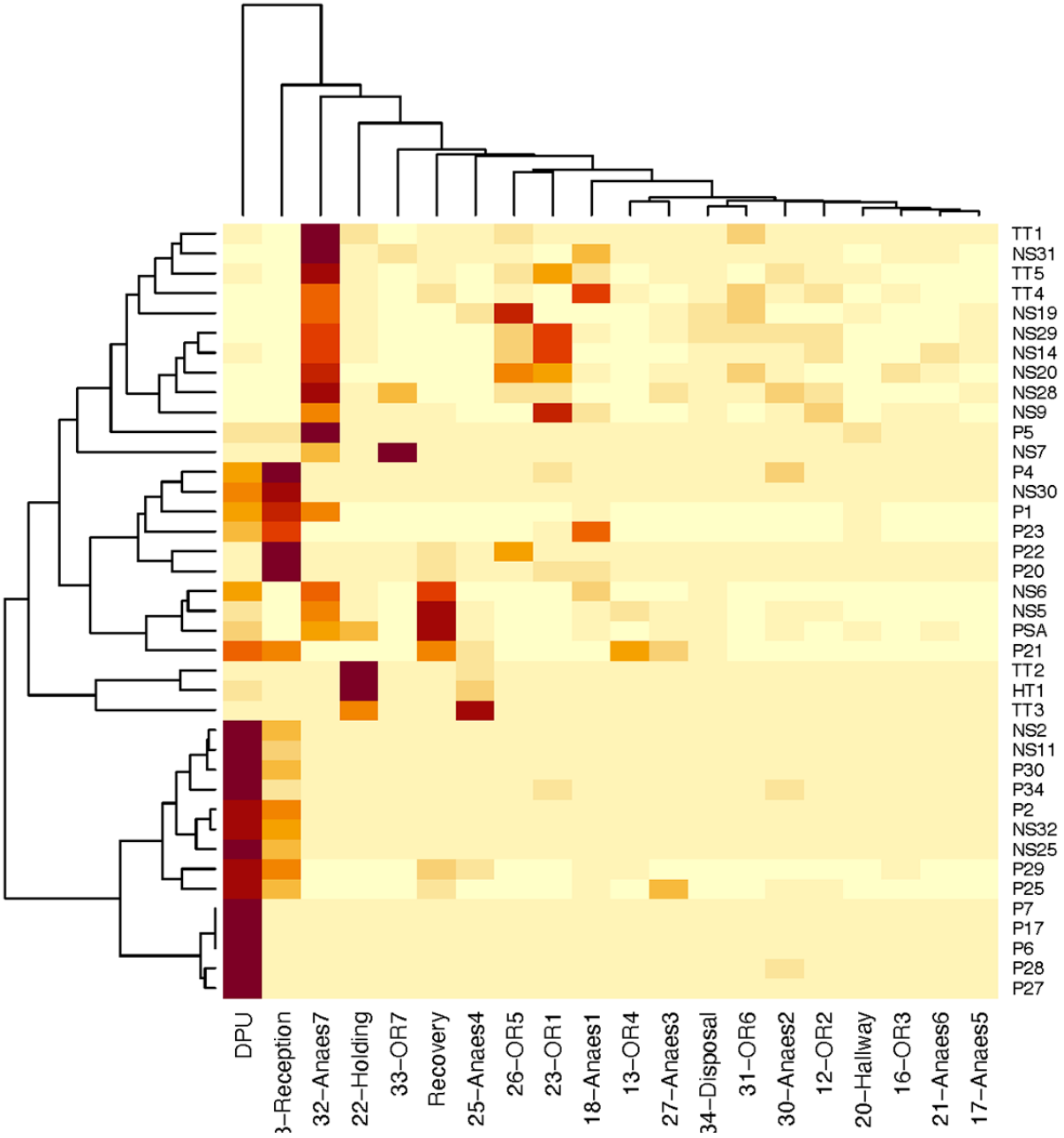
[c]DPU: daily process unity.

From our analysis, we could calculate how much time, on average, patients spent in different rooms of the hospital, which is indicative of how long each stage of the process took. For instance, we find that patients spent an average of 42 min (SD 55 min) at reception, 8 to 50 min in an anesthetic room (SD 5-28 min), 5 to 20 min in an operating room (SD 2-8 min), 6 to 9 min in recovery (SD 2-8 min), and 3 to 4 hours (SD 14-22 hours) in the DPU before they were discharged.

Finally, our results show that each person may spend a different amount of time in different rooms. Effectively, the variations to these patterns can be thought of as a *signature* of the person's role and function. To demonstrate this, we applied a very simplistic, unsupervised clustering to the data from Figures 13 and 14, and the results are shown in Figure 15. We fed the clustering algorithm an array of occupancy rates that describe how each person spent their time across different rooms. The algorithm then clusters people based on these values (shown as the dendrogram on the y-axis). We found that the algorithm, without any additional input or training from us, could identify 2 clusters of patients and 2 clusters of staff (medical staff and technical support staff).

XSL•FO
RenderX

Figure 15. Using hierarchical clustering to group people and rooms into similar clusters. The data used for clustering is the time spent per room by each person tracked. NSxx: medical staff; Pxx: patients; TTxx/HTxx: technical staff.



Discussion

Indoor Localization in Clinical Settings

The use of indoor localization technologies is not new to clinical settings. In the past, most of the applications in hospitals and clinical settings focused on real-time localization. In fact, there are multiple commercial systems or indoor localization technologies in clinical settings. For instance, IBM developed the Real-Time Asset Locator

System that relies on ultrasound provides high spatial accuracy at a substantial infrastructure investment [23]; a similar system was developed by CenTrak using Bluetooth to provide real-time localization services [24]. In addition, a number of research papers have investigated the use of Bluetooth for real-time localization [25], while also focusing on particular spaces such as operating rooms [26] or particular patients, such as newborns [27].

The popularity of real-time localization systems is increasing because they can provide a number of valuable capabilities. For instance, they allow staff to quickly call for help, raise alerts when a patient has wandered into an inappropriate room, and quickly find mobile equipment or assets when needed. However, these systems do not typically provide an aggregated overview and analysis of operational measures and do not provide enough information to make grounded judgments about the processes being followed in the hospital. For this reason, hospitals typically undergo short, intense bursts of observations to map their activities and movements. These typically last a couple of weeks and rely on manual data collection, which can be costly, time-consuming, and subject to the observer effect. In addition, some systems such as TimeCaT [28] do allow staff to indicate

XSL•FO
RenderX

their location and activity, but this requires ongoing manual data entry, which can be time-consuming for staff.

Manual observation stints remain popular because longitudinal data are useful for measuring and improving efficiency. In fact, in other process-based disciplines, efficiency has been improved using longitudinal analyses, including public transport [29], traffic routing [30], and construction [5]. Therefore, our paper's premise is that analyzing long-term mobility patterns can help us quantify a hospital's operational efficiency in multiple ways.

System Performance

System accuracy was assessed using our ground-truth data before the main deployment. During the deployment, we considered the reliability and robustness of the system. In general, we found that the system accuracy was consistently high across all ground-truth tests (96%) in terms of room-level localization [31]. During the actual deployment, there were a number of challenges faced. For example, some staff lost or misplaced their beacon and had to be issued another one. In these cases, we noted the time when the old beacon was lost and when the new beacon was issued, and these 2 data points allowed us to seamlessly analyze the movement of any staff member. However, these incidents point to a weakness of the system and indoor localization in general [32], in that it does not work for people who do not carry their tags. We also found that our algorithm seems to identify trips that are temporally shifted by a few seconds from the ground truth. This is a side effect of using a temporal window to apply the median filter [33,34], and we believe that in terms of longitudinal analysis it does not significantly affect the findings. We also found that the sampling rate of the beacons was adequate, and the beacons did not face any battery or power issues during our deployment, as expected [35].

Operational Insights

Our study demonstrates how longitudinal data from a proximity-based localization system can be filtered, aggregated, and analyzed to estimate relevant operational metrics related to mobility and occupancy rates. Specifically, we showed that the system could estimate how much time, on average, patients spent at each stage of their treatment process, which rooms they spent the most time in, and which stage seemed to take most time. With a larger sample, it would be possible to break down these results by particular conditions (eg, heart surgery and hip surgery) and identify trends or outliers within those.

Our system could also provide the same metrics for staff, and we could estimate where the staff spent their time and identify daily and weekly patterns in their behavior. We showed that different types of staff (medical vs technical support) exhibited different movement patterns, and in fact, a simple hierarchical clustering was able to identify the presence of these 2 staff groups.

These metrics can be used to assess the impact of a new strategy or protocol in the hospital. For instance, the metrics can be compared before and after a new scheduling system is deployed at the hospital. The comparison could be used to judge whether patient journeys are affected (eg, they spend less time at reception) and whether staff working patterns have substantially shifted (eg, staff spends relatively more of their time in the operating theater). In addition, it is also possible to characterize and study the behavior of other relevant subgroups of staff, such as junior medical staff. Their behavior can be aggregated and analyzed to identify how and where they spend their time, and to ensure that their time is spent effectively and with adequate support.

Localization Technologies and Representations

From a localization standpoint, our work does not break new ground in terms of accurately determining the location of people. Perhaps one unique aspect of our work is our decision to use an *inverted* deployment, whereby Bluetooth tags were given to people, and phones were glued to the wall. Typically, the opposite occurs; for example, in music festivals [7] or museums and galleries [8], Bluetooth tags are installed near items of interest while people carry phones that display information for nearby items. Our decision meant that there was minimal disruption to staff and patients, who had to simply carry a light Bluetooth tag.

However, a technical contribution of our work is our development of a movement-centered representation that is flexible enough to work with a variety of localization systems and settings. This allows our system to be easily redeployed to other settings, such as schools, universities, and galleries, with minimal changes. This is possible because we have developed a data representation that allows us to calculate all the measures we have presented in this paper, but which remains agnostic of the space and environment where the deployment takes place. Effectively, this representation allows researchers to study people's flow across abstracted rooms and spaces. Our work bears a resemblance to temporal abstraction rules [10], which can be applied to a variety of environments [36], for example, previous research [11] on a knowledge-based temporal abstraction of clinical phenomena. However, this approach relies on time-stamped clinical data, which are, in part, electronic patient records. Similarly, work on business process management [37] relies on electronic records to model corporate operations and functions, and recently, this work has been applied to clinical settings [38].

Our work, on the other hand, could capture information that was not necessarily part of electronic patient records, and which would be too costly to acquire manually. By capturing movement, we could make inferences about patient journeys and staff work practices. Similar work has recently looked at using Bluetooth to assess the levels of physical activity of people moving inside buildings and consider abstract spaces in graph form [9]. That work was confined to a handful of small trials, each lasting a few minutes, whereas our deployment lasted a month and had dozens of participants. Nevertheless, conceptually, our work used a similar approach to model space in an agnostic manner and captured people's transition between the various spaces being observed.

Limitations

The deployment we report took place at a single hospital, lasted 30 days, and did not include all patients and staff at the hospital. We expect that a different hospital or setting might pose

XSL•FO
RenderX

different challenges to the study. Still, we also point out that our algorithm analyses do not rely on any particular characteristic or aspect of the hospital itself. We also acknowledge that our data are sparse, meaning that there were many patients and staff that did not have a beacon in our study and were not observed by the system. For this reason, we argue that absolute occupancy rates were not particularly meaningful in our case per se, but we could still compare rooms in terms of how much time our participants spent there, which is a relative measure. Finally, the deployment period was not long enough to capture any potential seasonal effects or any substantial changes to the policies and protocols used at this hospital. A long-term deployment, possibly as part of an A to B study design, would generate more insights on that front.

Conclusions

In this paper, we described the development, deployment, and evaluation of an indoor localization system for hospitals and clinical settings. We demonstrated how the analysis of longitudinal data could provide operational insights regarding how people inside the hospital move, where they spend their time, and how much do various rooms get occupied. We also discussed how these metrics can be adapted to measure a number of clinical efficiency measures and how they can be used to evaluate changes to policies and protocols in these settings. As part of our future work, we plan to extend our system's algorithms to evaluate the accumulated *exposure* of clinical staff to infected patients, particularly during the COVID-19 crisis.

Acknowledgments

GM is supported by a PhD scholarship provided by the Commonwealth Scientific and Industrial Research Organisation Data61.

Conflicts of Interest

None declared.

References

1. Lorcan C, Steve N, Paddy N, Aaron Q. Sensor Aggregation and Integration in Healthcare Location Based Services. In: Pervasive Health Conference and Workshops. 2006 Presented at: PCTHEALTH'06; November 29-December 1, 2006; Innsbruck, Austria. [doi: 10.1109/pcthealth.2006.361698]
2. Davidson P, Piche R. A survey of selected indoor positioning methods for smartphones. IEEE Commun Surv Tutorials 2017;19(2):1347-1370 [FREE Full text] [doi: 10.1109/comst.2016.2637663]
3. Miclo R, Fontanili F, Marquès G, Bomert P, Lauras M. RTLS-Based Process Mining: Towards an Automatic Process Diagnosis in Healthcare. In: International Conference on Automation Science and Engineering. 2015 Presented at: CASE'15; August 24-28, 2015; Gothenburg, Sweden. [doi: 10.1109/coase.2015.7294294]
4. Kamel Boulos MN, Berry G. Real-time locating systems (RTLS) in healthcare: a condensed primer. Int J Health Geogr 2012 Jun 28;11:25 [FREE Full text] [doi: 10.1186/1476-072X-11-25] [Medline: 22741760]
5. YNOMIA and CSIRO Collaboration. URL: https://ynomia.io/technology [accessed 2020-09-22]
6. Health Victoria. URL: https://www2.health.vic.gov.au [accessed 2020-05-05]
7. Jung SH, Park SM, Sim CB. A novel of mobile commerce and information notice system for local festival content activation based on location information. J Korean Phys Soc 2015 Jun 30;18(6):780-791. [doi: 10.9717/kmms.2015.18.6.780]
8. Yoshimura Y, Sobolevsky S, Ratti C, Girardin F, Carrascal JP, Blat J, et al. An analysis of visitors' behavior in the louvre museum: a study using bluetooth data. Environ Plann B Plann Des 2014 Dec;41(6):1113-1131. [doi: 10.1068/b130047p]
9. Magistro D, Sessa S, Kingsnorth AP, Loveday A, Simeone A, Zecca M, et al. A novel algorithm for determining the contextual characteristics of movement behaviors by combining accelerometer features and wireless beacons: development and implementation. JMIR Mhealth Uhealth 2018 Apr 20;6(4):e100 [FREE Full text] [doi: 10.2196/mhealth.8516] [Medline: 29678806]
10. Jensen CS, Lu H, Yang B. Graph Model Based Indoor Tracking. In: Tenth InternationalConference on Mobile Data Management: Systems, Services and Middleware. 2009 Presented at: MDM'09; May 18-20, 2009; Taipei, Taiwan URL: https://doi.org/10.1109/MDM.2009.23 [doi: 10.1109/mdm.2009.23]
11. Shahar Y, Musen MA. Knowledge-based temporal abstraction in clinical domains. Artif Intell Med 1996 Jul;8(3):267-298. [doi: 10.1016/0933-3657(95)00036-4] [Medline: 8830925]
12. Krumm J. Ubiquitous Computing Fundamentals. New York, USA: Chapman Hall/CRC; 2009:-.
13. Deak G, Curran K, Condell J. A survey of active and passive indoor localisation systems. Comput Commun 2012 Sep;35(16):1939-1954. [doi: 10.1016/j.comcom.2012.06.004]
14. Roussos G, Kostakos V. RFID in pervasive computing: state-of-the-art and outlook. Pervas Mobile Comput 2009 Feb;5(1):110-131 [FREE Full text] [doi: 10.1016/j.pmcj.2008.11.004]
15. Acharya D, Ramezani M, Khoshelham K, Winter S. BIM-tracker: a model-based visual tracking approach for indoor localisation using a 3D building model. ISPRS J Photogrammetry Remote Sense 2019 Apr;150:157-171 [FREE Full text] [doi: 10.1016/j.isprsjprs.2019.02.014]
16. Radius Networks. URL: https://store.radiusnetworks.com/products/radbeacon-dot [accessed 2020-05-05]

XSL•FO
RenderX

17. Lin X, Ho T, Fang Z, Yen B, lai F. A Mobile Indoor Positioning System Based on iBeacon Technology. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2015 Presented at: EMBC'15; August 25-29, 2015; Milan, Italy URL: https://doi.org/10.1109/EMBC.2015.7319507 [doi: 10.1109/embc.2015.7319507]
18. Wang Y, Yang X, Zhao Y, Liu L. Bluetooth Positioning Using RSSI and Triangulation Methods. In: 10th Consumer Communications and Networking Conference. 2013 Presented at: CCNC'13; January 11-14, 2013; Las Vegas, NV, USA URL: https://doi.org/10.1109/CCNC.2013.6488558 [doi: 10.1109/ccnc.2013.6488558]
19. Faragher R, Harle R. Location fingerprinting with Bluetooth low energy beacons. IEEE J Select Areas Commun 2015 Nov;33(11):2418-2428 [FREE Full text] [doi: 10.1109/jsac.2015.2430281]
20. Faragher R, Harle R. An analysis of the accuracy of bluetooth low energy for indoor positioning applications. In: Proceedings of the 27th International Technical Meeting of the Satellite Division of the Institute of Navigation. 2014 Presented at: ION GNS'14; September 8-12, 2014; Miami, Florida.
21. Deak G, Curran K, Condell J. Evaluation of Smoothing Algorithms for a RSSI-Based Device-Free Passive Localisation. Image Processing and Communications Challenges 2 2010:469-476. [doi: 10.1007/978-3-642-16295-4_52]
22. Loo M. The stringdist package for approximate string matching. R J 2014;6(1):111. [doi: 10.32614/rj-2014-011]
23. IBM Support. URL: https://www-01.ibm.com/common/ssi/ShowDoc.wss?docURL=/common/ssi/rep_sm/3/897/ENUS5725-A93/index.html [accessed 2020-05-05]
24. Real Time Location Services. CenTrak: Accurate Hospital Real Time Location System. 2020. URL: https://centrak.com/products/real-time-location-services [accessed 2020-05-05]
25. Yang J, Wang Z, Zhang X. An iBeacon-based indoor positioning systems for hospitals. Int J Smart Home 2015 Jul 31;9(7):161-168. [doi: 10.14257/ijsh.2015.9.7.16]
26. Han G, Klinker J, Ostler D, Schneider A. Testing a Proximity-based Location Tracking System With Bluetooth Low Energy Tags for Future Use in the oR. In: 17th International Conference on E-health Networking, Application Services. 2015 Presented at: HEALTHCOM'15; October 14-17, 2015; Boston, USA. [doi: 10.1109/healthcom.2015.7454466]
27. Lu ZY. Using iBeacon for Newborns Localization in Hospitals. In: 10th International Symposium on Medical Information and Communication Technology. 2016 Presented at: ISMICT'16; March 20-23, 2016; Worcester, MA, USA. [doi: 10.1109/ismict.2016.7498906]
28. TimeCaT. URL: https://lopetegui.net/timecat/39/login/ [accessed 2020-05-05]
29. Kostakos V, Camacho T, Mantero C. Wireless Detection of End-to-end Passenger Trips on Public Transport Buses. In: Conference on Intelligent Transportation Systems. 2010 Presented at: ITS'10; September 19-22, 2010; Funchal, Madeira Island, Portugal URL: https://doi.org/10.1109/ITSC.2010.5625062 [doi: 10.1109/itsc.2010.5625062]
30. Kostakos V, Ojala T, Juntunen T. Traffic in the smart city: exploring city-wide sensing for traffic control center augmentation. IEEE Internet Comput 2013 Nov;17(6):22-29 [FREE Full text] [doi: 10.1109/mic.2013.83]
31. Athanasios IK, Panagiotis K, Michel D, Dimitri K. A BLE-based Probabilistic Room-level Localization Method. In: International Conference on Localization and GNSS. 2016 Presented at: ICL-GNSS'16; June 28-30, 2016; Barcelona, Spain. [doi: 10.1109/icl-gnss.2016.7533848]
32. Rainer M. The Challenges of Indoor Environments and Specification on Some Alternative Positioning Systems. In: 6th Workshop on Positioning, Navigation and Communication. 2009 Presented at: WPNC'09; March 19, 2009; Hannover, Germany. [doi: 10.1109/wpnc.2009.4907800]
33. Brownrigg DR. The weighted median filter. Commun ACM 1984 Aug;27(8):807-818. [doi: 10.1145/358198.358222]
34. Jokinen P, Tarhio J, Ukkonen E. A comparison of approximate string matching algorithms. In: Software: Practice and Experience. New York, USA: Routedge; 1996:1439-1458.
35. Artem D, Steve H, Stuart T, Joshua S. Power Consumption Analysis of Bluetooth Low Energy, Zigbee and ANT Sensor Nodes in a Cyclic Sleep Scenario. In: IEEE International Wireless Symposium. 2013 Presented at: IWS'13; April 13-18, 2013; Beijing, China. [doi: 10.1109/ieee-iws.2013.6616827]
36. Yaeli A, Bak P, Feigenblat G, Nadler S, Roitman H, Saadoun G, et al. Understanding customer behavior using indoor location analysis and visualization. IBM J Res Dev 2014 Sep;58(5/6):3:1-312 [FREE Full text] [doi: 10.1147/jrd.2014.2337552]
37. Dumas M, La Rosa M, Mendling J, Reijers HA. 2013. Fundamentals of business process management. Vol. 1. Springer 2013:-. [doi: 10.1007/978-3-642-33143-5]
38. Andrews R, Suriadi S, Wynn MT. Comparing static and dynamic aspects of patient flows via process model visualisations. QUT ePrints. 2016. URL: https://eprints.qut.edu.au/102848/ [accessed 2020-09-30]

Abbreviations

BLE: Bluetooth low energy
DPU: daily process unity
RFID: radio-frequency identification
RSSI: received signal strength indication

XSL•FO
RenderX

Edited by G Eysenbach; submitted 05.05.20; peer-reviewed by Q Fang, W Zhang, S Amini; comments to author 02.08.20; revised version received 18.08.20; accepted 14.09.20; published 27.10.20.

Please cite as:
Marini G, Tag B, Goncalves J, Velloso E, Jurdak R, Capurro D, McCarthy C, Shearer W, Kostakos V
Measuring Mobility and Room Occupancy in Clinical Settings: System Development and Implementation
JMIR Mhealth Uhealth 2020;8(10):e19874
URL: http://mhealth.jmir.org/2020/10/e19874/
doi:10.2196/19874
PMID:33107838

©Gabriele Marini, Benjamin Tag, Jorge Goncalves, Eduardo Velloso, Raja Jurdak, Daniel Capurro, Clare McCarthy, William Shearer, Vassilis Kostakos. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 27.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Feasibility and Acceptability of Wearable Sleep Electroencephalogram Device Use in Adolescents: Observational Study

Jessica R Lunsford-Avery[1], PhD; Casey Keller[1], MA; Scott H Kollins[1], PhD; Andrew D Krystal[2], MD; Leah Jackson[1], MS; Matthew M Engelhard[1], MD, PhD

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

[2]Departments of Psychiatry and Neurology, University of California San Francisco School of Medicine, San Francisco, CA, United States

Corresponding Author:
Jessica R Lunsford-Avery, PhD
Department of Psychiatry and Behavioral Sciences
Duke University School of Medicine
2608 Erwin Rd Suite 300
Durham, NC,
United States
Phone: 1 919 681 0035
Email: jessica.r.avery@duke.edu

Abstract

Background: Adolescence is an important life stage for the development of healthy behaviors, which have a long-lasting impact on health across the lifespan. Sleep undergoes significant changes during adolescence and is linked to physical and psychiatric health; however, sleep is rarely assessed in routine health care settings. Wearable sleep electroencephalogram (EEG) devices may represent user-friendly methods for assessing sleep among adolescents, but no studies to date have examined the feasibility and acceptability of sleep EEG wearables in this age group.

Objective: The goal of the research was to investigate the feasibility and acceptability of sleep EEG wearable devices among adolescents aged 11 to 17 years.

Methods: A total of 104 adolescents aged 11 to 17 years participated in 7 days of at-home sleep recording using a self-administered wearable sleep EEG device (Zmachine Insight+, General Sleep Corporation) as well as a wristworn actigraph. Feasibility was assessed as the number of full nights of successful recording completed by adolescents, and acceptability was measured by the wearable acceptability survey for sleep. Feasibility and acceptability were assessed separately for the sleep EEG device and wristworn actigraph.

Results: A total of 94.2% (98/104) of adolescents successfully recorded at least 1 night of data using the sleep EEG device (mean number of nights 5.42; SD 1.71; median 6, mode 7). A total of 81.6% (84/103) rated the comfort of the device as falling in the comfortable to mildly uncomfortable range while awake. A total of 40.8% (42/103) reported typical sleep while using the device, while 39.8% (41/103) indicated minimal to mild device-related sleep disturbances. A minority (32/104, 30.8%) indicated changes in their sleep position due to device use, and very few (11/103, 10.7%) expressed dissatisfaction with their experience with the device. A similar pattern was observed for the wristworn actigraph device.

Conclusions: Wearable sleep EEG appears to represent a feasible, acceptable method for sleep assessment among adolescents and may have utility for assessing and treating sleep disturbances at a population level. Future studies with adolescents should evaluate strategies for further improving usability of such devices, assess relationships between sleep EEG–derived metrics and health outcomes, and investigate methods for incorporating data from these devices into emerging digital interventions and applications.

Trial Registration: ClinicalTrials.gov NCT03843762; https://clinicaltrials.gov/ct2/show/NCT03843762

(JMIR Mhealth Uhealth 2020;8(10):e20590) doi:10.2196/20590

KEYWORDS

sleep; wearable; mHealth; adolescents; EEG; feasibility; acceptability; tolerability; actigraphy

XSL•FO
RenderX

Introduction

Adolescence is a critical period in human development, associated with rapid physical growth, brain and cognitive development, and marked changes in social roles as youth transition from dependent roles within their family of origin to the independence of adulthood. Given the importance of these developmental processes for shaping adult behavior, interventions supporting healthy behaviors during adolescence may be particularly likely to improve and support health outcomes across the lifespan [1]. Sleep is one health behavior that undergoes significant alternations in adolescence, including shifts toward shortened sleep periods, delayed circadian rhythms, and changes in sleep architecture, including reductions in deep sleep (ie, slow wave sleep [SWS]) [2]. Poor sleep during this period is associated with risk for detrimental physical and mental health outcomes such as higher rates of obesity and onset of psychiatric disorders [3], which in turn may place individuals at risk for a range of diseases over the long-term, including cardiometabolic illness and cancer, and result in high rates of health care use and socioeconomic costs [4].

Despite the centrality of sleep to health in adolescence and beyond, sleep health among adolescents is rarely adequately assessed or treated in routine health care settings. For example, in pediatric primary care, the most common setting in which adolescents interact with health care providers, screening for sleep problems is infrequent, and when it does occur, it is typically limited in scope and subject to reporter bias (ie, one question to the adolescent or parent querying whether sleep concerns are present) [5]. Barriers to thorough and accurate sleep assessment are myriad and include lack of awareness of the importance of sleep among both patients and physicians [6] as well as the cost, invasiveness, and inaccessibility associated with polysomnography (PSG), the gold standard evaluation of sleep [7]. Thus, identifying and combating sleep disturbances in adolescence—along with their varied and profound consequences—requires a new approach that supports accurate, convenient, low-cost sleep assessment that can be deployed at the population level in real-world, health care settings.

With the rise of mobile health (mHealth), wearable technologies have offered unique opportunities to assess and shape health behaviors among adolescents [8-10]. For this age group specifically, researchers have long touted the importance of feasible, at-home collection of sleep information using devices that are acceptable to adolescents, in part due to PSG-related alterations to sleep (eg, diminished sleep quality) that occur for youth in clinic or laboratory settings [11]. This led to the widespread research use of wristworn activity monitors, such as actigraphs, to assess sleep behaviors in pediatric populations [12,13], followed by popular use of commercial grade devices, such as Fitbit or Garmin watches, that estimate sleep based on activity patterns [14]. However, because these devices estimate sleep based on other metrics (eg, movement, heart rate) rather than providing a direct measure of sleep (ie, electroencephalogram [EEG]), wristworn devices tend to underestimate some important sleep parameters, particularly in those with disturbed sleep (eg, frequency/duration of night awakenings) [13], and significantly deviate from EEG devices in their estimation of transitions between sleep stages (ie, between light, deep, and rapid eye movement [REM] sleep). Specifically, a recent study found that wristworn devices overestimate the probability of remaining in a specific sleep stage and underestimate the probability of transitioning between sleep stages [15].

To address this limitation, ambulatory (at-home) full PSG was developed and has been used in pediatric research settings [11] to characterize maturation of sleep architecture across adolescence [16,17] as well as links between sleep and healthy behaviors (eg, exercise, nutrition) [18-20], psychiatric [21] and physical [22] health, and academic and social functioning [18,23] during this developmental period. However, these devices remain expensive and require medical professionals or trained research staff to administer, restricting their utility in routine care settings and at the population health level. Consumer-grade wearables incorporating a single or a few channels of EEG have been developed to assess sleep architecture in the home environment and are relatively low cost compared with traditional and ambulatory PSG [24]. Obtaining recordings with these devices can be achieved by the subjects themselves without the involvement of a PSG technologist. This is possible because they obtain EEG data exclusively from outside the hairline (typically from the forehead and sometimes the mastoids), whereas obtaining scalp EEG data from the usual locations employed for in-laboratory studies involves placing electrodes within the hairline, which requires a PSG technologist involvement in order to obtain adequate recordings. Some provide automatic scoring of sleep behaviors and architecture that are user-friendly for patients and providers [25,26]. Among adults, consumer-grade sleep EEG wearables have been used in research and clinic settings [27,28]. However, the feasibility and acceptability of collecting sleep data using self-administered, consumer-grade sleep EEG wearables among adolescents is currently unknown.

This study had two aims. First, we examined the feasibility of collecting sleep data using a self-applied, at-home, wearable sleep EEG recording device among adolescents aged 11 to 17 years. We hypothesized that most adolescents would be capable of recording sleep EEG at home (ie, successfully recording at least 1 night), and indeed, record sufficient data to provide a stable measure of sleep architecture and physiology (ie, at least 3 nights of successful recording [29]). We simultaneously collected information about sleep behavior using a device previously associated with high feasibility and acceptability in this age group (ie, wristworn actigraph [13]) and hypothesized that most adolescents would be able to record sufficient data to provide a stable measure of sleep behavior with this method (ie, at least 5 nights of successful recording [30]). We included both sleep EEG and actigraphy measures to demonstrate the feasibility of collecting data about multiple aspects of sleep (ie, sleep physiology and architecture via sleep EEG and 24-hour sleep behavior and circadian activity patterns via the actigraphy) via wearables in this age group. In addition, multiple nights of sleep data were collected due to prior research suggesting that the stability of sleep measures increases with more nights of collection [29,30]. Second, we examined the acceptability of the wearable sleep EEG device among adolescents and

XSL•FO
RenderX

hypothesized that these devices would be associated with high ratings of acceptability including a high frequency of endorsements of study satisfaction (>50%, or majority, of the sample) and a low frequency of endorsements of significant (ie, moderate to severe) device-related sleep disturbances and discomfort (<50%, or minority, of the sample). Again, we conducted a parallel query about the acceptability of a commonly used wearable in this age group (wristworn actigraph).

Methods

Participants

Adolescents between the ages of 11 years 0 months and 17 years 11 months were recruited through online and school advertisements, a participant contact database maintained by the Duke ADHD (Attention-Deficit/Hyperactivity Disorder) Program, and word of mouth. Inclusion criteria included the ability to follow written and verbal instructions in English, access to a smartphone in order to complete daily sleep diaries, and ability to comply with all study requirements and procedures. Because we were interested in evaluating feasibility of wearable sleep EEG device use among typical adolescents drawn from the community (rather than a clinical sample), participants were excluded if they indicated a diagnosis of sleep apnea or periodic leg movement syndrome, current use of prescribed or over-the-counter sleep aids (eg, sedatives, melatonin), diagnosis of an acute or chronic medical illness or use of medication that may interfere with sleep as determined by the research team, and/or inability to comply with study requirements. If subjects were currently taking melatonin supplements on an as-needed basis (ie, not daily), they were given the option to initiate a 7-day washout period in order to participate in the study (n=2). Participants and their guardians were compensated up to $200 for completing all study-related activities. Compensation was not based on the usability of data provided by participants. This study was approved by the Duke University School of Medicine institutional review board and registered with ClinicalTrials.gov [NCT03843762], and the study protocol was conducted according to Good Clinical Practice standards.

Procedure

After passing a brief prescreen assessment via telephone, participants attended an in-person intake visit accompanied by a parent or legal guardian. During the intake visit, participants and guardians signed an electronic informed consent form. Verbal assent was obtained from children younger than age 12 years. Eligible participants were provided with a sleep EEG device and actigraph and given detailed instructions on how to use and care for both devices. Specifically, participants were instructed to wear the actigraph on their nondominant wrist 24 hours a day for the study duration unless the device was going to be submerged in water, in which case participants were asked to remove the actigraph for as little time as possible. In order to ensure that the sleep EEG devices were used properly, a study coordinator reviewed the step-by-step instructions located in the device manual with each participant. Participants were also provided with a copy of the manual and directions on how to locate additional instructional videos on the manufacturer's website.

Sleep recordings were collected for 1 week beginning the evening of the in-person intake visit. In addition to using the wearable devices, participants also completed an electronic sleep diary each morning via Research Electronic Data Capture (REDCap), a secure web platform. The link to the daily sleep diary was sent each morning via text message, and an evening reminder was sent automatically if the diary was not completed. As is standard procedure for actigraphy [13], the sleep diary was used to ensure completeness of data collected by the wearable sleep devices. On day 8 of using the wearables, participants and their guardians completed an in-person follow-up visit, during which the actigraphs and sleep EEG devices were collected and adolescents completed a survey querying device acceptability via REDCap.

Measures

Wearable Sleep Devices

Sleep EEG was collected using a Zmachine Insight+ device (General Sleep Corporation). The Zmachine Insight+ is a single-channel EEG device with accompanying software that provides sleep staging scoring [25,26]. Participants were instructed to place one disposable (ie, one-time use) sensor behind each ear and one on the center of their neck directly prior to getting into bed each evening, and remove the sensors immediately following their final awakening. Sensors were connected to the Zmachine Insight+ device via wires, and participants were instructed to place the device next to their bed on a bedside table or similar. Actigraphy was collected using a wGT3X-BT (ActiGraph LLC) worn on participants' nondominant wrists. These actigraphic devices are easy to use, acceptable to adolescents, sufficiently reliable and valid to assess sleep behavior, and commonly used in clinical research settings [13]. The daily Consensus Sleep Diary queried previous night bedtime, sleep-onset latency, wake time during the night, time of final awakening, and final rising time [31].

Feasibility and Acceptability of Wearable Sleep Devices

Feasibility was evaluated by the rate of study completion and the number of nights with viable data (full/complete recordings). Acceptability was assessed using the Wearable Acceptability Survey for Sleep (WASS; Multimedia Appendix 1) developed by authors JRLA and ADK. This 10-item scale is designed to optimize face validity and measures comfort of wearable sleep devices while asleep and awake, sleep disturbances resulting from device use, and satisfaction with the devices. The first 3 questions were adapted from those developed by ADK and colleagues for a previous study [32], and the scale was further expanded upon with 7 additional items by JRLA and ADK [33]. Acceptability of sleep EEG and actigraphy were queried separately. Internal consistency of the scaled items (items 1, 2, 3, 7, and 8) was high when queried about actigraphy (Cronbach alpha=.79) and EEG (Cronbach alpha=.80). This measure has not been previously published.

XSL•FO
RenderX

Results

Demographics of the Sample

A total of 123 adolescents and their guardians expressed interest in study participation and completed an eligibility prescreening via telephone. Of these individuals, 3 were excluded due to use of sleep medications and 13 declined to participate in the study following the phone screen. A total of 107 participants completed the initial in-person visit. One individual attended the in-person visit but declined participation following the consent process. Two additional individuals completed the study in full but are not included in this analysis due to difficulty retrieving the devices attributable to circumstances surrounding the novel coronavirus (ie, North Carolina COVID-19 shelter-in-place order), leaving a final sample of 104. See Table 1 for a summary of the demographics of the sample, which is reflective of the population demographics of Raleigh-Durham-Chapel Hill, North Carolina.

Table 1. Demographics of the study sample (n=104).

Characteristic	Value
Age in years, mean (SD)	14.43 (1.77)
Sex, female, n (%)	49 (47.1)
Race, n (%)	
African American	29 (27.9)
White	67 (64.4)
Asian	3 (2.9)
Native American	1 (1.0)
More than 1 race	4 (3.8)
Ethnicity, n (%)	
Hispanic	7 (6.7)
Non-Hispanic	97 (93.3)

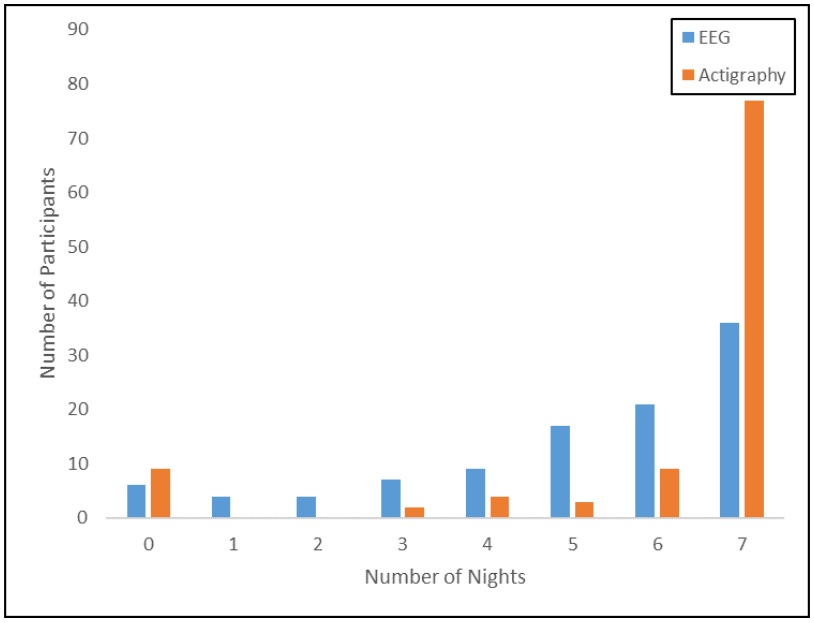
Feasibility of Wearable Sleep Devices

All participants (100%) who enrolled in the study following the intake visit completed the study. Regarding the sleep EEG device, 94.2% (98/104) of adolescents successfully achieved at least 1 full night of sleep recording. Of these individuals, the mean number of nights of successful recording was 5.42 (SD 1.71), the median number was 6, and the modal number was 7. A total of 86.5% (90/104) successfully recorded at least 3 full nights of sleep EEG recording, the amount recommended for obtaining stable measures of sleep architecture physiology using EEG. Of the 6 adolescents who were unable to successfully record EEG data, 4 adolescents appeared not to comply with sensor app instructions, 1 adolescent decided not to wear the device due to initial discomfort, and 1 adolescent accidently broke the device while it was not in use. It is notable that among the adolescents who successfully recorded at least 1 full night of data but less than 7 full nights of data, sleep recording was typically attempted by the adolescent on the other nights but the recording was incomplete (there were partial recordings because, for example, wires became unplugged so recording stopped).

A similarly high rate of adolescents successfully recorded at least 1 full night of sleep recording using the wristworn actigraph device (95/104, 91.3%). Of these individuals, the mean number of nights of successful recording was 6.63 (SD 0.90), the median number was 7, and the modal number was 7. A total of 85.6% (89/104) recorded at least 5 nights of actigraphy sleep recording, the number of nights recommended for obtaining stable measures of sleep behavior using actigraphy. Of the adolescents who were unable to unsuccessfully record sleep data via actigraphy, the reasons were device/battery failure (5/9), lost device (2/9), and decision not to wear due to initial discomfort (2/9). See Figure 1.

XSL•FO
RenderX

Figure 1. Number of nights of successful sleep recording by sleep device.

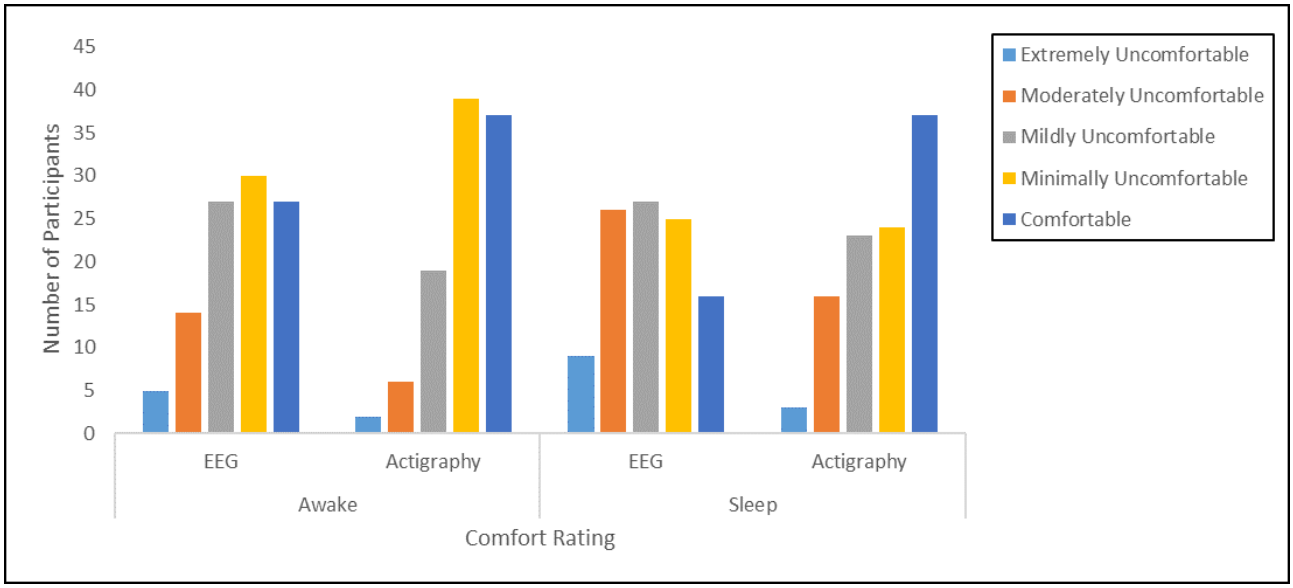


Acceptability of Wearable Sleep Devices

Comfort

The majority of adolescents rated both devices in the comfortable to mildly uncomfortable range while awake (sleep EEG: 81.6% (84/103); actigraphy: 92.2% (95/103) and sleeping (sleep EEG: 66.0% (68/103); actigraphy: 81.6% (84/103); Figure 2). When asked what led to discomfort when using the sleep EEG device, adolescents most frequently mentioned issues related to the wires (27/104; eg, getting caught/tangled or becoming detached/requiring reattachment due to movements during sleep) or the adhesive used to attach the sensors (28/104; eg, left residue or was itchy). A handful of participants mentioned discomfort due to difficulty changing sleep positions while wearing the device (5/104) and expressed concerns that they would mess up the recording while sleeping (3/104). When asked what led to discomfort while using the actigraph wristwatch, adolescents most often cited issues related to the light (18/104), which they found distracting while initiating sleep, and the wristband (26/104; eg, too loose or too tight, itchy/scratchy, too bulky). For both devices, it is notable that most adolescents rated any discomfort as minimal or mild.

Figure 2. Ratings of comfort for sleep wearables while awake and sleeping.



Wearable-Related Sleep Disturbances

Regarding perceived device-related interference with sleep, most adolescents indicated that their sleep was either typical of their normal sleep while wearing the sleep EEG device (42/103, 40.8%) or that the sleep EEG device resulted in minimal to mild sleep disturbances (41/103, 39.8%). Similarly, the majority of youth rated their sleep while wearing the wristworn actigraph as typical (64/103, 62.1%) or minimally to mildly disturbed (31/103, 30.1%). The most frequently reported device-related sleep disturbances due to EEG were difficulties falling asleep (24/104) and waking more frequently (23/104), followed by poorer sleep quality (18/104) and early waking (11/104). Most

XSL•FO
RenderX

of the reported disturbances associated with the wristworn actigraph regarded difficulties falling asleep (20/104), followed by more night awakenings (7/104), poorer sleep quality (6/104), and early waking (6/104; Figure 3).

A minority of adolescents reported that they slept in a different position due to use of the sleep EEG device (32/104, 30.8%) or wristworn actigraph (16/104, 15.4%; Figure 4). Of the participants who endorsed a change in sleeping position due to sleep EEG, endorsements of preferred typical sleeping position included (participants could select more than one) sleeping on their back (3/32), stomach (5/32), side (13/32), changing positions frequently (18/32), or sleeping in all positions equally (4/32). Of the participants who endorsed a change in sleeping position due to actigraphy, endorsements of preferred typical sleeping position included sleeping on their stomach (3/16), side (6/16), changing positions frequently (9/16), or sleeping in all positions equally (1/16). These results suggests that across both devices, device-related changes in sleep position were most frequently endorsed by youth who already experience frequent position changes followed by those who sleep on their sides.

Figure 3. Wearable-related sleep disturbances by device.

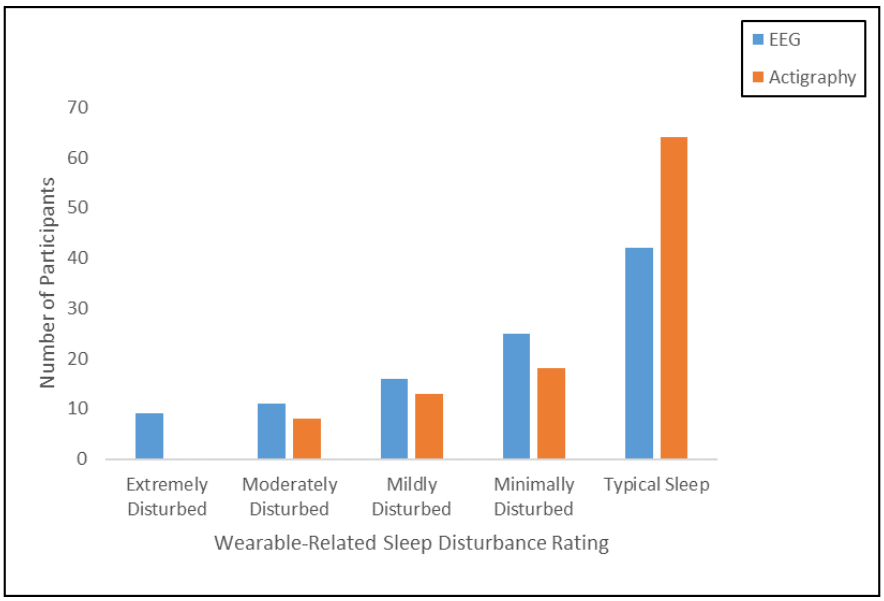
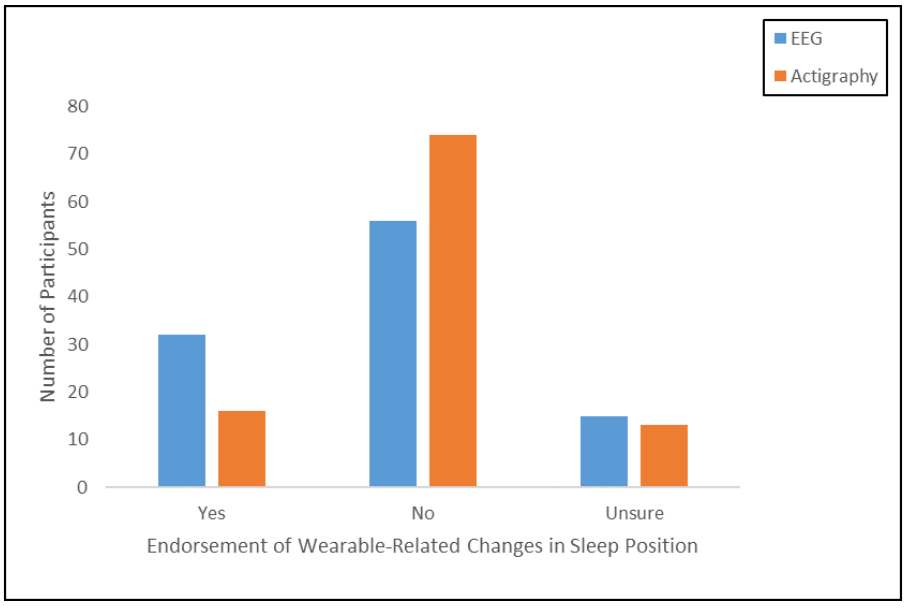


Figure 4. Endorsement of wearable-related changes in sleep position by device.



Satisfaction

Adolescents were most likely to rate themselves as satisfied with participation in the study based on their experience with the sleep EEG (53/103, 51.5%) and actigraphy (71/103, 68.9%) devices, followed by ambivalence (neither satisfied nor dissatisfied; 39/103, 37.9%, for the sleep EEG; 31/103, 30.1%, for the actigraph). Very few adolescents rated themselves as dissatisfied with their participation in the study based on experience with the sleep EEG (11/103, 10.7%) or actigraph (1/103, 1.0%; Figure 5).

Interestingly, there was a statistically significant association between satisfaction and device comfort while sleeping

XSL•FO
RenderX

($\chi^2_{4,n=103}$=17.21, *P*=.002); the small number of adolescents who rated themselves as dissatisfied with the study based on sleep EEG use were also more likely to have experienced significant (moderate to severe) device-related discomfort (number observed = 9, number expected = 3.74). Similarly, there was a statistically significant association between satisfaction and significant (moderate to severe) device-related sleep disturbances ($\chi^2_{4,n=103}$=21.28, *P*=.002), such that adolescents who rated themselves as dissatisfied with sleep EEG use were more likely to have experienced significant device-related sleep disturbances (number observed = 9, number expected = 2.14).

The majority of adolescents would recommend the study to a friend based on their experience using the sleep EEG (77/98, 79%) and actigraphy (88/98, 90%; Figure 6).

Figure 5. Overall study satisfaction by device.

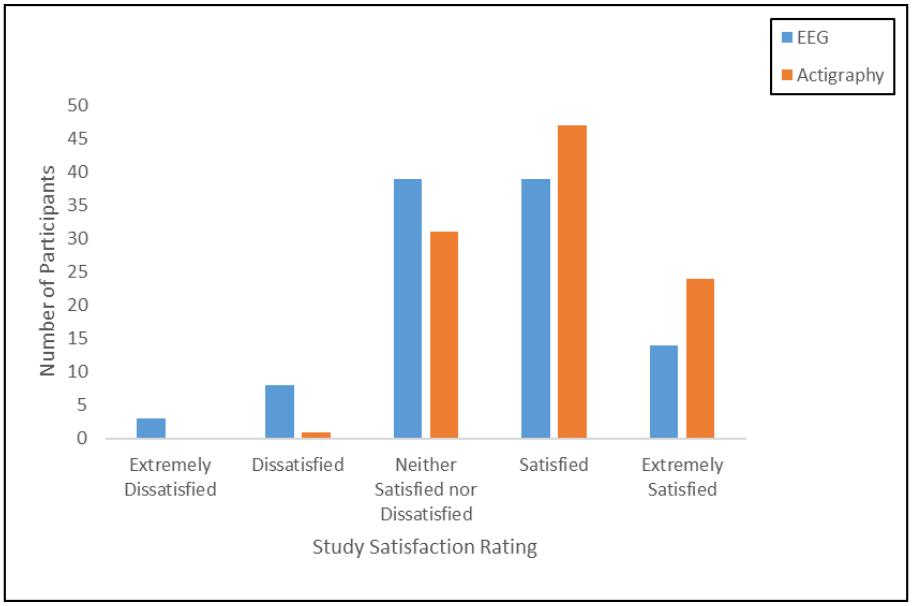
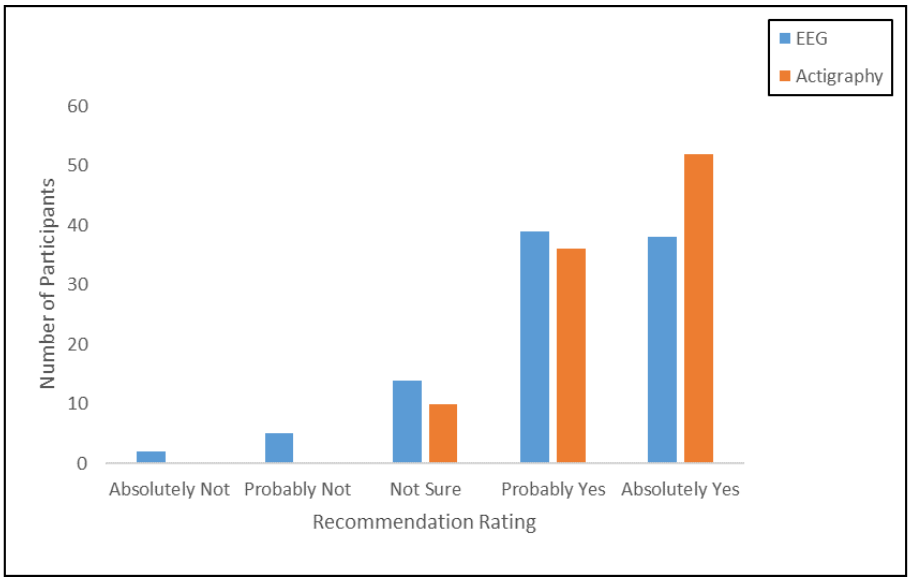


Figure 6. Number of adolescents recommending study to a friend by device.



Discussion

Principal Findings

Despite its importance in supporting overall health, sleep is infrequently evaluated during adolescence, a period characterized by shifts in sleep architecture and behavior. Adolescence has also been identified as a critical period for initiating lifelong health behaviors with the potential for a significant, long-term impact for the individual and societal/health care systems. Wearable sleep EEG devices have the potential to overcome obstacles to sleep assessments in routine health care settings and may allow for evaluation and treatment at a population level. However, to date, the usability of such devices in this age group has been unknown. Our findings suggest that the vast majority of adolescents were able to successfully record multiple nights of sleep data using the sleep EEG device. In addition, while adolescents described some discomfort and device-related sleep disturbances associated with wearing the sleep EEG device, most rated any discomfort as minimal to mild. Very few adolescents described

XSL•FO
RenderX

dissatisfaction with their experience in the study. Notably, this pattern of results was similar to the pattern observed when using a commonly administered wearable (wristworn actigraph) assessing sleep behavior for this age group although satisfaction was numerically higher and discomfort and disturbances numerically lower for the actigraph compared with the sleep EEG device. In addition, it is notable that sleep EEG device-related disturbances were more likely to be endorsed by adolescents who sleep in certain sleep positions (eg, side sleepers, individuals who already change positions frequently), suggesting some individuals may be better suited to use the device than others.

These results have implications for the use of wearable sleep EEG devices with adolescents in both clinical and research settings. First, although wristworn activity monitors have utility for providing information about some aspects of sleep [12,13], EEG measures provide a more robust picture of sleep health by providing more accurate measures of sleep onset latency and night awakenings, particularly in those with sleep disturbance, and SWS and REM sleep, aspects of sleep architecture that have been associated with both psychiatric and physical disease in this age group [34]. For example, the suppression of SWS and REM sleep has been shown to be associated with insulin resistance in adolescents [35], less REM sleep has been found among adolescents with asthma [36], and early onset of REM and increased REM density may characterize the sleep of adolescents with depression [37]. Thus, sleep EEG measures have the potential to provide a more comprehensive representation of adolescent sleep as it relates to physical and mental health outcomes, and this study suggests that commercial grade wearables may provide a feasible, acceptable method for assessing sleep architecture in this age group. However, demonstration of feasibility and acceptability of sleep EEG among adolescents is only the first step toward incorporation of sleep EEG into routine clinical settings, as additional challenges (eg, validation of automated scoring algorithms, validation of findings from recording EEG data from forehead or mastoid versus standard placement of EEG electrodes within the hairline, training primary care physicians to dispense and interpret sleep EEG devices) would need to be addressed in order to facilitate meaningful clinical application.

Second, given costs associated with traditional PSG assessments, including not only the devices themselves but also access to a specialized sleep clinic and trained personnel, assessment of adolescent sleep architecture at a population level has been, to date, unattainable. Access to sleep EEG assessments in routine health care settings, including pediatric primary care, has been similarly limited. Wearable sleep EEG devices provide a potential opportunity to expand access to sleep evaluations for adolescents in both research and clinical settings, which may in turn allow for the acceleration of sleep-based mHealth apps in this age group. As mentioned above, interest in developing user-friendly, efficacious mHealth apps among adolescents is increasing at a rapid pace [8-10], in part due to adolescent engagement and facility with emerging technologies [38]. Incorporation of sleep EEG assessment via wearables may allow for refinement of mHealth apps targeting sleep in this age group—for example, by helping to identify adolescents most likely to benefit from sleep-based mHealth and/or to aid in monitoring sleep improvements resulting from such interventions.

Future Directions

Results from this study offer many exciting new directions for future research. For example, although ambulatory PSG has been linked to health outcomes among adolescents, to date no studies have examined associations between consumer-grade sleep-EEG, which is scored using commercial algorithms and obtained from nonstandard scalp locations, and physical and psychiatric health in this age group. Given the importance of adolescence for shaping healthy behaviors across the lifespan, elucidating relationships between wearable sleep EEG and health outcomes is a critical next step for driving the development of mHealth apps targeting sleep and overall health in this age group. Second, future studies may examine the concordance between measures of sleep including consumer-grade sleep EEG, actigraphy, and self-report, as well as how these measures may be best used in conjunction to assess and support sleep health in this age group. Specifically, while sleep EEG has benefits in accurately assessing sleep stages, as a 24-hour measure of sleep behavior, actigraphy allows for the calculation of estimates of circadian patterns [39,40] as well as sensitive measures of the regularity of sleep patterns [41,42]. These findings suggest that both sleep EEG and wristworn wearables are feasible and tolerable in this age group and future studies should investigate how to optimize and integrate information collected from sleep wearables to enhance health outcomes among adolescents. Third, these findings may suggest avenues for improving sleep EEG technology. Namely, both feasibility and acceptability of sleep EEG devices in this age group were negatively impacted by specific features of the device, such as the wires, which contributed to both lost data (by becoming unplugged) and some, usually mild, discomfort. Wireless EEG systems have been developed and tested [43], and if successfully adapted, may have utility for assessing sleep in adolescents. Relatedly, unlike other commercial-grade devices, the device was attached to the head only via sensors rather than with a headband or cap [24]. The relative acceptability of these different features among adolescents is still an open question and may inform further refinement of sleep EEG systems. In addition, facilitating communication regarding sleep EEG data directly to adolescents and/or health care providers through smartphone apps or electronic health records may also support evaluation and treatments in this population and should be the focus of future work.

Limitations

This study is subject to several limitations. First, this study focused on feasibility and acceptability of collecting sleep EEG data in generally healthy adolescents who might be seen in routine health care settings. Relatedly, the selected device did not measure respiratory function. Thus, the feasibility and acceptability of sleep EEG devices, including those with respiratory measures, in adolescent clinical populations (eg, obstructive sleep apnea), remain open questions. Second, it was beyond the scope of this study to assess the performance of wearable sleep EEG devices—for example, as compared with

XSL•FO
RenderX

the gold standard sleep assessment (PSG). The selected device (Zmachine Insight+) has been shown to compare favorably with PSG in adults [26], and future studies should evaluate the performance of wearable sleep EEG devices in this age group.

Conclusions

Wearable sleep EEG devices are feasible and acceptable for use with adolescents, a group characterized by developmental sleep changes and whose health behaviors may have a long-term impact on their well-being. Additionally, these devices may have utility in both clinical and research settings and for assessing and treating sleep disturbances at a population level. Future studies with adolescents should evaluate strategies for further improving usability of such devices, assess relationships between sleep EEG–derived metrics and health outcomes, and investigate methods for incorporating data from these devices into emerging mHealth interventions and apps.

Acknowledgments

This study was funded by the Duke Institute for Brain Sciences through their Research Incubator Award mechanism. Further funding for JRLA was provided by National Institute of Mental Health (grant K23MH108704). We are grateful to the participants and their families for their contribution to research.

Conflicts of Interest

Authors do not have any conflicts of interest to report in regard to the devices used in this study. ADK has received research funding from Janssen Pharmaceuticals, Axsome Pharmaceutics, Reveal Biosensors, The Ray and Dagmar Dolby Family Fund, and the National Institutes of Health. He has received consulting fees from Adare, Axsome Therapeutics, Big Data, Eisai, Evecxia, Ferring Pharmaceuticals, Galderma, Harmony Biosciences, Janssen Pharmaceuticals, Jazz Pharmaceuticals, Millennium Pharmaceuticals, Merck, Neurocrine Biosciences, Pernix, Otsuka Pharmaceuticals, Sage, and Takeda. Authors JRLA, CEK, SHK, LJ, and MME do not have any conflicts of interest to report.

Multimedia Appendix 1
Wearable acceptability survey for sleep.
[DOCX File , 14 KB - mhealth_v8i10e20590_app1.docx]

References

1. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S, Dick B, Ezeh AC, et al. Adolescence: a foundation for future health. Lancet 2012 Apr 28;379(9826):1630-1640. [doi: 10.1016/S0140-6736(12)60072-5] [Medline: 22538178]
2. Crowley SJ, Wolfson AR, Tarokh L, Carskadon MA. An update on adolescent sleep: new evidence informing the perfect storm model. J Adolesc 2018 Aug;67:55-65 [FREE Full text] [doi: 10.1016/j.adolescence.2018.06.001] [Medline: 29908393]
3. Owens J, Adolescent Sleep Working Group, Committee on Adolescence. Insufficient sleep in adolescents and young adults: an update on causes and consequences. Pediatrics 2014 Sep;134(3):e921-e932. [doi: 10.1542/peds.2014-1696] [Medline: 25157012]
4. Garbarino S, Lanteri P, Durando P, Magnavita N, Sannita WG. Co-morbidity, mortality, quality of life and the healthcare/welfare/social costs of disordered sleep: a rapid review. Int J Environ Res Public Health 2016 Aug 18;13(8) [FREE Full text] [doi: 10.3390/ijerph13080831] [Medline: 27548196]
5. Meltzer LJ, Plaufcan MR, Thomas JH, Mindell JA. Sleep problems and sleep disorders in pediatric primary care: treatment recommendations, persistence, and health care utilization. J Clin Sleep Med 2014 Apr 15;10(4):421-426 [FREE Full text] [doi: 10.5664/jcsm.3620] [Medline: 24733988]
6. Perry GS, Patil SP, Presley-Cantrell LR. Raising awareness of sleep as a healthy behavior. Prev Chronic Dis 2013 Aug 08;10:E133 [FREE Full text] [doi: 10.5888/pcd10.130081] [Medline: 23928458]
7. Van de Water ATM, Holmes A, Hurley DA. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography: a systematic review. J Sleep Res 2011 Mar;20(1 Pt 2):183-200 [FREE Full text] [doi: 10.1111/j.1365-2869.2009.00814.x] [Medline: 20374444]
8. Lee AM, Chavez S, Bian J, Thompson LA, Gurka MJ, Williamson VG, et al. Efficacy and effectiveness of mobile health technologies for facilitating physical activity in adolescents: scoping review. JMIR Mhealth Uhealth 2019 Feb 12;7(2):e11847 [FREE Full text] [doi: 10.2196/11847] [Medline: 30747716]
9. Jeminiwa RN, Hohmann NS, Fox BI. Developing a theoretical framework for evaluating the quality of mHealth apps for adolescent users: a systematic review. J Pediatr Pharmacol Ther 2019;24(4):254-269 [FREE Full text] [doi: 10.5863/1551-6776-24.4.254] [Medline: 31337988]
10. 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]
11. Stores G, Crawford C, Selman J, Wiggs L. Home polysomnography norms for children. Technol Health Care 1998 Nov;6(4):231-236. [Medline: 9924950]

XSL•FO
RenderX

12. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev 2011 Aug;15(4):259-267. [doi: 10.1016/j.smrv.2010.10.001] [Medline: 21237680]
13. Meltzer LJ, Montgomery-Downs HE, Insana SP, Walsh CM. Use of actigraphy for assessment in pediatric sleep research. Sleep Med Rev 2012 Oct;16(5):463-475 [FREE Full text] [doi: 10.1016/j.smrv.2011.10.002] [Medline: 22424706]
14. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
15. Liang Z, Chapa-Martell MA. Accuracy of Fitbit wristbands in measuring sleep stage transitions and the effect of user-specific factors. JMIR Mhealth Uhealth 2019 Jun 06;7(6):e13384 [FREE Full text] [doi: 10.2196/13384] [Medline: 31172956]
16. Feinberg I, de Bie E, Davis NM, Campbell IG. Topographic differences in the adolescent maturation of the slow wave EEG during NREM sleep. Sleep 2011 Mar 01;34(3):325-333 [FREE Full text] [doi: 10.1093/sleep/34.3.325] [Medline: 21358849]
17. Baker FC, Turlington SR, Colrain I. Developmental changes in the sleep electroencephalogram of adolescent boys and girls. J Sleep Res 2012 Feb;21(1):59-67 [FREE Full text] [doi: 10.1111/j.1365-2869.2011.00930.x] [Medline: 21668552]
18. Brand S, Gerber M, Beck J, Hatzinger M, Pühse U, Holsboer-Trachsler E. Exercising, sleep-EEG patterns, and psychological functioning are related among adolescents. World J Biol Psychiatry 2010 Mar;11(2):129-140. [doi: 10.3109/15622970903522501] [Medline: 20109114]
19. Brand S, Beck J, Gerber M, Hatzinger M, Holsboer-Trachsler E. Evidence of favorable sleep-EEG patterns in adolescent male vigorous football players compared to controls. World J Biol Psychiatry 2010 Mar;11(2 Pt 2):465-475. [doi: 10.1080/15622970903079820] [Medline: 19606405]
20. Awad KM, Drescher AA, Malhotra A, Quan SF. Effects of exercise and nutritional intake on sleep architecture in adolescents. Sleep Breath 2013 Mar;17(1):117-124 [FREE Full text] [doi: 10.1007/s11325-012-0658-6] [Medline: 22331514]
21. Santangeli O, Porkka-Heiskanen T, Virkkala J, Castaneda AE, Marttunen M, Paunio T, et al. Sleep and slow-wave activity in depressed adolescent boys: a preliminary study. Sleep Med 2017 Oct;38:24-30. [doi: 10.1016/j.sleep.2017.06.029] [Medline: 29031752]
22. Stores G, Fry A, Crawford C. Sleep abnormalities demonstrated by home polysomnography in teenagers with chronic fatigue syndrome. J Psychosom Res 1998 Jul;45(1):85-91. [doi: 10.1016/s0022-3999(98)00024-5]
23. Perfect MM, Levine-Donnerstein D, Archbold K, Goodwin JL, Quan SF. The contribution of sleep problems to academic and psychosocial functioning. Psychol Schools 2014 Jan 28;51(3):273-295. [doi: 10.1002/pits.21746]
24. Bianchi MT. Sleep devices: wearables and nearables, informational and interventional, consumer and clinical. Metabolism 2018 Dec;84:99-108. [doi: 10.1016/j.metabol.2017.10.008] [Medline: 29080814]
25. Wang Y, Loparo KA, Kelly MR, Kaplan RF. Evaluation of an automated single-channel sleep staging algorithm. Nat Sci Sleep 2015;7:101-111 [FREE Full text] [doi: 10.2147/NSS.S77888] [Medline: 26425109]
26. Kaplan RF, Wang Y, Loparo KA, Kelly MR, Bootzin RR. Performance evaluation of an automated single-channel sleep-wake detection algorithm. Nat Sci Sleep 2014;6:113-122 [FREE Full text] [doi: 10.2147/NSS.S71159] [Medline: 25342922]
27. Szczygiel EJ, Cho S, Tucker RM. Multiple dimensions of sweet taste perception altered after sleep curtailment. Nutrients 2019 Aug 27;11(9) [FREE Full text] [doi: 10.3390/nu11092015] [Medline: 31461917]
28. Dunn LK, Kleiman AM, Forkin KT, Bechtel AJ, Collins SR, Potter JF, et al. Anesthesiology resident night float duty alters sleep patterns: an observational study. Anesthesiology 2019 Aug;131(2):401-409. [doi: 10.1097/ALN.0000000000002806] [Medline: 31149926]
29. Wohlgemuth WK, Edinger JD, Fins AI, Sullivan RJ. How many nights are enough? The short-term stability of sleep parameters in elderly insomniacs and normal sleepers. Psychophysiology 1999 Mar;36(2):233-244. [Medline: 10194970]
30. Acebo C, Sadeh A, Seifer R, Tzischinsky O, Wolfson AR, Hafer A, et al. Estimating sleep patterns with activity monitoring in children and adolescents: how many nights are necessary for reliable measures? Sleep 1999 Feb 01;22(1):95-103. [doi: 10.1093/sleep/22.1.95] [Medline: 9989370]
31. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
32. Dieffenderfer J, Krystal A, Bozkurt A. Shedding light to sleep studies. Proc SPIE 2017;10352. [doi: 10.1117/12.2275625]
33. Lunsford-Avery JR, Kollins SH, Krystal AD. Sleeping at home: feasibility and tolerability of ambulatory polysomnography for use with adolescents with attention-deficit/hyperactivity disorder. Sleep 2018 Apr;41:A283-A284. [doi: 10.1093/sleep/zsy061.761]
34. Brand S, Kirov R. Sleep and its importance in adolescence and in common adolescent somatic and psychiatric conditions. Int J Gen Med 2011;4:425-442 [FREE Full text] [doi: 10.2147/IJGM.S11557] [Medline: 21731894]
35. Dutil C, Chaput J. Inadequate sleep as a contributor to type 2 diabetes in children and adolescents. Nutr Diabetes 2017 May 08;7(5):e266 [FREE Full text] [doi: 10.1038/nutd.2017.19] [Medline: 28481337]
36. Khassawneh B, Tsai SC, Meltzer LJ. Polysomnographic characteristics of adolescents with asthma and low risk for sleep-disordered breathing. Sleep Breath 2019 Sep;23(3):943-951. [doi: 10.1007/s11325-018-01774-3] [Medline: 30689097]
37. Augustinavicius JLS, Zanjani A, Zakzanis KK, Shapiro CM. Polysomnographic features of early-onset depression: a meta-analysis. J Affect Disord 2014 Apr;158:11-18. [doi: 10.1016/j.jad.2013.12.009] [Medline: 24655760]

XSL•FO
RenderX

38. Anderson M, Jiang J. Teens, social media & technology. Washington: Pew Internet and American Life Project; 2018 May 31. URL: http://assets.pewresearch.org/wp-content/uploads/sites/14/2018/05/31102617/PI_2018.05.31_TeensTech_FINAL.pdf [accessed 2020-05-05]
39. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003 May 01;26(3):342-392. [Medline: 12749557]
40. Littner M, Kushida CA, Anderson WM, Bailey D, Berry RB, Davila DG, Standards of Practice Committee of the American Academy of Sleep Medicine. Practice parameters for the role of actigraphy in the study of sleep and circadian rhythms: an update for 2002. Sleep 2003 May 01;26(3):337-341. [doi: 10.1093/sleep/26.3.337] [Medline: 12749556]
41. Lunsford-Avery JR, Engelhard MM, Navar AM, Kollins SH. Validation of the sleep regularity index in older adults and associations with cardiometabolic risk. Sci Rep-UK 2018 Sep 21;8(1) [FREE Full text] [doi: 10.1038/s41598-018-32402-5] [Medline: 30242174]
42. Phillips A, Clerx W, O'Brien C, Sano A, Barger L, Picard R, et al. Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing. Sci Rep-UK 2017 Jun 12;7 [FREE Full text] [doi: 10.1038/s41598-017-03171-4] [Medline: 28607474]
43. Kam JWY, Griffin S, Shen A, Patel S, Hinrichs H, Heinze H, et al. Systematic comparison between a wireless EEG system with dry electrodes and a wired EEG system with wet electrodes. Neuroimage 2019 Jan 01;184:119-129 [FREE Full text] [doi: 10.1016/j.neuroimage.2018.09.012] [Medline: 30218769]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
EEG: electroencephalogram
mHealth: mobile health
PSG: polysomnography
REDCap: Research Electronic Data Capture
REM: rapid eye movement sleep
SWS: slow-wave sleep
WASS: Wearable Acceptability Survey for Sleep

Edited by G Eysenbach; submitted 25.05.20; peer-reviewed by J Zhang, E Kutafina; comments to author 29.06.20; revised version received 05.08.20; accepted 06.08.20; published 01.10.20.

Please cite as:
Lunsford-Avery JR, Keller C, Kollins SH, Krystal AD, Jackson L, Engelhard MM
Feasibility and Acceptability of Wearable Sleep Electroencephalogram Device Use in Adolescents: Observational Study
JMIR Mhealth Uhealth 2020;8(10):e20590
URL: https://mhealth.jmir.org/2020/10/e20590
doi:10.2196/20590
PMID:33001035

©Jessica R Lunsford-Avery, Casey Keller, Scott H Kollins, Andrew D Krystal, Leah Jackson, Matthew M Engelhard. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 01.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User Experiences of a Smartphone-Based Attentive Eating App and Their Association With Diet and Weight Loss Outcomes: Thematic and Exploratory Analyses From a Randomized Controlled Trial

Victoria Whitelock[1,2], BSc, PhD; Inge Kersbergen[3], BSc, MSc, PhD; Suzanne Higgs[4], BA, MA, PhD; Paul Aveyard[5], PhD, FRCP, FRCGP, FFPH; Jason CG Halford[1], BSc, PhD, CPsychol; Eric Robinson[1], PhD

[1]Department of Psychological Sciences, University of Liverpool, Liverpool, United Kingdom
[2]Cancer Intelligence, Cancer Research UK, London, United Kingdom
[3]School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom
[4]The School of Psychology, University of Birmingham, Birmingham, United Kingdom
[5]Nuffield Department of Primary Care Health Services, University of Oxford, Oxford, United Kingdom

Corresponding Author:
Eric Robinson, PhD
Department of Psychological Sciences
University of Liverpool
Eleanor Rathbone Building
Bedford Street South
Liverpool, L69 7ZA
United Kingdom
Phone: 44 (0)151 794 1187
Email: eric.robinson@liv.ac.uk

Abstract

Background: Short-term laboratory studies suggest that eating attentively can reduce food intake. However, in a recent randomized controlled trial we found no evidence that using an attentive eating smartphone app outside of the laboratory had an effect on energy intake or weight loss over 8 weeks.

Objective: This research examined trial participants' experiences of using an attentive eating smartphone app and whether app usage was associated with energy intake and weight loss outcomes over 8 weeks.

Methods: We conducted thematic analysis of semistructured interviews (N=38) among participants in the attentive eating smartphone app group of the trial who completed the 8-week assessment. Linear regression models examined the associations between energy intake and weight loss outcomes at 8 weeks and app usage.

Results: Participants reported several barriers and facilitators to using the smartphone app, including repetition of app content, social setting, motivation, and habitual use of the app. Participants believed that using the app had some beneficial effects on their eating behavior and diet. Exploratory analyses indicated that more frequent recording of eating episodes in the app was associated with lower body weight (B=–0.02, P=.004) and greater self-reported energy intake (B=5.98, P=.01) at 8 weeks, but not body fat percentage or taste-test energy intake. Total audio clip plays, gallery views, and percentage of food entries recorded using an image were not significantly associated with energy intake or weight.

Conclusions: Frequent recording of eating episodes in a smartphone app was associated with greater weight loss. There are barriers and facilitators to frequent use of an attentive eating smartphone app that may be useful to address when designing dietary behavior change smartphone apps.

Trial Registration: ClinicalTrials.gov NCT03602001; https://clinicaltrials.gov/ct2/show/NCT03602001; Open Science Framework DOI 10.17605/osf.io/btzhw; https://osf.io/btzhw/

(JMIR Mhealth Uhealth 2020;8(10):e16780) doi:10.2196/16780

XSL•FO
RenderX

KEYWORDS

attentive eating; weight loss; smartphone app; eHealth; mHealth; food intake; obesity; overweight; focused attention; participant experience

Introduction

Lifestyle factors are the biggest contributors to disability-adjusted life years (DALYs) lost, with most DALYs lost attributable specifically to dietary factors (including excess body weight) having increased significantly from 2005 to 2015 [1]. In high-income English-speaking countries, obesity currently exceeds 30% in men and women [2]. Interventions are therefore needed that support people to change their diet and lose weight.

Smartphones are a viable way to deliver health interventions. For example, a large percentage (76%) of UK adults now own a smartphone [3], and smartphone apps can be used to deliver behavior change interventions [4]. However, sustained use of such apps may be needed for behavior change [5], as minimal usage could undermine effects of potentially successful interventions [4,6]. For example, Mattila and colleagues [5] found that sustained users of technology-based tools targeting a range of health behavior risk factors showed a greater reduction in body fat percentage and waist circumference than nonsustained users. Qualitative research has been used to explore factors associated with the use of smartphone apps designed to promote behavior change [7,8]. The results of such research can be used to identify reasons for poor or nonsustained use and can be used to inform future development of health behavior change smartphone apps. For example, the use of health behavior change apps has been found to depend on the app being user friendly, intuitive, accessible, and well structured, as well as having personalized features [7,9,10]. The ability to access novel and updated content and receiving feedback on progress toward goals have also been found to motivate app use [7,11].

Paying more attention to food being consumed, termed *attentive eating*, has been found to reduce snack intake 2-3 hours later in some [12-15], but not other [16-18], laboratory studies. If eating attentively is sustained over longer periods, it may support people in eating less and losing weight. We developed and tested the feasibility of a smartphone-based attentive eating app designed to encourage a more attentive eating style in daily life. In a feasibility trial, participants with overweight and obesity used the app and reported in qualitative interviews that they found the app easy and acceptable to use, and it increased their awareness of what they had been eating [19]. We subsequently conducted a randomized controlled trial testing the efficacy of the multicomponent, attentive eating smartphone app plus standard dietary advice, compared to standard dietary advice only, to support weight loss in adults with overweight and obesity [20]. We found no significant effect of using the attentive eating smartphone app, compared to the control condition, on weight loss (−0.10 kg, 95% CI −1.6 to 1.3) or energy intake after 8 weeks of usage. At the end of the trial period, semistructured interviews were conducted with participants who had used the attentive eating smartphone app in order to understand user experiences. Here we report the results of qualitative analyses of these interviews. These results showed that some participants experienced benefits of using the attentive eating app, which may suggest that specific patterns of app use may be associated with dietary change and weight loss. Therefore, we also conducted exploratory analyses to examine whether greater use of specific app functions was associated with greater reductions in energy intake and weight during the trial; the results are reported here.

Methods

Design and Sample

The trial was prospectively registered on the Open Science Framework [21] and retrospectively registered at ClinicalTrials.gov (NCT03602001). A full description of the trial design and methods are reported elsewhere [20]. In brief, this was an 8-week randomized controlled trial. Adults with overweight and obesity (mean BMI ≥25 kg/m^2) were randomized to use a multicomponent attentive eating smartphone app, along with receipt of standard dietary advice delivered via a booklet and text messages once a week (intervention group), or to receive standard dietary advice only (control group). Standard dietary advice only for the control group was similar to previous studies [22]. Participants were also required to have no history of eating disorders or food allergies, self-reported by participant; be aged 18-65 years; be fluent English speakers; not be taking medication that affects appetite; not be pregnant; not be scheduled for weight loss surgery during the trial; own an Android or Apple smartphone (Android operating system versions 4.4-7.1; Apple operating systems iOS 8-10); and not currently be on a structured weight loss program or using other weight loss apps.

Intervention Components

Standard Dietary Advice

All participants were provided with a booklet based on British Heart Foundation materials [23] containing tips on healthy eating (eg, the importance of a balanced diet, reducing calories and lower energy swaps, eating fruits and vegetables, reducing intake of high-fat and high-sugar foods and drinks, shopping, and eating out) and brief information about physical activity. Participants also received a text message once a week that repeated the information in the dietary advice booklet.

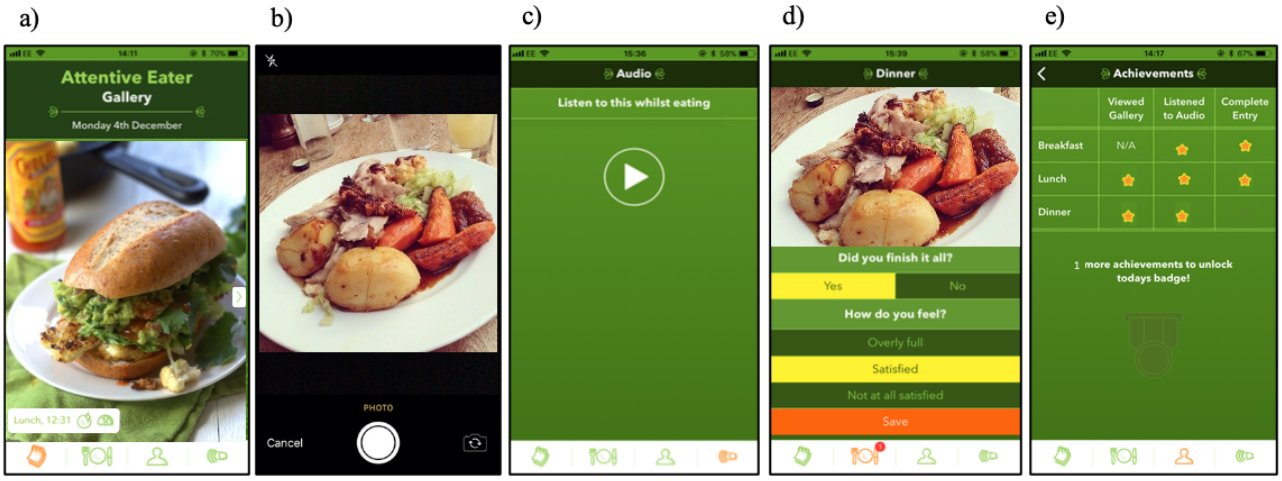
Attentive Eating App

The attentive eating smartphone app was designed to encourage a more attentive eating style by requiring users to photograph food and drink being consumed and then review this information when making dietary decisions throughout the day. Prior to eating and drinking, users accessed the camera function of their phones via the app and took a photograph of what they were about to consume. Alternatively, users could enter a description of the food or drink or upload a photo taken previously. Once they had finished eating and drinking, users answered questions about their consumption experience (ie, “Did you finish it all?”

XSL•FO
RenderX

and "How do you feel?"). Once these questions had been answered, the food and drink entry was completed and the consumption episode, along with the photograph and answers to the consumption experience questions, were logged in a food gallery section of the app. Prior to deciding what and how much to eat for a meal, users of the app reviewed their consumption episodes for the day up to that point by navigating through the photographs and information about their consumption experiences stored in the food gallery. While eating and drinking, users were encouraged to listen to a 2.5-minute audio clip that prompted them to pay more attention to what they were eating. The audio clip did this by asking listeners to pay attention to the smell, taste, and texture of the food, as well as how full they felt. Users were also encouraged to eat slowly, one mouthful at a time, and to periodically think about how much food was on their plate at the beginning and how much they had eaten. To foster positive feedback, users received daily in-app *stars* for using the key functions of the app at appropriate times (eg, reviewing the gallery shortly before a meal) and received a daily badge if they obtained all of the stars in a day (see Figure 1 for screenshots of the app). Participants also received a short leaflet that explained the principles of attentive eating and other ways to eat attentively, such as avoiding eating while distracted.

Figure 1. Screenshots of the key functions of the smartphone app: a) The food gallery; b) Photographing a meal; c) The attentive eating audio clip; d) Consumption experience questions; and e) Star and daily badge achievements.



Outcomes

We used four measures collected at baseline and at the 8-week follow-up: (1) body weight, (2) body fat percentage (3), self-reported 24-hour energy intake, and (4) laboratory-measured bogus taste-test energy intake. Body weight and body fat percentage were measured with the Tanita BC-418 MA body composition analyzer (Tanita Corporation). Self-reported 24-hour energy intake was measured using myfood24, an online, automated 24-hour dietary assessment system developed and validated for use in the United Kingdom [24,25]. In the laboratory-measured bogus taste test, participants were provided with three bowls of three different kinds of biscuits, 50 g each—Maryland chocolate chip cookies, ~249 kcal; Cadbury's chocolate fingers, ~240 kcal; and McVitie's digestives, ~241 kcal—broken up into small pieces. In the laboratory, participants were given 10 minutes to rate the biscuits on 100-point visual analog scales, ranging from *not at all* to *extremely*, on a number of features (eg, crunchiness and flavorsome) and were told that they could eat as many biscuits as they wished [26]. Participants were asked not to eat for 1 hour prior to the assessment sessions in order to standardize hunger. Hunger was measured immediately before the taste test using a 100-point visual analog scale, ranging from *not at all* to *extremely*. The biscuits were weighed afterward in order to calculate the amount of biscuits consumed, and weights were converted to total kcals.

Interviews

During the study visit at the 8-week follow-up, semistructured interviews were conducted with participants in the intervention group (ie, standard dietary advice and attentive eating app) to understand their experiences of using the attentive eating app. Participants were asked to be as honest as possible, were told that the researcher was interested in their experiences and opinions, and were told that there were no right or wrong answers. Participants were then asked what using the app was like for them and how they managed to include using the app into their lives. Participants were also asked what influenced how regularly they were able to use the app and whether using it affected awareness of what they were eating and/or what they ate. Interviews lasted approximately 10-30 minutes.

Analyses

Semistructured interviews were audio recorded, transcribed verbatim, and imported into NVivo (QSR International) for analysis. Thematic analysis was conducted, following the steps of Braun and Clark [27]. The data were first coded using a systematic and inductive approach; initial codes were then aggregated into overarching themes by a single researcher (IK). The appropriateness of each theme and each code aggregated into a theme was then reviewed by a second researcher (VW). Any disagreements were resolved through discussion.

Ordinary least squares linear regressions were used to examine the relationship between app usage variables (ie, independent variables: total eating episodes recorded, total audio clip listens,

XSL•FO
RenderX

total gallery views, and percentage of eating episodes recorded using an image) and each 8-week trial outcome (ie, dependent variables: body weight, body fat percentage, self-reported 24-hour energy intake, and taste-test energy intake), while controlling for baseline measurement of the dependent variables, total number of days in the trial, and hunger (in the analyses with taste-test energy intake as the dependent variable). The rationale for inclusion of independent variables was derived from the thematic analysis presented in the Summary of Thematic Analysis section. We analyzed the total number of food entries, audio plays, and gallery views, as participants were asked to complete these functions as often as possible. We analyzed the proportion of food entries that were photographed, as participants were asked to photograph as many of their meals as possible. A Bonferroni correction was used to account for multiple comparisons in statistical testing (0.05/4 regression models = 0.0125).

Data Availability

The dataset for the quantitative analyses is available on the Open Science Framework [28]. The qualitative interview data are available upon request.

Results

Sample

Interviews were conducted with all 39 participants in the intervention group that attended the 8-week visit; however, due to equipment malfunction we were unable to transcribe one of the interviews, leaving 38 interviews. The exploratory quantitative analyses included all participants in the intervention group who completed all assessment sessions (N=39); see Table 1 for characteristics of the analyzed sample.

Table 1. Baseline characteristics of the analyzed sample.

Characteristic	Value (N=39), mean (SD) or n (%)
Age (years), mean (SD)	41.7 (10.3)
Gender (female), n (%)	31 (79)
Ethnicity (White), n (%)	35 (90)
Education level, n(%)[a]	
Entry level or equivalent	0 (0)
General Certificate of Secondary Education or equivalent	6 (15)
Advanced (A) or Advanced Subsidiary (AS) level or equivalent	8 (21)
Undergraduate degree or equivalent	15 (38)
Higher degree or equivalent	7 (18)
Other	3 (8)
Baseline BMI (kg/m^2), mean (SD)	35.2 (7.2)
Baseline weight (kg), mean (SD)	99.5 (22.7)
Baseline body fat (%), mean (SD)	42.4 (8.5)
Baseline taste-test energy intake (kcal), mean (SD)	118.3 (98.4)
Baseline self-reported energy intake (kcal), mean (SD)	2047.0 (726.0)

[a]Percentages do not add up to 100 due to rounding.

Summary of Thematic Analysis

Two overarching themes were identified: (1) barriers and facilitators to app usage and (2) effects of the app. See Table 2 for example quotes.

XSL•FO
RenderX

Table 2. Table of themes and supporting quotes.

Theme	Supporting quotes
Theme 1: barriers and facilitators to app usage	
Subtheme 1: believing that the app was effective	PP[a] 103: "I didn't really see any effect of doing it so it just seemed like it wasn't worth the effort."
Subtheme 2: motivation to get the daily badge	PP 34: "It made me more aware to use the app rather than just like, 'oh, I'll just put that in later,' it was like 'no, I need to get my badge today.'"
	PP 64: "If you've messed up once in a day you tend to then think 'well, I can't get all the stars now today' so you know you're not going to get the trophy. So you sort of don't feel quite as competitive about getting them, but if you get them from the beginning of the day, you think 'right, I want to get them all today.'"
Subtheme 3: habitual, routine app use and distractions	PP 53: "I'd got into a routine that when I was in the kitchen or when I was doing meals the phone came, the picture came, everything came, but if I went out of the routine I was in, then I wouldn't think to do it."
Subtheme 4: not using full app in social situations	PP 53: "If I'm out socially with friends, family, or whatever and then to just kind of interrupt and say, 'I have to listen to this for five minutes,' it didn't feel right to do it."
	PP 35: "I think that's a real age thing, I am not someone who posts my food on Instagram. Also feeling a bit embarrassed that I didn't want to have to explain to people if they said, 'why are you taking a picture of your dinner?'"
Subtheme 5: wanting a good-looking food gallery	PP 90: "I wanted it to look like a good gallery and not a bad one just for my own personal, I don't know, pride, satisfaction" and "it would make me want to eat more of those nice fresh homemade things."
	PP 96: "I kind of stopped using it 'cause I was like 'don't judge me.'"
Subtheme 6: meal descriptions used for retrospective entries	PP 64: "I ended up writing the description in more than the picture towards the end."
	PP 03: "Where I'd forgotten to take a photo I was trying to write in the description. But I don't think that that worked as well as it would have done if I had of taken a photo in terms of looking back over it."
Subtheme 7: repetition of the audio clip	PP 103: "It was quite repetitive so I didn't want to listen to the same thing over and over."
Theme 2: effects of the app	
Subtheme 1: making healthier food choices	PP 92: "I had leeks last night on my carvery, which I never would have picked."
Subtheme 2: eating smaller portions	PP 7: "I had rice and a sweetcorn dish last night, and normally I'd put two big spoons on, and last night it was just the one and the one spoon was plenty."
	PP 71: "I bought a smaller plate; I said, 'right that's mine,' so I am more aware of the portion size."
Subtheme 3: reviewing the gallery informed eating	PP 78: "If I was hungry, I would look back to see what I've eaten, 'ah yeah, you've ate plenty,' and that would stop me."
	PP 34: "I've noticed I'm very very bad at eating late at night."
Subtheme 4: greater attention to hunger and fullness	PP 63: "I'll stop because I'm full, and I'm not eating the full thing because I'm actually full halfway through it."
Subtheme 5: eating more slowly	PP 4: "It's helped me to slow down when I'm eating."
Subtheme 6: eating attentively when not listening to the audio clip	PP 56: "Even though I'm not listening to it, I know what it asks me to do every time. So, yes, it would get ingrained in you and you would do it."
	PP 98: "Even when I'm not out with the app, I definitely chew slower."
Subtheme 7: enhanced eating experience	PP 22: "I noticed that I was tasting the food more and experiencing the food more."

[a]PP: participant; participant numbers exceed 39, as 104 participants (across arms) took part in the trial.

Barriers and Facilitators to App Usage

Key barriers to using the app included feeling that using the app in social settings was inappropriate, in particular, not wanting to be seen taking pictures of one's food. We noted that older participants were more likely to describe these as barriers to using the app (see Table 2). Believing that the app was effective was identified as a motivator and facilitator to using the app, whereas appearing to not get anything out of using the app reduced motivation to use it. When a person was following their usual daily routine and had incorporated using the app into this routine, they found it easier to remember to use the app, while distractions reduced its use. Those who found the in-app stars and daily badge rewarding described these as motivators to use the app more frequently. However, missing out on a star achievement early in the day reduced motivation to continue using the app for the rest of the day. For some, wanting to have a good-looking diary motivated them to eat "healthier," to eat

XSL•FO
RenderX

nicer looking food, and to make sure they recorded this. For others, wanting to impress the researcher or feeling like their food gallery was going to be judged meant that they recorded food intake less often if they were going to eat food they considered incompatible with their goals. Being able to write descriptions of food consumed when someone had forgotten or was unable to take a photograph was helpful because it meant that participants could still complete their task even if they did not photograph the food. However, participants tended to take fewer photographs as time went on, despite finding it more difficult to remember how much they had eaten from the description than the photograph. Users found it boring to listen to the same audio clip repeatedly and this demotivated them to keep doing so.

Effects of the App

Participants reported that using the app helped them eat more slowly, pay more attention to hunger and fullness, make healthier food choices, and eat smaller portion sizes, and it enhanced the enjoyment of eating. Some participants described continuing to follow the focused attention principles described in the audio clip even when not listening to it during a meal, such as eating slowly. Participants also described using the food gallery to inform their eating choices and to increase awareness of their eating patterns.

Exploratory Analyses

Usage of the app functions was generally reasonable: median total food entries was 136.0 (IQR 98), median total times listened to audio was 42.0 (IQR 66), and median percentage of food entries recorded using an image was 78.6% (IQR 28.1). However, median total gallery views (45.0, IQR 56) was low—the trial period lasted 56 days. VIF (variance inflation factor) (<0.96) and tolerance statistics (<4.33) showed that levels of multicollinearity were acceptable. There was no evidence that the number of audio clip plays, gallery views, or percentage of food entries recorded using an image were associated with taste-test or self-reported energy intake, weight, or body fat percentage at the 8-week follow-up (see Table 3). However, recording more food entries was associated with lower body weight and greater self-reported energy intake at 8 weeks. The number of entries recorded was not significantly associated with taste-test energy intake or body fat percentage.

Table 3. Regression results for app usage variables as predictors of trial outcomes at 8 weeks.

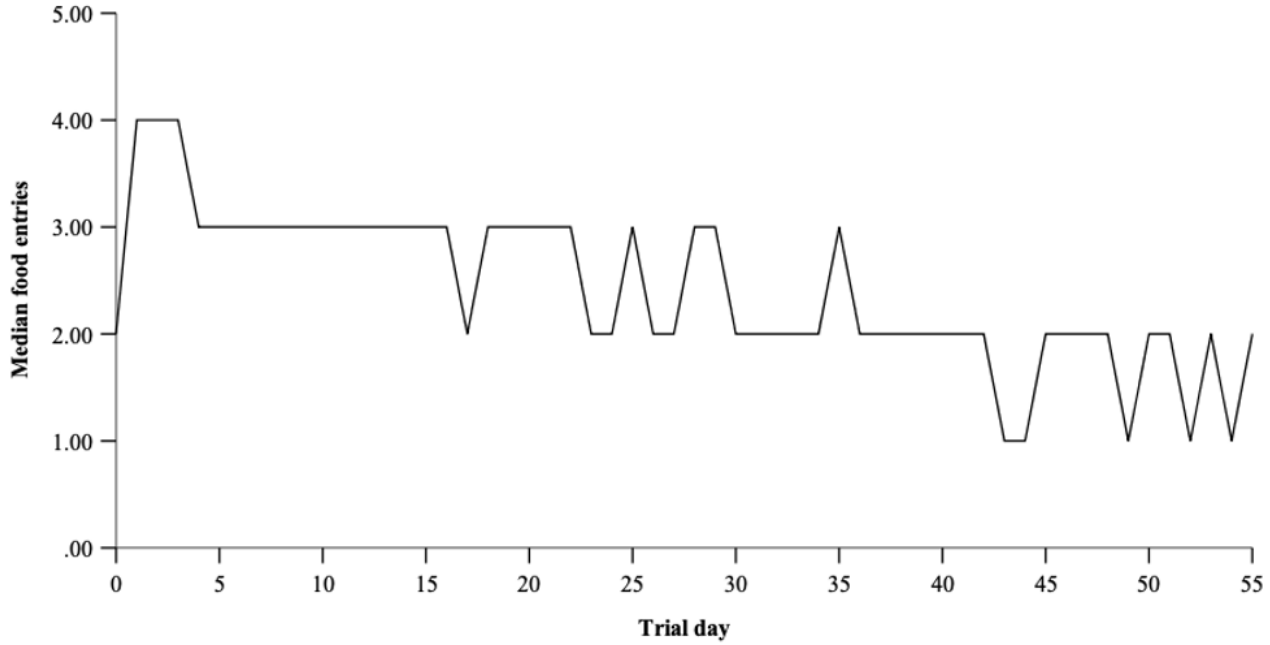
Outcome	Total food entries		Total audio plays		Total gallery views		Percentage of food entries using an image	
	B (95% CI)	*P* value	*B* (95% CI)	*P* value	*B* (95% CI)	*P* value	*B* (95% CI)	*P* value
Weight (kg)	–0.02 (–0.04 to –0.01)	.004	–0.01 (–0.03 to 0.01)	.46	0.04 (–0.002 to 0.08)	.06	0.01 (–0.03 to 0.04)	.66
Body fat (%)	–0.01 (–0.02 to –0.0007)	.05	0.004 (–0.01 to 0.02)	.62	0.01 (–0.02 to 0.04)	.50	0.002 (–0.02 to 0.03)	.84
Self-reported energy intake (kcal)	5.98 (1.50 to 10.46)	.01	–4.59 (–11.12 to 1.93)	.16	–7.55 (–20.49 to 5.38)	.24	4.50 (–5.88 to 14.89)	.38
Taste-test energy intake (kcal)	–0.06 (–0.58 to 0.46)	.82	–0.21 (–0.96 to 0.54)	.58	0.94 (–0.56 to 2.44)	.21	–0.19 (–1.37 to 0.99)	.74

App Use Across the Trial Period

Total number of food entries was the only app usage variable significantly associated with any trial outcomes and was, therefore, plotted to explore usage over the trial period. The median number of food entries recorded per day declined over the trial period, with between three and four food items being recorded per day at the beginning of the trial, down to one to two per day toward the end of the trial (see Figure 2).

XSL•FO
RenderX

Figure 2. Median number of food entries across trial days for participants enrolled in the trial (0-55 days).



Discussion

Principal Findings

We aimed to examine participants' experiences of using an attentive eating smartphone app, and we conducted exploratory analyses into the associations between usage of individual app functions and trial outcomes. Overall usage of app functions was reasonable (ie, total food entries, audio listens, and percentage of meals recorded using an image), except for total food gallery views, which was lower than expected: the median was less than one gallery view per day. Participants reported several barriers and facilitators to app use that could inform future research. Participants found that incorporating the use of the attentive eating app into one's routine meant they were more likely to continue using it. This is in line with Ahtinen and colleagues [11], who found that once use of an app became routine, the app was easy to use. Future research should, therefore, include design of smartphone apps in such a way that facilitates their inclusion into a person's routine; researchers may benefit from encouraging participants to use the app as part of their day-to-day habits early on in order to foster greater app usage throughout a study. The ability to access new and updated content has been found to motivate continued app usage [7,11], which is in line with our finding that boredom with the same audio clip may have led to reduced use of this function of the app. Further, believing that the app was effective was a motivator for continued use in this trial, as it was in another trial [11]. One way for future research studies to demonstrate app effectiveness to participants is to provide feedback on progress toward goals [7,11,29]. For example, feedback on weight loss would demonstrate app effectiveness to participants and likely bolster motivation to continue using the app. This study also found that some participants felt uncomfortable using the app in social situations. There was a notable age difference here, where younger participants were less likely to be concerned about being seen taking photographs of their food in social situations. Therefore, mobile phone apps to promote dietary change will benefit from being designed in ways that overcome concerns about their use in public. Further, some participants found the in-app rewards and badges motivating, whereas others did not. Future work should make such features optional or allow participants to personalize them.

Participants reported effects of using the app that were in line with the hypothesized mechanisms of action of the app. For example, participants reported using the food gallery to introspect on what they had eaten earlier and to inform subsequent food choices. Participants also described eating more slowly and stopping eating before finishing their food due to greater awareness of hunger and fullness, most likely due to the attentive eating audio clip that encouraged listeners to eat slowly and pay attention to their hunger and fullness levels. These findings are in contrast to the results of our main trial analyses, which found no effects on trial outcomes at the 8-week assessment, including weight and energy intake in the full sample and among a subsample of participants that were categorized as having used the app as intended [20]. Our exploratory analyses reported here found that the frequency of recording food consumed was significantly associated with lower body weight at the 8-week assessment but was not significantly associated with body fat percentage. This was a clinically relevant effect, with each additional food entry equating to 0.02 kg of weight loss. Recording four entries per day (ie, three meals and one snack) compared to recording only one entry per day would equate to an additional 3.6 kg of weight loss over a 2-month period. This amount of weight loss would be in line with recommendations on healthy weight loss of 0.5-1 kg per week [30]. We also found that the frequency of recording food consumed was significantly associated with greater self-reported 24-hour energy intake but not with taste-test energy intake. This was a small effect, with each additional food entry recorded equating to an additional 6 kcal consumed. Given the association with weight, we presume this may be a measurement

XSL•FO
RenderX

artefact (ie, participants who diligently recorded meals in the app may have been more likely to diligently complete the dietary recalls).

Further examination of the number of food entries recorded over the entire trial period showed that while the number of food entries recorded was high at the beginning of the trial, it tended to decline over time. The decline in recording of food consumed may partially explain why the trial did not find any effects on main trial outcomes in the planned analyses. This finding is in line with previous work that found sustained use of technology-based interventions to be associated with better health outcomes [5]. More frequent entry of eating episodes allows for more frequent self-monitoring, and this may explain the association with weight, as self-monitoring is a known predictor of weight loss [31]. We found no association between how frequently participants viewed entries or listened to the attentive eating audio clip and weight, which may suggest that the act of recording behavior, as opposed to eating more attentively (ie, remembering food eaten and paying attention during eating), may explain why a greater number of eating episodes was associated with lower body weight. However, an alternative explanation is that more highly motivated participants were more likely to record their consumption episodes and were also more likely to lose weight.

Strengths and Limitations

The use of qualitative research methods to examine participants' experience of an attentive eating smartphone app is a strength of this work, as we have identified barriers and facilitators to using smartphone apps designed for dietary change and weight loss. Although our sample size was reasonable for qualitative analyses [32], it was small for our exploratory quantitative analyses and, for this reason, these results should be interpreted accordingly. Future confirmatory work is needed. A further limitation is that our measures of energy intake are prone to error and may, therefore, not accurately reflect energy intake during the trial, as one was self-reported and the other an objective laboratory measure of consumption of a single type of snack food.

Conclusions

In conclusion, frequent recording of eating episodes using an attentive eating smartphone app was associated with greater weight loss during an 8-week randomized controlled trial. There are a number of barriers and facilitators to frequent use of an attentive eating mobile phone app that may be useful to address when designing dietary behavior change mobile phone apps.

Acknowledgments

This work was supported by the Economic and Social Research Council (project reference: ES/N00034X/1). The funding body had no role in the design of the trial, data collection, analysis and interpretation, or writing the manuscript. PA was supported by a National Institute for Health Research (NIHR) senior investigator award, the NIHR Oxford Biomedical Research Centre, and the Collaboration for Leadership in Applied Health Research and Care (CLAHRC).

Authors' Contributions

All authors contributed to the design of the trial. VW and IK collected and analyzed the data, and VW wrote the manuscript draft. All authors contributed to the manuscript and have read and approved the final manuscript.

Conflicts of Interest

ER and JH have received research funding from the American Beverage Association and Unilever for projects unrelated to this research. PA done research on a weight loss intervention (unrelated to this paper's intervention) funded by Cambridge Weight Plan, and has done half a day's consultancy for Weight Watchers. PA also spoke at a symposium at the RCGP conference funded by Novo Nordisk. None of these activities resulted in payments to this author.

References

1. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: A systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 08;388(10053):1659-1724 [FREE Full text] [doi: 10.1016/S0140-6736(16)31679-8] [Medline: 27733284]
2. NCD Risk Factor Collaboration (NCD-RisC). Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19·2 million participants. Lancet 2016 Apr 02;387(10026):1377-1396 [FREE Full text] [doi: 10.1016/S0140-6736(16)30054-X] [Medline: 27115820]
3. Communications Market Report 2017 - United Kingdom. London, UK: Ofcom; 2017. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0021/105438/uk-internet-online.pdf [accessed 2020-09-13]
4. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: A systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]

XSL•FO
RenderX

5. Mattila E, Orsama A, Ahtinen A, Hopsu L, Leino T, Korhonen I. Personal health technologies in employee health promotion: Usage activity, usefulness, and health-related outcomes in a 1-year randomized controlled trial. JMIR Mhealth Uhealth 2013 Jul 29;1(2):e16 [FREE Full text] [doi: 10.2196/mhealth.2557] [Medline: 25098385]
6. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
7. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: Qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
8. Morrison LG, Hargood C, Lin SX, Dennison L, Joseph J, Hughes S, et al. Understanding usage of a hybrid website and smartphone app for weight management: A mixed-methods study. J Med Internet Res 2014 Oct 22;16(10):e201 [FREE Full text] [doi: 10.2196/jmir.3579] [Medline: 25355131]
9. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers best engage and support users? A qualitative investigation. Br J Health Psychol 2015 Feb;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
10. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. Telemed J E Health 2011 Dec;17(10):801-803. [doi: 10.1089/tmj.2011.0055] [Medline: 22010977]
11. Ahtinen A, Mattila E, Vaatanen A, Hynninen L, Salminen J, Koskinen E, et al. User experiences of mobile wellness applications in health promotion: User study of Wellness Diary, Mobile Coach and SelfRelax. In: Proceedings of the 3rd International Conference on Pervasive Computing Technologies for Healthcare. 2009 Presented at: 3rd International Conference on Pervasive Computing Technologies for Healthcare; April 1-3, 2009; London UK. [doi: 10.4108/icst.pervasivehealth2009.6007]
12. Higgs S, Donohoe JE. Focusing on food during lunch enhances lunch memory and decreases later snack intake. Appetite 2011 Aug;57(1):202-206. [doi: 10.1016/j.appet.2011.04.016] [Medline: 21569808]
13. Robinson E, Kersbergen I, Higgs S. Eating 'attentively' reduces later energy consumption in overweight and obese females. Br J Nutr 2014 Aug 28;112(4):657-661. [doi: 10.1017/S000711451400141X] [Medline: 24933322]
14. Seguias L, Tapper K. The effect of mindful eating on subsequent intake of a high calorie snack. Appetite 2018 Feb 01;121:93-100. [doi: 10.1016/j.appet.2017.10.041] [Medline: 29104106]
15. Allirot X, Miragall M, Perdices I, Baños RM, Urdaneta E, Cebolla A. Effects of a brief mindful eating induction on food choices and energy intake: External eating and mindfulness state as moderators. Mindfulness 2017 Sep 25;9(3):750-760. [doi: 10.1007/s12671-017-0812-0]
16. Whitelock V, Higgs S, Brunstrom JM, Halford JC, Robinson E. No effect of focused attention whilst eating on later snack food intake: Two laboratory experiments. Appetite 2018 Sep 01;128:188-196 [FREE Full text] [doi: 10.1016/j.appet.2018.06.002] [Medline: 29870758]
17. Tapper K, Seguias L. The effects of mindful eating on food consumption over a half-day period. Appetite 2020 Feb 01;145:104495. [doi: 10.1016/j.appet.2019.104495] [Medline: 31629718]
18. Whitelock V, Gaglione A, Davies-Owen J, Robinson E. Focused attention during eating enhanced memory for meal satiety but did not reduce later snack intake in men: A randomised within-subjects laboratory experiment. Appetite 2019 May 01;136:124-129 [FREE Full text] [doi: 10.1016/j.appet.2019.01.021] [Medline: 30690053]
19. Robinson E, Higgs S, Daley AJ, Jolly K, Lycett D, Lewis A, et al. Development and feasibility testing of a smart phone based attentive eating intervention. BMC Public Health 2013 Jul 09;13:639 [FREE Full text] [doi: 10.1186/1471-2458-13-639] [Medline: 23837771]
20. Whitelock V, Kersbergen I, Higgs S, Aveyard P, Halford JCG, Robinson E. A smartphone based attentive eating intervention for energy intake and weight loss: Results from a randomised controlled trial. BMC Public Health 2019 May 21;19(1):611 [FREE Full text] [doi: 10.1186/s12889-019-6923-x] [Medline: 31113400]
21. Whitelock V. A smartphone based attentive eating intervention for energy intake and weight loss: A randomised controlled trial. Open Science Framework. 2019 Mar 27. URL: https://osf.io/btzhw/ [accessed 2020-09-09]
22. Ahern AL, Wheeler GM, Aveyard P, Boyland EJ, Halford JCG, Mander AP, et al. Extended and standard duration weight-loss programme referrals for adults in primary care (WRAP): A randomised controlled trial. Lancet 2017 Jun;389(10085):2214-2225. [doi: 10.1016/s0140-6736(17)30647-5]
23. British Heath Foundation. Facts Not Fads: Your Simple Guide to Healthy Weight Loss. Birmingham, UK: British Heart Foundation; 2014. URL: https://www.bhf.org.uk/~/media/files/publications/healthy-eating-and-drinking/m2_facts_not_fads_weightv2.pdf [accessed 2020-09-13]
24. Carter MC, Albar SA, Morris MA, Mulla UZ, Hancock N, Evans CE, myfood24 Consortium. Development of a UK online 24-h dietary assessment tool: myfood24. Nutrients 2015 May 27;7(6):4016-4032 [FREE Full text] [doi: 10.3390/nu7064016] [Medline: 26024292]
25. Albar SA, Alwan NA, Evans CEL, Greenwood DC, Cade JE. Agreement between an online dietary assessment tool (myfood24) and an interviewer-administered 24-h dietary recall in British adolescents aged 11-18 years. Br J Nutr 2016 May;115(9):1678-1686. [doi: 10.1017/S0007114516000593] [Medline: 26975650]

XSL•FO
RenderX

26. Robinson E, Haynes A, Hardman CA, Kemps E, Higgs S, Jones A. The bogus taste test: Validity as a measure of laboratory food intake. Appetite 2017 Sep 01;116:223-231 [FREE Full text] [doi: 10.1016/j.appet.2017.05.002] [Medline: 28476629]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Whitelock V. User experiences of a smartphone based attentive eating application and their association with diet and weight loss outcomes: Thematic and exploratory analyses from a randomised controlled trial. Open Science Framework. 2020 Jun 19. URL: https://osf.io/tyrk8/ [accessed 2020-09-09]
29. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
30. Start losing weight: Healthy weight. NHS. 2019 Dec 02. URL: https://www.nhs.uk/live-well/healthy-weight/start-losing-weight/ [accessed 2020-09-13]
31. Zheng Y, Klem ML, Sereika SM, Danford CA, Ewing LJ, Burke LE. Self-weighing in weight management: A systematic literature review. Obesity (Silver Spring) 2015 Feb;23(2):256-265 [FREE Full text] [doi: 10.1002/oby.20946] [Medline: 25521523]
32. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide for Beginners. London, UK: SAGE Publications; 2013.

Abbreviations

CLAHRC: Collaboration for Leadership in Applied Health Research and Care
DALY: disability-adjusted life year
NIHR: National Institute for Health Research
VIF: variance inflation factor

Edited by G Eysenbach; submitted 29.10.19; peer-reviewed by J Alvarez Pitti, K Tapper, A Nomura; comments to author 10.03.20; revised version received 19.06.20; accepted 07.07.20; published 02.10.20.

Please cite as:
Whitelock V, Kersbergen I, Higgs S, Aveyard P, Halford JCG, Robinson E
User Experiences of a Smartphone-Based Attentive Eating App and Their Association With Diet and Weight Loss Outcomes: Thematic and Exploratory Analyses From a Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(10):e16780
URL: https://mhealth.jmir.org/2020/10/e16780
doi:10.2196/16780
PMID:33006564

©Victoria Whitelock, Inge Kersbergen, Suzanne Higgs, Paul Aveyard, Jason CG Halford, Eric Robinson. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 02.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Role of Information Technology Mindfulness in the Postadoption Stage of Using Personal Health Devices: Cross-Sectional Questionnaire Study in Mobile Health

Pouyan Esmaeilzadeh[1], PhD

Department of Information Systems and Business Analytics, Florida International University, Miami, FL, United States

Corresponding Author:
Pouyan Esmaeilzadeh, PhD
Department of Information Systems and Business Analytics
Florida International University
Modesto A. Maidique Campus 11200 S.W. 8th St
Miami, FL, 33199
United States
Phone: 1 3053483302
Email: pesmaeil@fiu.edu

Abstract

Background: Although personal health devices (for example, smartwatches, fitness trackers and intelligent bracelets) offer great potential to monitor personal fitness and health parameters, many users discontinue using them after a few months. Thus, it is critical to study the postadoption behaviors of current users to enhance their engagement with personal health devices and use behaviors. However, there is little empirical research on the factors affecting users' engagement in beneficial use behaviors. Mindfulness and identity are not new topics, but the applications of these concepts in the field of information systems are emerging themes. Information technology (IT) mindfulness has been conceptualized in previous studies; however, little is known about the antecedents and consequences of IT mindfulness in the mobile health (mHealth) context.

Objective: The main aim of this study is to explore both IT identity and IT mindfulness to develop a new ground for research in the domain of mHealth postadoption. Thus, we aim to explain why users should be fully mindful of their engagement with PHDs and what could be the consequences and implications.

Methods: This study proposes that IT mindfulness can play an important role in improving the use behaviors of users. Through a web-based survey with 450 current users of a personal health device, this paper tests the relationship between IT identity and IT mindfulness in the postadoption stage of using personal health devices.

Results: We found that IT identity significantly shapes IT mindfulness associated with PHDs. Moreover, the IT identity–IT mindfulness relationship is negatively moderated by individuals' perceived health status (P=.003). Finally, the results of this study show that IT mindfulness can significantly predict automatic use behaviors (eg, continued intention to use), active use behaviors (eg, feature use and enhanced use behaviors), and commitment behaviors in using personal health devices (eg, positive word-of-mouth intention).

Conclusions: The findings of this study provide implications for both research and practice. This study can contribute to our current understanding of IT mindfulness by developing and empirically testing a research model that explains the determinants and outcomes of the IT mindfulness construct in the context of personal health devices. The results imply that IT mindfulness significantly helps individuals express their alertness, awareness, openness, and orientation in the present in their postadoption interactions with smart devices used for health care purposes. Finally, our findings may assist practitioners and IT developers in designing mindfulness-supporting PHDs. Owing to the impact of IT mindfulness on postadoption behaviors, its 4 dimensions could be used for developing PHD technologies. Moreover, PHD developers may need to direct their efforts toward increasing IT mindfulness by reinforcing IT identity to serve and retain a wide range of target users.

(JMIR Mhealth Uhealth 2020;8(10):e18122) doi:10.2196/18122

KEYWORDS

IT identity; IT mindfulness; personal health devices; perceived health status; post-adoption behaviors; mHealth; mobile phone

XSL•FO
RenderX

Introduction

Background

Mobile health (mHealth) has tremendously changed how people are involved in performing different roles in their social relations [1]. In recent years, the use of personal health devices (PHDs) as platforms for monitoring and controlling health conditions has grown significantly. Many manufacturers (eg, Samsung, Jawbone, and Fitbit) have developed smartwatches and intelligent bracelets to help individuals fulfill their health-related goals. Use of information technology (IT) allows individuals to feel competent and accepted in social networks of relationships in which their roles are defined based on cultural expectations and norms [2]. For instance, PHDs can enable individuals to monitor their health status, manage chronic diseases, control fitness as well as wellness, and view personal health information in real time [3]. In this study, PHDs are defined as smart devices used for monitoring and controlling health status, not for providing diagnoses and medical treatment remotely. People may continuously engage with their PHDs to monitor their health status or check their health care information [4]. Thus, the use of PHDs is beneficial to people because these devices provide them with additional resources in managing personal health care, which is highly important to them. These devices can verify identity by providing broad applications across a wide range of technology networks and social contexts. These smart devices can be integrated into an individual's sense of self, as one may spend a significant portion of their time interacting with the devices as a repeated behavior. This is consistent with the findings of Reychav et al [5] that repeated behaviors can directly contribute to identity construction.

Technology can be manifested in different ways, for instance, technology interface or what the technology can afford. This study does not focus on the technology interface, but we examine technology in terms of how people use the platform to identify the underlying application, purpose, and function. Individuals may interact with many technologies daily but may consider only certain IT devices as an inherent part of themselves, which can influence their behavioral choices [6]. A PHD (eg, a wearable smart device or an intelligent bracelet) is a consumer IT, which could become part of people's identities because of its everyday use for self-monitoring and self-management purposes [7]. The adoption and the use of patient self-management tools are consistent with the notion of patient-centered eHealth apps, which revolves around patients as pivotal actors in the health care ecosystem [8]. However, recent studies highlight that many users discontinue using PHDs after only a few months [9]. Little is known about the factors that motivate users to actively use such electronic devices and explore more features and functions for additional health-related tasks [10]. Thus, further research is required to investigate individuals' beliefs and behaviors during the postadoption of PHDs to uncover users' requirements and preferences for eHealth services and to propose suggestions for the construction of effective smart devices, interactive mHealth apps, and efficient eHealth systems.

To fill this gap, we use the concept of IT mindfulness as a theoretical foundation to delineate how PHD users perform postadoption behaviors. Mindfulness is a multidimensional concept that explains constant examinations, expectation refinement, recognition of new changes, and exploration of novel aspects of a phenomenon [11]. According to Weick and Sutcliffe [12], mindfulness refers to focusing on the present, paying attention to details, considering alternative perspectives, and willingness to investigate for understanding system failures. In line with a study by Ndubisi [13], people who are mindful are very sensitive to different contexts and have the ability to continuously create new categories in awareness and interpretation of the world. In contrast, individuals who engage in mindless behaviors have a premature commitment to rigid beliefs and preexisting categories [14].

Mindfulness promotes being in a watchful and vigilant state of mind [14]. A study by Sun et al [15] suggests that IT mindfulness may influence technology adoption decisions and continuance. Their results indicate that mindful users continuously compare an IT with existing technologies to be aware of its uniqueness and make more rational postadoption decisions in the light of task-technology fit. A mindful process of experimenting with IT can articulate how individuals extract value from technologies [16]. IT mindfulness may reduce the addictive use of technology and help individuals feel comfortable with their use of IT systems. Previous research indicates that domain-specific mindfulness (eg, in situations related to personal health) could predict individual behaviors [17]. IT mindfulness, as an IT-specific trait, influences users' behaviors with a given technology [18]. As IT mindfulness is considered a dynamic IT-specific trait, this class of trait likely influences technology-related behaviors such as user acceptance of technology and postadoption use behaviors.

Previous studies provide significant evidence to show that IT mindfulness shapes future behavior by explaining how awareness and flexibility of users can alter their future interactions with the IT system [19]. People with high levels of mindfulness tend to continually monitor the current situation to find new ways of using IT, which helps them complete their tasks [20]. Beliefs and behaviors are affected by personality traits (at different hierarchical levels). Previous studies indicate that dynamic and context-specific traits (rather than broad traits) are more likely to change specific behaviors [21]. However, little is known about the effects of information system (IS)–specific traits (such as IT mindfulness) on IT-specific beliefs and use behaviors. Moreover, there is an increase in studying the causes and consequences of mindfulness in different research areas [22]. However, most of these studies are conducted in the organizational context and are considered organizational factors and interventions [22]. Another stream of research focuses on mindfulness in a collective context, such as a group setting, rather than examining mindfulness at the individual level [23]. Thus, more studies are required to examine the antecedents and resultant effects of IT mindfulness at the individual level in the context of using consumer IT.

XSL•FO
RenderX

Objectives

The main goal of this study is to highlight the ways in which the application of the IT mindfulness concept in studies of mHealth device design and use can contribute to the realization of its antecedents and outcomes in the PHD context. We propose that IT mindfulness may provide a possible basis for answering questions about how individuals can hope to efficiently use smart devices to achieve reliable health-related results. We argue that positive use behaviors are possible when a mindful approach permeates an individual user. In doing so, we develop a research model by drawing on the recent appearance of the concept of mindfulness in the IS literature and adapting it for application to postadoption of PHDs. In brief, this study addresses the following research questions:

1. How does IT mindfulness influence individuals' postadoption interactions with PHDs?
2. What are the antecedents of IT mindfulness in the context of PHDs?
3. How do health factors (ie, perceived health status) moderate the influencing chain from IT identity to IT mindfulness?

Literature Review

IT Mindfulness

Mindfulness is a psychological trait that has roots in the cognitive abilities of the individual [20]. Mindfulness has been used in IS studies with different themes, such as IT innovation, IT management, IT use, and outcomes [24]. Swanson and Ramiller [20] suggest that the concept of mindfulness can be incorporated into the adoption, implementation, and assimilation of an IT innovation. As a result, IT mindfulness arises when people are working with IT. When people are aware of IT capabilities and open to its various functions, they elevate their mental mindset to become mindful of value-adding applications of IT. People with cultivated IT mindfulness are likely to focus on the present IT functionalities, search for more details about its applications, explore more uses of IT, and examine IT features [25]. When a person mindfully accepts a technology, they are aware of the given technology, its functions, and their needs. Thus, they are more likely to search for more details and information about the technology and the acceptance decision and possible implications [26].

Sternberg [27] describes the concept of mindfulness as (1) alertness to distinction, (2) awareness of multiple perspectives, (3) openness to novelty, and (4) orientation in the present. Consistent with previous studies, we consider IT mindfulness as a second-order construct that consists of 4 reflective dimensions [25,28]. The first first-order dimension is alertness to distinction, which refers to the ability to define, appreciate, and make judgments about IT applications and their potential. This factor helps individuals identify the differences between the old and new features of an IT application and seek new ways to use the system. The second dimension is awareness of multiple perspectives, which helps individuals analyze IT system applications and features from different or even opposing viewpoints. This dialectical thinking may lead to innovative solutions to IT-related problems. The third dimension is the openness to novelty, which involves curiosity and flexibility in a user's interactions with an IT system's features and applications. This factor is instrumental in cultivating in a user the ability to reason about new types of stimuli, consider a large number of IT applications, and explore fewer familiar features. The last factor is an orientation in the present, which manifests the degree to which an individual pays more of his/her attention to his/her current situation instead of envisioning future possibilities or concentrating on past events. This factor may increase people's sensitivity to the immediate context and adapt their responses and system used to the current situation. As a second-order construct with reflective dimensions, IT mindfulness requires capturing all 4 dimensions. Higher levels of IT mindfulness will lead to greater levels of the 4 dimensions. We cannot assume that a change in IT mindfulness will lead to the same amount of change across the 4 dimensions [29]. Although the indicators of dimensions may covary, each dimension has a separate conceptual foundation [30].

IT Identity

Prior researchers have studied the topic of IT and identity and their relationship based on different approaches. Carter and Grover [31] conceptualize IT identity using theories on social structures and self-concept to describe how people categorize themselves in relation to an IT object. They define IT identity as the extent to which the use of IT is saliently related to who people think they are (self-identification). IT can change individuals' self-perceptions by recognizing their original selves in using the capabilities and resources offered by the IT device. For instance, in the presence of IT, people may feel empowered, productive, autonomous, and accessible. According to Carter et al [32], adults' interactions with their mobile phones lead to enhanced perceptions of empowerment, self-authenticity, and autonomy. In the context of PHDs, the self-concept and personal resources of users can be expanded by the capabilities provided by the devices. For example, wearable smart devices can be used to reduce time and place constraints in controlling fitness and wellness, measuring different physical changes, and handling emergency cases. These functionalities may enhance individuals' original self-perceptions and may make them feel independent, empowered, and smart.

Previous studies on IT identity suggest that the construct of IT identity is expressed through 3 first-order factors: relatedness, dependence, and emotional responses of individuals in relation to IT [31]. Relatedness refers to a sense of connection felt when interacting with an IT device. For instance, a strong feeling of connection with an IT device can turn individuals' perceptions about the self to what they can do with the IT [33]. Emotional energy indicates the levels of emotional attachment, enthusiasm, and confidence that an individual attributes to an IT when thinking about his/her interaction with it. Long-term interactions with an IT system can raise individuals' levels of emotion and confidence in relation to the IT experienced and help them to be more spontaneous with the IT [34]. Dependence explains the extent to which people rely on IT to represent their self-perceptions. For instance, people rely on digital forms of communication to manage their interactions and relationships with others to satisfy social expectations [35].

XSL•FO
RenderX

Hypothesis Development

IT Identity and IT Mindfulness

Using the hierarchical structure of personality traits, broader IS-specific traits (ie, qualities or characteristics belonging to a user that developed and were enacted because of using a particular technology) can influence narrower dynamic traits [36]. For instance, previous studies suggest that context-specific traits are closely related to dynamic IT-specific traits (such as IT mindfulness) [22]. Owing to the malleability of this trait, personal interactions with technology may nurture IT mindfulness and encourage users to gain more value from PHDs. In line with this rationale and based on the scope of IT identity and IT mindfulness, we propose that IT identity will help shape IT mindfulness. IT identity is a broad trait because it is the first step that users will attach dependence, emotion, and relatedness to a technology. This IS-specific trait may encourage users to gather more information about the applications and functions of that technology, analyze its applications and features, and seek new ways to use the system.

Users who hold a higher level of IT identity associated with a PHD will be generally dependent on their device, have an overall sense of connection when interacting with it, and develop emotional energy toward using it. IT identity can change individuals' self-perceptions by recognizing their original selves in using the capabilities and resources offered by the IT device. For instance, in the presence of IT, people may feel empowered, productive, autonomous, and accessible. These feelings function broadly because of the sense of independence, and the feeling of being smart or productive can nourish their overall self-perceptions [37]. The self-concept (which is influenced by IT) can then reinforce the users' readiness to increase their understanding of the functions of a specific IT, vigilance to its differences, sensitivity to its current task context, and curiosity to its features. Therefore, using the hierarchical structure of personality traits, we contend that IT identity may be placed at a higher hierarchical level, and it could be an antecedent of shaping IT mindfulness.

A PHD user with a strong IT identity is highly dependent on his/her device [38]. This leads the IT identity holder to be greatly alert to distinction and such a user tends to identify new ways to accomplish health-related tasks by using his/her device (alert to the distinction). A higher level of IT identity makes users have a strong feeling of connection with a PHD. This makes them more willing to get involved when using their PHD and to keep a constant eye on the big picture to differentiate between usage contexts (orientation to the present) [25]. Users with strong IT identity develop higher levels of emotional attachment, enthusiasm, and confidence that they attribute to a PHD when thinking about their interaction with it [31]. This feeling enables users to be open to new ways of using a PHD and be more eager to learn new ways of using it (awareness of multiple perspectives). IT identity holders rely highly on a PHD to manage their health metrics to satisfy their personal health-related expectations. This self-concept encourages users to explore new potential or features within their device (openness to novelty) [18].

Interacting with an IT device that contributes to self-identification can be a considerable opportunity for IT users. When a person considers a PHD as an integral part of the self, he/she may quickly recognize the capabilities and potential of a PHD, the differences between its features, and the utility of its new/updated features. A strong IT identity can motivate users to attribute higher levels of emotional attachment, enthusiasm, and confidence in his/her medical device and, in turn, become involved in the PHD usage through constant adoption of its new features. An IT identity holder may seek more information, increase his/her knowledge about its features, and hold diverse perspectives toward the potential usage of a PHD and become inspired to develop innovative solutions using these different perspectives. A higher level of IT identity may impart more confidence to the users to rely on PHDs to explore new features within a PHD and be flexible to new features of their PHD. Thus, we posit that IT identity affects the degree to which users develop mindfulness in using their PHDs.

Our first hypothesis is as follows:

- H1: IT identity positively influences the IT mindfulness of PHD users.

The Moderating Role of Perceived Health Status

According to Bansal and Gefen [39], individuals' characteristics (such as current health status) significantly affect the way they analyze the utility of using IT. Perceived health status is a common health factor that highlights overall individual health [40]. Previous research highlights the effects of perceived health status in different contexts. For instance, personal health status evokes privacy concerns related to health information disclosure and the tools used to share such information [41]. Previous research highlights that mindfulness is positively related to higher levels of well-being [42]. Most studies on mindfulness suggest that mindfulness will lead to pleasant psychological effects. For instance, acting mindfully is closely related to improved psychological health [43]. Brown and Ryan [44] reported that mindfulness could improve self-esteem and optimism and reduce anxiety. Cash and Whittingham [45] demonstrated a negative relationship between mindfulness and depression. A study by Raes et al [46] also showed that mindfulness positively influences cognitive reactivity. A review study contends that mindfulness-oriented interventions are significantly associated with positive psychological effects, such as better subjective well-being, reduced psychological distress, and improved behavioral regulation [47].

However, little is known about whether the overall evaluation of a person about his/her health status can affect the development of IT mindfulness and IT identity. We argue that perceived health status may change the way people think about themselves, their capabilities, and the world around them. Unhealthy individuals (eg, with chronic diseases) perceive more strain because of the presence of physical/mental infirmity, and this health condition makes them more anxious and vulnerable to the digital devices surrounding them. People in good health tend to assume less severe demands on their strength or abilities, perceive more control, and experience less mental/emotional strain [48].

XSL•FO
RenderX

Therefore, the impact of IT identity on IT mindfulness in a given situation may depend on a user's perception of his/her health status. We expect that the IT identity–IT mindfulness relationship may vary depending on the current health status of users. In doing so, the overall sense of well-being can influence the level of PHD users' attachment, reliance, and dependence on their devices to perform health-related tasks. In turn, it may induce changes in the degree of awareness of various features, attention to the present moment experience, alertness to differences, and openness to new information. On the basis of the moderating effect of perceived health status, we posit that people with high IT identity levels may not always tend to remain high in IT mindfulness, and their perceptions about their well-being may change the strength and direction of this relationship. For instance, in the PHD context, an IT identity holder may think of using the body temperature function, high heart rate notifications, or insulin delivery features to monitor their health status rather than making an appointment with a physician. By doing so, he/she will enact his/her identity as a competent person who can leverage a PHD to keep records of his/her health conditions. However, current studies cannot answer whether this high IT identity will always be translated into high IT mindfulness. Thus, we propose that the IT identity–IT mindfulness linkage could be moderated (augmented or attenuated) by perceived health status.

Our second hypothesis is as follows:

- H2: Perceived health status moderates the relationship between IT identity and IT mindfulness in using PHDs.

IT Mindfulness and Postadoption Behaviors

In line with previous research [44], mindfulness is a state of consciousness that facilitates the fulfillment of basic psychological needs. In turn, mindfulness is a good predictor of self-regulated behaviors. Consequently, IT mindfulness as an IT-specific trait can be used to study IT-related beliefs and behaviors. Previous studies examined the impact of mindfulness on the formation of people's beliefs about using technology [26]. For instance, IT mindfulness strongly shapes the perception of technostress [22]. Previous research demonstrates the significant effects of mindfulness on innovating with IT [20]. Thatcher et al [25] demonstrated that IT mindfulness significantly influences deep structure usage and attempts at innovation.

We also tested the relationships between IT mindfulness and important systems use constructs. In this study, the postadoption system use behavior is represented by feature use, enhanced use, continued intention, and positive word-of-mouth (WOM) intention. Active system use refers to a situation in which people ponder the system and knowingly modify how to use it [49]. In previous studies, automatic system use usually implies habitual behaviors where people use the system unconsciously without thoughtful assessments and focused analyses related to their use [50]. However, in this research, automated system use reflects continued intention to use a PHD based on a mindful consideration of alternatives, not the addictive use of technology. Therefore, more active system use is denoted by feature use as well as enhanced use behaviors, and automated system use is represented by continued intention to use.

According to Sun et al [15], mindful IT users are more willing to use different features of IT. When being mindful, the user is more likely to actively explore and discover additional useful features and functions of a technology [26]. Individuals who are mindfully engaged in a health-related task using a PHD are more motivated to explore a wide range of perspectives. Engagement in feature use behaviors requires sharp user alertness and dynamic awareness of how the use of various features and applications can contribute to task completion [51]. Involvement in enhanced use entails users to explore previously unused features of a PHD to use it for performing additional tasks [52]. Mindfulness helps people scan the context for interpreting the context-relevant information of all conditions [14]. People with higher levels of mindfulness tend to know their context as well as their ability, and they are more open to deliberately search for new features to complete further health-related tasks.

Hypotheses 3 and 4 are as follows:

- H3: IT mindfulness positively influences feature use behavior of PHD users.
- H4: IT mindfulness positively influences enhanced use behavior of PHD users.

Continued intention implies the extent to which people are likely to use familiar technology in the future. Mindful thinking increases people's willingness to process information and continue to use the features of IT in an alert and open way [28]. Mindful PHD users are more likely to have a sense of control when using this technology because they clearly know what they can and cannot do with their smart devices. According to Wong et al [53], mindfulness indicates being open to new information about the technology at hand, being aware of various perspectives, and being involved in the continuous creation of options. Mindful people are more likely to be receptive to new information and compare the technology being used with others [26]. In addition, mindful users are more likely to recognize all the consequences of their decisions (eg, both the pros and cons). Given that they have more information about the system, as long as the current technology seems beneficial, mindful users are likely to continue using the same system compared with individuals who use the system mindlessly.

Hypothesis 5 is as follows:

- H5: IT mindfulness positively influences continued intention to use PHD.

Mindfulness reinforces learning from interpreting related outcomes [24]. When acting in a mindful way, one pays more attention to every detail of IT applications at hand and becomes sensitive to the context. By exploring new aspects of IT and understanding its capabilities and potentials, the user will be open to resolving any challenging situation to accomplish his/her tasks more effectively. Thus, IT mindfulness may positively influence user satisfaction with the technology used to accomplish his/her tasks [18]. Consistent with Fiol and O'Connor [14], mindful users actively analyze how a PHD fits their own contexts and needs rather than blindly follow others in using it. As a result, they may be more inclined to describe the features and functions of PHDs to others and encourage them to use

XSL•FO
RenderX

these smart devices to fulfill their health-related needs. Achieving a fit between the technology and the task may encourage users to be committed to PHDs and become more likely to make positive comments about the mobile system they are using. Thus, we hypothesize that IT mindfulness can enhance positive postadoption behaviors (ie, active use behaviors, continued use behaviors, and positive WOM intention of PHD users).

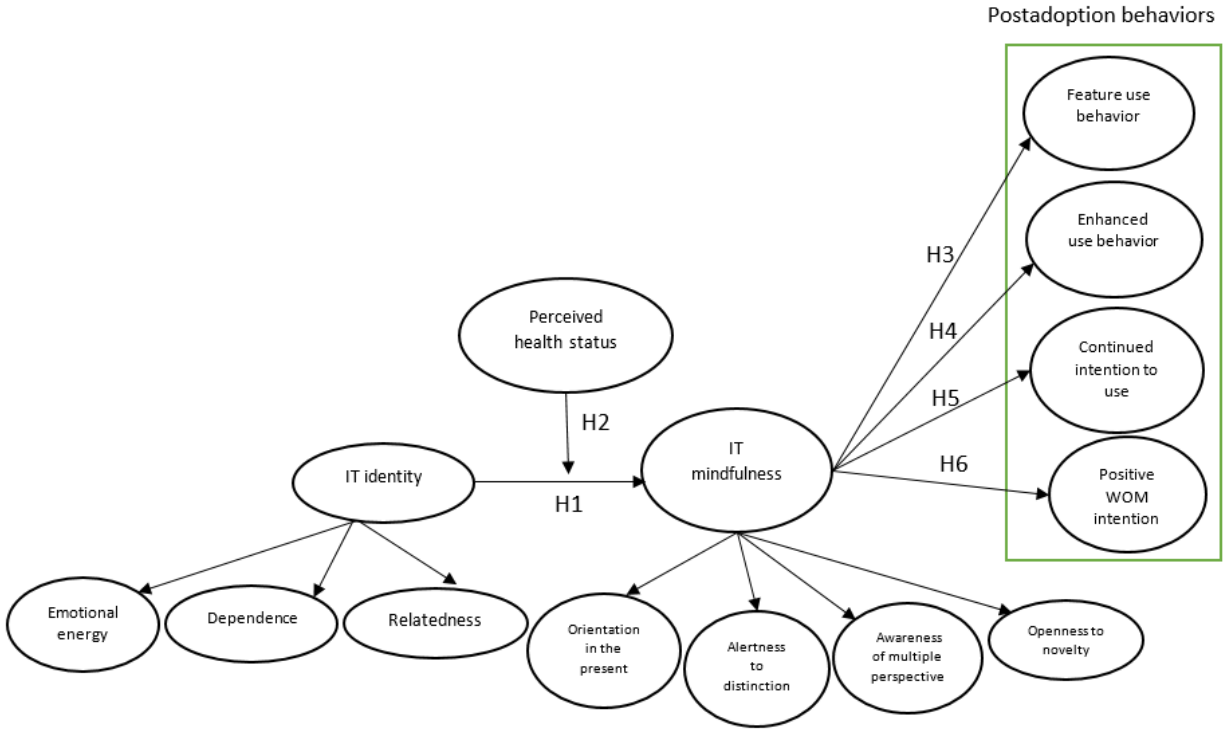
Hypothesis 6 is as follows:

- H6: IT mindfulness positively influences the positive WOM intention of PHD users.

Research Model

We bring IT identity and IT mindlessness together in a theoretical synthesis in which these concepts are seen to interact in ways that help shape the postadoption behaviors of PHD users. The research model indicates that IT identity with PHDs can build IT mindfulness and, in turn, will lead to positive postadoption use behaviors. However, according to Carter and Grover [31], individuals do not always attempt to use IT to exhibit who they are to others even when the IT is advantageous to them. Moreover, previous studies suggest that mindfulness skills are significantly related to aspects of health status [54]. Thus, we suggest that the verification of IT identity–IT mindfulness linkage in relation to the use of PHDs may depend on the health status of people. As the probability of IT identity and IT mindfulness may be evoked by perceived health status, the relationships between IT identity and IT mindfulness can be moderated by this health factor. Figure 1 shows the proposed research model.

Figure 1. Research model. IT: information technology; WOM: word-of-mouth.



Methods

Definition of Variables

The final measure items (ie, all the items included in the survey) are listed in Multimedia Appendix 1. Table 1 provides the definitions of the constructs used in this study.

XSL•FO
RenderX

Table 1. Operationalization of variables.

Construct	Construct definition	Source
IT[a] identity	The degree to which the use of an IT (ie, a PHD[b]) is meaningfully related to who people think they are (self-identification).	Carter and Grover [31]
IT mindfulness	The degree to which a user is involved in the present context, alert of details, aware of other potential uses, and open to investigating IT (ie, a PHD) features and failures.	Thatcher et al [25]
Perceived health status	The extent to which an individual believes that the overall status of his/her health and wellness is good.	Bansal and Gefen [39]
Positive word-of-mouth intention	The degree to which a user shares a positive assessment of his/her experience with a PHD with others.	Maxham III [55]
Continued intention to use	The degree to which a user feels he or she will keep using a PHD.	Bhattacherjee [56]
Feature use behavior	The extent to which an individual uses various features and functions of a PHD in different situations.	Lucas Jr and Spitler [57]
Enhanced use behavior	The extent to which an individual explores previously unused features of a PHD to use it for performing additional tasks.	Bagayogo et al [52]

[a]IT: information technology.

[b]PHD: personal health device.

Measurement of Variables

This study drew on the existing literature to measure the constructs included in the model, and minor changes were made to the instrument to fit the PHD context. Consistent with prior studies, we consider IT identity as a second-order construct with 3 reflective factors [58]. The rationale behind this measurement is that IT identity is reflective of the 3 dimensions as well as the expected interactions among them. Therefore, all these dimensions can reflect the same theme and may covary. To measure the 3 interrelated dimensions of IT identity (ie, relatedness, emotional energy, and dependence), we adapted the items reported in Carter and Grover [31]. According to Kayhan [59], reflective modeling is a better option than formative when first-order factors are expected to interact, correlate, or share a common theme. Thus, a set of interrelationships among these factors is an essential component for measuring IT identity. For instance, dependence, which defines a person's sense of reliance on a PHD, may be related to the emotional energy dimension that describes the feelings of attachment in relation to the device.

In line with a study by Langer and Ngnoumen [60], IT mindfulness is also modeled as a second-order construct composed of 4 reflective first-order dimensions: (1) alertness to distinction, (2) awareness of multiple perspectives, (3) openness to novelty, and (4) orientation in the present. Thus, each dimension has a distinct conceptual foundation, and the items of these 4 dimensions may covary and become interchangeable. On the basis of reflective modeling, individuals with higher levels of IT mindfulness are more likely to exhibit higher levels of alertness to distinction, awareness of multiple perspectives, openness to novelty, and orientation in the present. A change in the IT mindfulness construct may not lead to the same amount of change across all 4 dimensions [29]. Thus, in this study, IT mindfulness is operationalized as a construct that requires capturing all 4 dimensions. We adapted reflective items from a study by Thatcher et al [25] to measure the 4 dimensions of IT mindfulness.

The outcome variables studied in this research are feature use behavior, enhanced use behavior, continued intention to use, and positive WOM. These outcome variables are considered as different types of postadoption behaviors. Items reflecting feature use behaviors were adapted from a study by Lucas and Spitler [57], and items measuring enhanced use behaviors were adapted from a study by Bagayogo et al [52]. Items measuring continued intention to use were adapted from studies by Venkatesh and Goyal [61] and Bhattacherjee [56]. We adapted the items reported in a study by Hoehle and Venkatesh [62] to measure positive WOM intention. Finally, items measuring perceived health status were adapted from the scale developed by Bansal and Gefen [39].

Participant Recruitment

Data were collected in October 2019 from Amazon's Mechanical Turk (MTurk) to obtain a representative group of subjects in the United States. As PHDs may not be considered as a routine technology for many individuals, to obtain more robust, reliable, and applicable findings, we specified 2 more qualifications that individuals had to meet to participate in the survey. First, we defined a screening question to only include individuals who have been using a PHD. We attempted to distinguish between PHDs (hardware device) and mHealth apps (software app) and only include users of a mobile smart device. The logic behind this screening is that we defined IT as a unit of technology (hardware device and software app). Accordingly, the dimensions of IT mindfulness and IT identity can be properly measured and examined, resulting from interactions with devices (as a unit) and not through interaction with application environments or software apps. For example, participants of this study were users of any PHDs (such as wearable smart devices, wearable activity monitors, and intelligent bracelets). When individuals use a PHD, they are exposed to their features and characteristics. Therefore, the likelihood that they become more familiar with its functions and mechanisms is greater, and IT mindfulness as well as IT identity are more likely to be enacted. Thus, we ensured that the participants had used PHDs when they took part in this study. The incentive for participation

XSL•FO
RenderX

was a monetary reward (US $3). At the beginning of the web-based survey, a detailed description of PHDs was provided to ensure that respondents completely comprehended the context and purpose of the study. In total, 462 individuals attempted the survey.

Second, as mentioned in previous studies, one general concern in data collection is a potential lack of attention and random responses [63]. Consistent with other studies, we used some attention trap questions to prevent and identify careless, hurried, or haphazard answers [64]. On the basis of answers to these attention-trap questions, 12 responses were dropped. This ratio is similar to those reported in previous studies that used MTurk for data collection [65]. Thus, concerns that web-based respondents might reply randomly or haphazardly to complete the survey quickly were alleviated. After excluding responses that failed the response quality questions, the final sample consisted of 450 usable and valid questionnaires. We also used Mplus to assess the power of the analysis and determine the sample size [66]. Given the number of observed and latent variables in the model, the anticipated effect size (0.3), the desired probability (0.8), and statistical power levels (α=.05 and power β=.95), the minimum sample size for the model structure is 400. Therefore, this study was adequately powered, as 450 respondents could be sufficient to reduce possible sampling errors and minimize type 2 errors. This is consistent with both the ratio of indicators to latent variables approach and the function of minimum effect, power, and significance suggested by Westland [67].

Results

Descriptive Statistics

Table 2 depicts the respondents' characteristics. The demographic characteristics show that most respondents were female (270/450, 60.0%), White (301/450, 66.9%), with a full-time job (311/450, 69.1%), and had a bachelor's degree (257/450, 57.1%). Approximately 69.7% (314/450) of respondents were aged between 20 and 39 years, and approximately half had an annual household income between US $25,000 and US $74,999. Regarding experience, frequency, and length of use, the results imply that the respondents of this study were familiar with a PHD. All participants had used a PHD before, and most (276/450, 60.8%) rated themselves as either very experienced or extremely experienced with an mHealth device. Overall, 62.0% (279/450) of respondents used PHD daily, and approximately 52.2% (235/450) used PHD for more than a year. Finally, respondents were asked to indicate the type of their PHDs and the purpose of using them. Fitbit (176/450, 39.1%), Apple Watch (90/450, 20.0%), and Samsung Galaxy Fit (81/450, 18.0%) were the top 3 PHDs used by respondents. Controlling fitness and diet (203/450, 45.1%) followed by monitoring blood pressure and checking the cholesterol level (122/450, 27.1%) received the highest percentage of responses regarding the purpose of use.

XSL•FO
RenderX

Table 2. Sample characteristics (N=450).

Variables	Value, n (%)
Gender	
Male	180 (40.0)
Female	270 (60.0)
Age (years)	
<20	5 (1.1)
20-29	157 (34.9)
30-39	157 (34.9)
40-49	82 (18.3)
50-59	36 (8.0)
≥60	13 (2.9)
Annual household income (US $)	
<25,000	72 (16.0)
25,000-49,999	115 (25.5)
50,000-74,999	112 (24.8)
75,000-99,999	75 (16.6)
≥100,000	76 (16.8)
Education	
Less than high school	22 (4.9)
High school graduate	45 (10.0)
Some college	77 (17.1)
2-year degree	35 (7.7)
Bachelor's degree	257 (57.1)
Graduate degree	14 (3.1)
Employment status	
Employed full time	311 (69.1)
Employed part time	73 (16.3)
Unemployed	33 (7.4)
Retired	10 (2.2)
Student	23 (5.1)
Race and ethnicity	
White	301 (66.9)
African American	51 (11.4)
Asian	27 (6.0)
Hispanic	66 (14.6)
Mixed	5 (1.1)
Experience with mobile devices (eg, phone, tablets)	
Slightly experienced	8 (1.7)
Moderately experienced	77 (17.1)
Very experienced	145 (32.3)
Extremely experienced	220 (48.9)
Experience with PHDs[a]	
Slightly experienced	40 (8.9)

XSL•FO
RenderX

Variables	Value, n (%)
Moderately experienced	134 (29.7)
Very experienced	172 (38.3)
Extremely experienced	104 (23.1)
Frequency of use	
Rarely	8 (1.7)
Monthly	44 (9.7)
Weekly	119 (26.6)
Daily	279 (62.0)
Length of use	
<6 months	84 (18.6)
6 months to 1 year	131 (29.1)
1-2 years	110 (24.6)
>2 years	125 (27.7)
PHDs used by participants	
Fitbit	176 (39.1)
Apple Watch	90 (20.0)
Samsung Galaxy Fit	81 (18.0)
FitTech Smart Watches	58 (12.8)
Garmin Fitness Watches	27 (6.0)
Other Smart Fitness Trackers	18 (3.9)
Purpose of use	
Controlling fitness and diet	203 (45.1)
Monitoring blood pressure and checking the cholesterol level	122 (27.1)
Controlling or quitting smoking	54 (12.0)
Monitoring chronic diseases (eg, diabetes and heart disease)	45 (10.0)
Controlling depression or anxiety	26 (5.8)

[a]PHD: personal health device.

Instrument Validations

Before data were statistically analyzed, normality was evaluated, as this is important for the distribution of data to exhibit this trait, to facilitate unbiased and consistent models [68]. Thus, all the constructs used in the model were scrutinized against the normality assumptions. An examination of the skewness and kurtosis of the constructs showed a skewness range from 0.045 to 1.164 and a kurtosis range from 0.017 to 1.531. On the basis of these findings, all the values fall within the prescribed limit and maximum acceptable levels of 2 for skewness and 7 for kurtosis tests [69].

To test the proposed research model, we apply a two-step assessment process using SmartPLS: measurement model and structural model assessments [70]. The SmartPLS method simultaneously assesses the theoretical propositions and properties of the underlying measurement model. To validate the survey instrument, we performed a confirmatory factor analysis on all the constructs to assess the measurement model. We used SmartPLS (version 3.0) to test the convergent and discriminant validity. According to Gefen et al [71], convergent validity can be tested by examining the standardized factor loading, composite reliability, and average variance extracted (AVE). Table 3 shows the results of the convergent validity test. All values of composite reliabilities were more than the threshold value of 0.7, which highlighted that the reliability of the constructs was adequate [72]. According to Hair et al [73], a factor loading of 0.7 or greater is acceptable. In this study, all reported standardized factor loadings were greater than 0.7. The AVE of each construct was calculated using standardized factor loadings. All reported values of the AVE were also greater than 0.5, which met the minimum requirement [74]. These measures indicated that the convergent validity of the measurement model was acceptable. As the instrument validation results were satisfactory, the scales were not purified, and no items were excluded from further analysis. Thus, Table 3 includes all items used in the questionnaire.

XSL•FO
RenderX

Table 3. Results of convergent validity.

Construct and items	Standardized factor loading (>0.7)	Composite reliability (>0.7)	Average variance extracted (>0.5)
IT[a] identity			
ITI–REL1[b]	0.83	0.923	0.706
ITI–REL2	0.80	N/A[c]	N/A
ITI–REL3	0.87	N/A	N/A
ITI–REL4	0.86	N/A	N/A
ITI–REL5	0.84	N/A	N/A
ITI–EMO1[d]	0.80	0.917	0.689
ITI–EMO2	0.85	N/A	N/A
ITI–EMO3	0.83	N/A	N/A
ITI–EMO4	0.82	N/A	N/A
ITI–EMO5	0.85	N/A	N/A
ITI–DEP1[e]	0.86	0.920	0.698
ITI–DEP2	0.87	N/A	N/A
ITI–DEP3	0.76	N/A	N/A
ITI–DEP4	0.79	N/A	N/A
ITI–DEP5	0.89	N/A	N/A
IT mindfulness			
ITM–ALT1[f]	0.79	0.870	0.690
ITM–ALT2	0.82	N/A	N/A
ITM–ALT3	0.88	N/A	N/A
ITM–AW1[g]	0.80	0.869	0.690
ITM–AW2	0.86	N/A	N/A
ITM–AW3	0.83	N/A	N/A
ITM–OP1[h]	0.90	0.917	0.787
ITM–OP2	0.86	N/A	N/A
ITM–OP3	0.90	N/A	N/A
ITM–OR1[i]	0.80	0.836	0.629
ITM–OR2	0.79	N/A	N/A
ITM–OR3	0.79	N/A	N/A
Perceived health status			
PHS1[j]	0.80	0.918	0.737
PHS2	0.89	N/A	N/A
PHS3	0.84	N/A	N/A
PHS4	0.90	N/A	N/A
Feature use behavior			
FEAT1[k]	0.86	0.895	0.681
FEAT2	0.81	N/A	N/A
FEAT3	0.80	N/A	N/A
FEAT4	0.83	N/A	N/A
Enhanced use			

XSL•FO
RenderX

Construct and items	Standardized factor loading (>0.7)	Composite reliability (>0.7)	Average variance extracted (>0.5)
ENH1[l]	0.83	0.914	0.727
ENH2	0.87	N/A	N/A
ENH3	0.85	N/A	N/A
ENH4	0.86	N/A	N/A
Continued intention to use			
CIU1[m]	0.82	0.953	0.801
CIU2	0.92	N/A	N/A
CIU3	0.92	N/A	N/A
CIU4	0.88	N/A	N/A
CIU5	0.93	N/A	N/A
Positive word-of-mouth intention			
PWOM1[n]	0.82	0.881	0.712
PWOM2	0.87	N/A	N/A
PWOM3	0.84	N/A	N/A

[a]IT: information technology.
[b]ITI–REL: IT identity–relatedness.
[c]N/A: not applicable.
[d]ITI–EMO: IT identity–emotional energy.
[e]ITI–DEP: IT identity–dependence.
[f]ITM–ALT: IT mindfulness–alertness to distinction.
[g]ITM–AW: IT mindfulness–awareness of multiple perspectives.
[h]ITM–OP: IT mindfulness–openness to novelty.
[i]ITM–OR: IT mindfulness–orientation in the present.
[j]PHS: perceived health status.
[k]FEAT: feature use behavior.
[l]ENH: enhanced use.
[m]CIU: continued intention to use.
[n]PWOM: positive word-of-mouth intention.

We also tested the discriminant validity of the constructs (Table 4). All the diagonal values (the square roots of the AVEs) were greater than 0.7 and exceeded the correlations between any pair of constructs [75]. Therefore, the result indicates that the model fulfills the requirements of discriminant validity, and it is assumed that the model also has adequate discriminant validity.

XSL•FO
RenderX

Table 4. Results of discriminant validity.

Construct	Mean	SD	ITI-REL[a]	ITI-EMO[b]	ITI-DEP[c]	ITM-ALT[d]	ITM-AW[e]	ITM-OP[f]	ITM-OR[g]	PHS[h]	FEAT[i]	ENH[j]	CIU[k]	PWOM[l]
ITI-REL	3.37	1.00	*0.840* [m]	N/A[n]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ITI-EMO	3.34	1.03	0.590	*0.830* [m]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ITI-DEP	3.43	1.02	0.615	0.639	*0.835* [m]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ITM-ALT	3.46	1.01	0.323	0.479	0.464	*0.830* [m]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ITM-AW	3.96	0.85	0.356	0.421	0.378	0.663	*0.830* [m]	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ITM-OP	3.85	0.97	0.229	0.393	0.446	0.671	0.634	*0.887* [m]	N/A	N/A	N/A	N/A	N/A	N/A
ITM-OR	3.76	0.87	0.315	0.373	0.399	0.546	0.618	0.526	*0.793* [m]	N/A	N/A	N/A	N/A	N/A
PHS	3.86	1.03	0.143	0.123	0.122	0.168	0.284	0.198	0.184	*0.858* [m]	N/A	N/A	N/A	N/A
FEAT	3.89	0.83	0.451	0.464	0.515	0.478	0.596	0.522	0.553	0.272	*0.825* [m]	N/A	N/A	N/A
ENH	3.67	0.92	0.546	0.531	0.496	0.497	0.516	0.517	0.485	0.270	0.523	*0.852* [m]	N/A	N/A
CIU	4.25	0.84	0.244	0.257	0.322	0.271	0.523	0.453	0.454	0.401	0.545	0.511	*0.89* [m]	N/A
PWOM	4.00	0.84	0.451	0.449	0.457	0.417	0.544	0.570	0.490	0.336	0.561	0.547	0.559	*0.843* [m]

[a]ITI-REL: IT identity–relatedness.

[b]ITI-EMO: IT identity–emotional energy.

[c]ITI-DEP: IT identity–dependence.

[d]ITM-ALT: IT mindfulness–alertness to distinction.

[e]ITM-AW: IT mindfulness–awareness of multiple perspectives.

[f]ITM-OP: IT mindfulness–openness to novelty.

[g]ITM-OR: IT mindfulness–orientation in the present.

[h]PHS: perceived health status.

[i]FEAT: feature use behavior.

[j]ENH: enhanced use.

[k]CIU: continued intention to use.

[l]PWOM: positive word-of-mouth intention.

[m]Greater than 0.7 and higher than the correlations between any pair of constructs.

[n]N/A: not applicable.

Control Variables

Factors that do not represent the core variables (ie, those included in the causal model) of this study, which may affect the interrelationships between the core variables, have been controlled for. As mentioned previously, we controlled for age, gender, race, income, employment, education, the purpose of use, and experience with a PHD. Although the causal model seems to represent individuals' active, automatic, and commitment use behaviors, we found that the effects of control variables were not negligible. On the basis of the findings, age and education influence feature use (β=−.20; *P*=.008; and β=.12, *P*=.02), which implies that younger users with higher education levels may exhibit a greater extent and breadth of use. Among the control variables, only education level influenced enhanced use (β=.19; *P*=.006). This result indicates that users with higher education backgrounds are more likely to use a formerly unused set of features for additional tasks. Age was the only control variable affecting continued intention to use (β=−.13; *P*=.03), indicating that older users are more likely to continue to use their PHDs. Finally, gender positively influences positive WOM intention (β=.18; *P*=.004). However, no effects of race, income, and purpose of use were found on any of the 4 use behaviors.

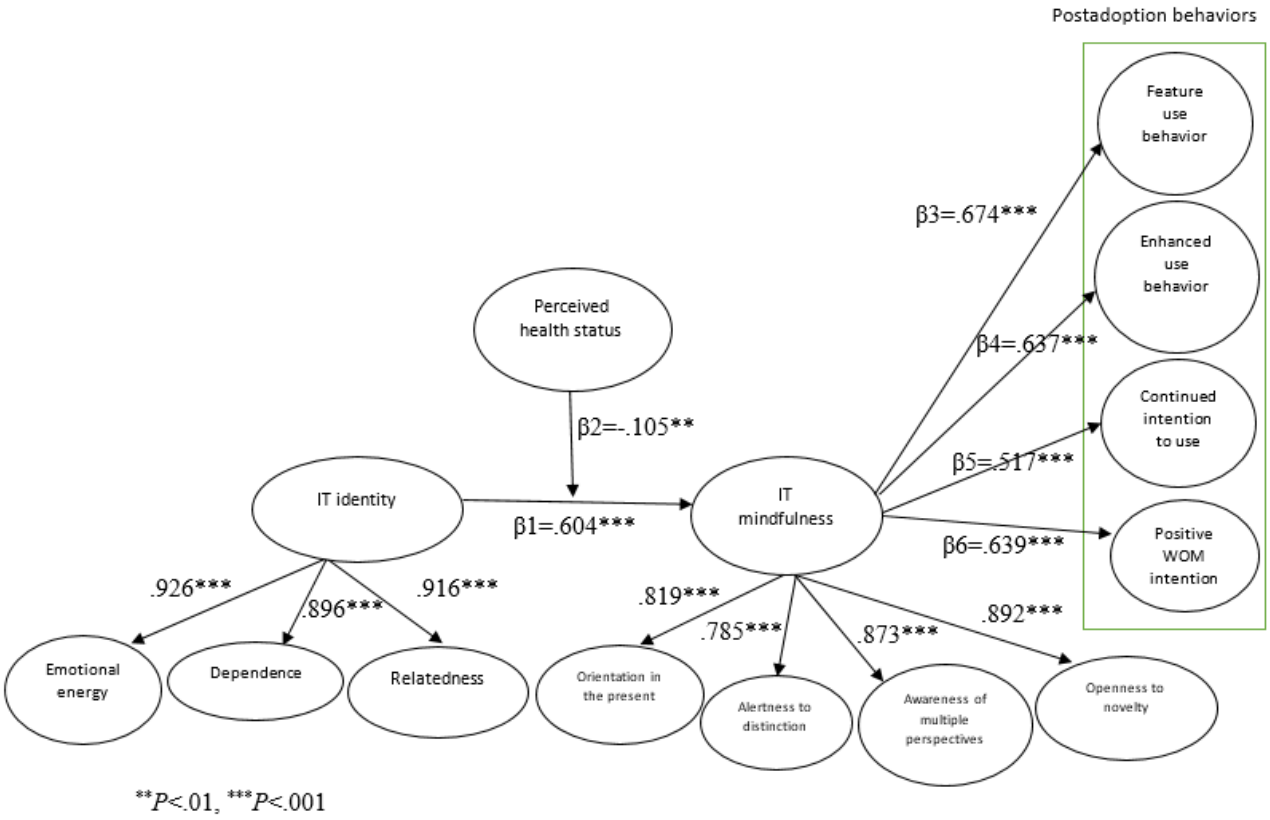
Structural Model

SmartPLS (version 3.0) was used to test the hypotheses within a structural equation modeling framework. According to Ho [76], the goodness-of-fit statistics can evaluate the entire structural model and assess the overall fit. The findings indicated the normed chi-square value of 2.5, which was between the recommended values of 1 and 3 [77]. The values for indices, that is, comparative fit index of 0.92, normed fit index of 0.91,

XSL•FO
RenderX

relative fit index of 0.93, and Tucker-Lewis index of 0.90, were above 0.9, and the index values for standardized root mean residual of 0.05 and root mean square error of approximation of 0.06 were below 0.08 [78]. The value of adjusted goodness-of-fit index (GFI) was 0.91, which exceeded 0.90. All these measures of fit were within the acceptable range, and only the GFI of 0.82 was marginal and could not meet the expected threshold value (which is >0.90). However, based on a study by Kline [79], at least four of the statistical values met the minimum recommended values, which supported a good fit between the hypothesized model and the observed data. Figure 2 displays the standardized path coefficients of the structural model under investigation.

Figure 2. Model paths. IT: information technology; WOM: word-of-mouth.



In this study, IT identity comprises 3 interrelated dimensions (relatedness, emotional energy, and dependence), and IT mindfulness comprises 4 interconnected dimensions (orientation in the present, alertness to distinction, awareness of multiple perspectives, and openness to novelty). The main reason for modeling the first-order constructs as reflective constructs is the expectation of interaction among the dimensions of the second-order construct [80]. This expectation was later confirmed by the presence of significant positive correlations between the 3 dimensions (Table 4). In a reflective construct, the dimensions have positive and significant intercorrelations as they share the same pattern [81]. Moreover, the findings show that all 3 dimensions of IT identity as first-order factors load significantly on the second-order construct, as the loadings were 0.92 for emotional energy, 0.89 for dependence, and 0.91 for relatedness. Similarly, the 4 dimensions of IT mindfulness as first-order factors load significantly on the second-order construct, as the loadings were 0.81 for orientation in the present, 0.78 for alertness to distinction, 0.87 for awareness of multiple perspectives, and 0.89 for openness to novelty. Thus, the combination of 4 dimensions reflects IT mindfulness in relation to PHDs. These characteristics are more indicative of a reflective construct.

To perform the partial least squares (PLS) structural equation modeling analysis, we determined one particular indicator per construct as a dominant indicator that correlates positively with the construct [82]. This approach avoids the issue of sign indeterminacy in PLS path modeling. The structural model was assessed by examining the path coefficients. We used bootstrapping to determine the significance of each path through *t* tests. The results of the hypotheses testing are summarized in Table 5. The findings support H1 by showing a significant positive relationship between IT identity and IT mindfulness (β=.604; *P*<.001). We analyzed the interaction terms to examine whether perceived health status moderates the impact of IT identity on IT mindfulness. One path goes from the interaction term to IT mindfulness; this path tests whether perceived health status moderates the relationship between IT identity and IT mindfulness (H2). To examine the moderating effect, we used the product indicator approach. As recommended by Henseler and Chin [83], when the exogenous or the moderator variable or both are formative constructs, the two-stage PLS approach for estimating moderating effects is a better approach than the product indicator approach. In this study, as none of the exogenous or moderator variables is formative, the product indicator method is preferred. Moreover, the product indicator

XSL•FO
RenderX

approach is suggested to be easily implementable in PLS path modeling [84]. The product term serves as an indicator of the interaction term in the structural model. This analysis indicates that the moderation hypothesis is supported (H2: β=−.105; *P*=.003). Therefore, the path from the interaction term has significant negative beta coefficients, indicating that the relationship between IT identity and IT mindfulness is negatively moderated by perceived health status. The *t* value for perceived health status's moderating effect was 2.623.

Table 5. Results of hypotheses testing.

Hypothesis	Path	Standardized coefficient	*P* value	Standard error	*t* statistics (*df*=25)	Results
H1	ITI[a]→ITM[b]	.604	<.001	0.040	15.086	Supported
H2	ITI→ITM (moderating effect of perceived health status)	−.105	.003	0.040	2.623	Supported
H3	ITM→feature use behavior	.674	<.001	0.042	16.141	Supported
H4	ITM→enhanced use	.637	<.001	0.038	16.719	Supported
H5	ITM→continued intention to use	.517	<.001	0.055	9.404	Supported
H6	ITM→positive WOM[c] intention	.639	.007	0.041	15.442	Supported

[a]ITI: information technology identity.

[b]ITM: information technology mindfulness.

[c]WOM: word-of-mouth.

H3 is also supported where higher IT mindfulness in relation to PHDs leads to feature use behaviors (β=.674; *P*<.001). The findings provide enough evidence to support H4, which indicates that IT mindfulness significantly reinforces enhanced use behaviors (β=.637; *P*<.001). The analysis also demonstrates that individuals' mindfulness with a PHD positively influences continued intention to use behaviors (β=.517; *P*<.001), and this positive linkage supports H5. The path coefficient of the relationship between IT mindfulness associated with PHDs and positive WOM intention is significant, supporting H6 (β=.639; *P*=.007).

Finally, the variables explained 48% of the variance in IT mindfulness, 30% of the variance in continued intention to use, 51% of the variance in enhanced use behaviors, 56% of the variance in feature use behaviors, and 45% of the variance in positive WOM intention. The R^2 scores reflect that the model provides relatively strong explanatory power to predict the variance in postadoption behaviors in the context of PHDs.

Discussion

Principal Findings

In this study, we adopt IT mindfulness as a theoretical lens to articulate factors affecting postadoption behaviors of PHD. We develop a research model (including determinants and outcomes) to gain a comprehensive view of the role of IT mindfulness during the postadoption usage of PHDs. Consistent with previous studies [22], the results confirm that IT mindfulness, as a domain-specific concept, can be used by IS studies to predict context-specific behaviors. By describing IT meaningfulness and explaining its relationships with active, automated, and commitment use behaviors, IS researchers can provide practitioners and developers with practical recommendations about how to advance users' value derived from PHD and how to retain and increase potential users. The results of this study contribute to the IS research on the area of mindfulness by examining the implications of IT mindfulness for PHD user performance.

We validated the second-order conceptualization of the IT mindfulness construct and demonstrated its utility in the context of smart health devices. As shown in the *Results* section, all 4 dimensions of IT mindfulness (ie, awareness, alertness, openness to novelty, and orientation in the present conditions) strongly contribute to the operationalization of this concept. With a better understanding of IT mindfulness, PHD system designers may be in a better position to design systems that support mindful use. Moreover, from a managerial perspective, characterizing PHD features in terms of mindfulness raises questions about how this aspect of IS operations should be managed. Practitioners and PHD developers can consider the malleability of these factors to cultivate IT mindfulness and improve consequent use behaviors. For instance, developers can add features to the PHD software to make users ready to become more mindful of PHD functionality. Previous research reports that critical thinking about how things can be done is likely to predict mindfulness, and some attributes of IT, such as the use of highly specific instructions, can hinder mindfulness [28]. One suggestion to raise IT mindfulness could be paying attention to flexible software structure so that instructions do not seem coercive to users and technical issues can be detected quickly.

Mindfulness theories indicate that excessive automation and routines are not desirable [28]. Offering customizable features and using gamification techniques coupled with defining an acceptable level of challenge for performing health-related tasks are likely to increase IT mindfulness. For example, PHD vendors can design health-related games that include multiple simple and complex steps and encourage users to participate in these challenges. The use of promotional efforts, such as providing participants with opportunities to gain points and redeem rewards in exchange for active participation, can enhance users' experience with smart devices. These features can elevate the state of users' awareness of PHD capabilities, alertness to the

XSL•FO
RenderX

device's distinction, engagement in the immediate health-related task context, and flexibility in system use. Another suggestion is to design factors that are integrated into social media platforms [85] that may improve IT mindfulness and enhance our understanding of how individuals explore and use PHD features. This suggestion is consistent with Junglas et al [86], indicating that adding social components and socially enabled features to digital devices can enhance the use behaviors of those technologies.

Theoretical and Practical Implications

One of the main theoretical contributions of this research is the identification of a cause of IT mindfulness. We show that IT identity is a strong antecedent of IT mindfulness. The results of this study contend that higher IT identity related to PHD will lead to stronger IT mindfulness associated with PHD use. Therefore, one practical way to enhance IT mindfulness in the context of smart devices could be by elevating users' IT identity. For instance, PHD vendors can stimulate users' sense of connection, levels of enthusiasm, and reliance on their PHD to increase their IT mindfulness, such as by continually introducing new features of PHDs [49]. It is valuable for practitioners to consider the dimensions of IT identity to establish specific guidelines and mechanisms to foster the IT mindfulness of PHD users.

Another theoretical implication of this research is to study the consequences of IT mindfulness and examine its effects on users' beliefs and behaviors. The findings shed more light on the explanatory power of IT mindfulness in predicting postadoption behaviors. Our study provides empirical evidence that IT mindfulness can be a significant factor affecting postadoption PHD use. On the basis of the results, IT mindfulness can explain additional variance in active system use (ie, enhanced use and feature use behaviors) than commitment behaviors and continued system use. Continued use defines the automatic extension of current PHD use, but active system use is finding new opportunities for changing existing use behaviors. Our results show that the relationship between IT mindfulness and active use behavior is stronger than its linkage with automated system use. In line with previous studies, mindfulness can nurture active rather than passive as well as a mechanical thinking process and motivate individuals to use an IT device to its fullest potential [50]. People with a higher level of IT mindfulness may pay more attention to their current context than obligations that restrict freedom of use behaviors [28]. Therefore, they may be more likely to accommodate their PHD use based on the situations they are experiencing. PHD developers can improve users' IT mindfulness by providing updated and well-formatted information about the features of this technology, explaining why their devices differ from others and articulating how these systems can be used for performing different health-related tasks.

Previous studies report that IT mindfulness empowers users to apply their knowledge in a flexible manner in new and unfamiliar situations [25]. In line with the literature, the results indicate that IT mindfulness enables individuals to innovate with their PHD to enhance their feature use. Therefore, IT-mindful people are expected to explore new and untested features to perform additional health-related tasks. Moreover, we show that IT mindfulness allows individuals to find new opportunities for using the current PHD features. The dimensions of IT mindfulness will increase the possibility of reaching more in-depth usage of familiar PHD features. Thus, we provide evidence that IT-mindful people have a greater tendency to use the existing features of their PHD in various situations.

This research is the first in the stream of IS use that hypothesizes a distinct impact of IT mindfulness on commitment behaviors. The findings demonstrate that IT mindfulness in the context of PHD use is a significant predictor of positive WOM intention. A possible justification is that IT-mindful individuals tend to monitor the task environment and keep abreast of new features and novel ways of using a system to perform different tasks [87]. Thus, they may not constrain themselves to current ways of using technology and will exhibit a greater likelihood of suggesting a PHD and its unique features to others. Their elevated awareness of the system's functionalities and applications may encourage them to form a larger level of commitment to their PHDs. This finding is consistent with results from previous studies, suggesting a strong relationship between awareness and WOM [88]. These results suggest that PHD developers consider IT mindfulness notions in their marketing campaign to promote usability as well as the value of their smart devices and increase the use rate. Specific marketing strategies in PHD companies can be developed to enhance users' state of being alert and aware of improving their affective commitment and positive WOM intention.

To highlight the health-related context of PHD usage and explore the contingent nature of the relationship between IT identity and IT mindfulness, we use perceived health status as a health factor. These findings imply that IT identity may lead to greater IT mindfulness, particularly among users with chronic physical or mental diseases. The moderating role of perceived health status demonstrates that IT mindfulness is dynamic and amenable to change through manipulation of individuals' health status perceptions. This is in line with previous studies, suggesting that IT mindfulness is a dynamic trait; thus, it can help vendors learn how IT can facilitate agility and flexibility rather than merely assuming that IT must benefit agility [89]. Our results show that the IT identity–IT mindfulness relationship is less substantial for users who perceive themselves to be healthier. Consequently, the likelihood that they will engage in active, automated, and commitment behaviors is lower. Therefore, users of a PHD who perceive a poor health status will exhibit higher IT identity and develop further awareness and alertness about its applications and tend to engage in a more nuanced use. The possible rationale is that poor perceptions of health status may drive unhealthier individuals to attach themselves more to their PHD and become mindful users of it in hopes of receiving promising health consequences.

There is considerable interest in understanding the interplay between perceived health status and IT mindfulness. We believe that our findings can be a useful means for exploring this relationship in greater depth. Regarding the moderating role of perceived health status, the IT identity–IT mindfulness linkage,

XSL•FO
RenderX

and in turn, the use behaviors of individuals are likely to vary. This makes it difficult for a single vendor to be able to generate health-related content and features that are comprehensive enough to embrace a wide range of health issues and topics. Thus, vendors need to decide the optimal scope of health-related functions and features on which their PHDs desire to focus. They can choose to offer a broader set of functions to cover a variety of health conditions or to focus on specific health issues (chronic diseases or typical ailments). Therefore, PHDs need to be relevant to the users, and developers should consider the target audience when designing their features, functions, and applications. For instance, a PHD may only offer the features and applications required to quit smoking or alcohol use. Another example would be a device that is required to monitor more severe issues such as cancer or HIV. According to the findings, we can argue that more focused devices with functions and applications devoted to a particular health situation may increase the chances of encouraging target users to exhibit beneficial use behaviors.

Limitations and Future Studies

It should be mentioned that the study is based only on a sample of respondents drawn from the United States. Therefore, the results may not be generalizable to all users of PHDs. It is recommended that future studies consider drawing samples from wider geographical areas, including other countries. Our study used a web-based survey to recruit participants digitally. As a self-rated sample of participants on MTurk was used, there is a small chance that some individuals were not completely aware of mobile technology and formed their mental construal of the IT artifact. Therefore, we suggest that further studies use a different method to ensure that subjects are knowledgeable about PHDs. For instance, future research can recruit informed patients who are directly referred by providers using patient self-management tools. Moreover, our study used a web-based survey to recruit participants digitally, which might induce sample selection bias. Thus, we only considered individuals who could access a computer, mobile devices, and the internet to participate in the web-based survey. Future studies can use other data collection means and sampling strategies to reach out to a sample that is generalizable to a wide range of health care consumers.

This study could also serve as a starting point for design science studies in the context of individual adoption of smart devices. In addition, this study could be viewed as an opening gate for research in the design of technology and assessing how investigated factors could shape actual performance and use of technology. In this study, no specific PHD was examined, but the general concept of a PHD was studied. For instance, it would be interesting to investigate how alternative PHD brands influence IT mindfulness enactment and, in turn, affect user positive WOM intention. Moreover, as there are many forms of consumer technologies (such as smartphones, tablets, and computers) with different IT characteristics, one promising research avenue would be to explore the effects of IT mindfulness in other contexts rather than PHDs. In this study, we defined IT as a unit of technology (hardware and software), and IT identity as well as IT mindfulness were examined in the context of general PHDs. Another promising area of research is expanding this study by examining IT identity and IT mindfulness for users of mHealth apps (software app) and analyzing the plausible differences. Our results are interpreted as personal identities and attitudes, as demonstrated by the instrument used in this study. By using this instrument, we cannot explain the results as diagnostic of some neurobiological or invariant cognitive constructs to which the individuals are condemned. It would be interesting for other studies to consider this area of research.

It should be mentioned that 3 demographic factors (ie, age, gender, and education level) that directly influence outcome variables are considered as control variables in our conceptual model. These effects could be viewed as a limitation of this study, as they may have affected the results. Future research could include these factors in the model and test their direct relationships with outcome variables. In this study, we discussed, modeled, and examined a positive relationship between IT identity and IT mindfulness. However, as a prospect for future studies, we also suggest that further research can investigate the possible effects of IT mindfulness on IT identity. This study highlights the significant moderating role of perceived health status between IT identity and IT mindfulness. Future studies could expand this moderating effect. For instance, additional research with a new study design is required to address what dimensions of IT identity could play a more significant role in shaping IT mindfulness in light of perceived health status effects. Future research should also compare the effects of specific health status (eg, physical and mental health stability) to deeply articulate whether a lack of physical and mental health stability could exert different effects on IT identity and its relationship with IT mindfulness. Furthermore, the results indicate that together, the factors were able to explain 48% of the variance in IT mindfulness. Although we controlled for confounding variables through randomization, we need to acknowledge the possible confounding effects of age, gender, education level, technology experience, and employment status in the proposed model. Another research avenue to consider is examining additional factors that may enhance the amount of variance in IT mindfulness explained (eg, trust in smart devices, personal innovativeness in IT, and computer self-efficacy).

Conclusions

IT mindfulness is a relatively new concept in IS research. This study contributes to IS research by validating the concept of IT mindfulness as a second-order construct with 4 reflective dimensions. We also develop a research model to examine the antecedents and implications of IT mindfulness for user performance in the context of PHDs. Through an empirical study, we offer evidence to highlight the importance of the IT mindfulness construct for studying individuals' resultant adoption behaviors within the domain of wearable health devices. Our results suggest that IT mindfulness could be cultivated through IT identity and relate closely to postadoptive PHD use. Furthermore, we demonstrate that perceived health status negatively moderates the relationship between IT identity and IT mindfulness associated with PHDs. Thus, we suggest that the link between IT identity and IT mindfulness is stronger for individuals who perceive themselves as unhealthier. The findings of this study provide insights into the phenomenon of

XSL•FO
RenderX

IT mindfulness formation and add to the literature on IT mindfulness, eHealth, mHealth, self-management tools, and health informatics. Owing to the impact of IT mindfulness on postadoption behaviors, its 4 dimensions could be used for designing PHD technologies. Moreover, vendors may need to put their efforts into means of increasing IT mindfulness by reinforcing IT identity to serve and retain a wide range of target users. Theoretical and practical contributions of this study are noticeable because they could result in a deeper understanding of human beings in relation to IT systems in an evolving digital world.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Online survey.
[DOCX File , 17 KB - mhealth_v8i10e18122_app1.docx]

References

1. Silva BM, Rodrigues JJ, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
2. Ling R. From ubicomp to ubiex(pectations). Telematics Info 2014 May;31(2):173-183. [doi: 10.1016/j.tele.2013.09.001]
3. Sun N, Rau PP. The acceptance of personal health devices among patients with chronic conditions. Int J Med Inform 2015 Apr;84(4):288-297. [doi: 10.1016/j.ijmedinf.2015.01.002] [Medline: 25655783]
4. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
5. Reychav I, Beeri R, Balapour A, Raban DR, Sabherwal R, Azuri J. How reliable are self-assessments using mobile technology in healthcare? The effects of technology identity and self-efficacy. Comput Hum Behav 2019 Feb;91:52-61. [doi: 10.1016/j.chb.2018.09.024]
6. Turel O, Serenko A, Bontis N. Family and work-related consequences of addiction to organizational pervasive technologies. Info Manag 2011 Mar;48(2-3):88-95. [doi: 10.1016/j.im.2011.01.004]
7. Metcalf D, Milliard ST, Gomez M, Schwartz M. Wearables and the internet of things for health: wearable, interconnected devices promise more efficient and comprehensive health care. IEEE Pulse 2016;7(5):35-39. [doi: 10.1109/MPUL.2016.2592260] [Medline: 28113167]
8. Zhang X, Yu P, Yan J, Ton A M Spil IR. Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: a case study in a primary care clinic. BMC Health Serv Res 2015 Feb 21;15:71 [FREE Full text] [doi: 10.1186/s12913-015-0726-2] [Medline: 25885110]
9. Attig C, Franke T. Abandonment of personal quantification: a review and empirical study investigating reasons for wearable activity tracking attrition. Comput Hum Behav 2020 Jan;102:223-237. [doi: 10.1016/j.chb.2019.08.025]
10. Esmaeilzadeh P. How does IT identity affect individuals' use behaviors associated with personal health devices (PHDs)? An empirical study. Info Manag 2020 May:103313. [doi: 10.1016/j.im.2020.103313]
11. Langer EJ. Minding matters: the consequences of mindlessness. J Exp Soc Psychol 1989:137-173. [doi: 10.1016/s0065-2601(08)60307-x]
12. Weick KE, Sutcliffe KM. Managing the Unexpected: Resilient Performance in an Age of Uncertainty. San Francisco, USA: Jossey-Bass; 2001.
13. Ndubisi NO. Mindfulness, reliability, pre-emptive conflict handling, customer orientation and outcomes in Malaysia's healthcare sector. J Bus Res 2012 Apr;65(4):537-546. [doi: 10.1016/j.jbusres.2011.02.019]
14. Fiol C, O'Connor E. Waking Up! Mindfulness in the Face of Bandwagons. AMR 2003 Jan;28(1):54-70. [doi: 10.5465/amr.2003.8925227]
15. Sun H, Fang Y, Zou H. Choosing a fit technology: understanding mindfulness in technology adoption and continuance. J Assoc Inf Syst 2016 Jun;17(6):377-412. [doi: 10.17705/1jais.00431]
16. Poston RS, Kettinger WJ. Mindfully experimenting with IT: cases on corporate social media introduction, Journal of Database Management (JDM), 25 (2014) 2020:29. [doi: 10.4018/jdm.2014040102]
17. Langer EJ, Moldoveanu M. The construct of mindfulness. J Social Isssues 2000 Jan;56(1):1-9. [doi: 10.1111/0022-4537.00148]
18. Ioannou A. Using IT Mindfulness to Mitigate the Negative Consequences of Technostress. In: Twenty-third Americas Conference on Information Systems. 2017 Presented at: AMCIS'17; August 10-12, 2017; Boston, Massachusetts, USA.
19. Wittmann M, Peter J, Gutina O, Otten S, Kohls N, Meissner K. Individual differences in self-attributed mindfulness levels are related to the experience of time and cognitive self-control. Pers Individ Differ 2014 Jul;64:41-45. [doi: 10.1016/j.paid.2014.02.011]

XSL•FO
RenderX

20. Swanson EB, Ramiller NC. Innovating mindfully with information technology. MIS Q 2004;28(4):553. [doi: 10.2307/25148655]
21. McElroy EJ, Marien C, Meyers JJ, Irschick DJ. Do displays send information about ornament structure and male quality in the ornate tree lizard, urosaurus ornatus? Ethology 2007 Nov;113(11):1113-1122. [doi: 10.1111/j.1439-0310.2007.01420.x]
22. Maier C, Laumer S, Wirth J, Weitzel T. Technostress and the hierarchical levels of personality: a two-wave study with multiple data samples. Eur J Inform Syst 2019 May 28;28(5):496-522. [doi: 10.1080/0960085x.2019.1614739]
23. Salovaara A, Lyytinen K, Penttinen E. High reliability in digital organizing: mindlessness, the frame problem, and digital operations. MIS Q 2019 Jan 1;43(2):555-578. [doi: 10.25300/misq/2019/14577]
24. Dernbecher S, Beck R. The concept of mindfulness in information systems research: a multi-dimensional analysis. Eur J Inform Syst 2017 Dec 19;26(2):121-142. [doi: 10.1057/s41303-016-0032-z]
25. Thatcher JB, Wright RT, Sun H, Zagenczyk TJ, Klein R. Mindfulness in information technology use: definitions, distinctions, and a new measure. MIS Q 2018 Mar 3;42(3):831-847. [doi: 10.25300/misq/2018/11881]
26. Sun H, Fang Y. Toward a Model of Mindfulness in Technology Acceptance. ICIS 2010:121 [FREE Full text]
27. Sternberg RJ. Images of mindfulness. J Social Isssues 2000 Jan;56(1):11-26. [doi: 10.1111/0022-4537.00149]
28. Butler BS, Gray PH. Reliability, mindfulness, and information systems. MIS Q 2006;30(2):211. [doi: 10.2307/25148728]
29. Polites GL, Roberts N, Thatcher J. Conceptualizing models using multidimensional constructs: a review and guidelines for their use. Eur J Inform Syst 2017 Dec 19;21(1):22-48. [doi: 10.1057/ejis.2011.10]
30. Collier JE, Bienstock CC. Model misspecification: contrasting formative and reflective indicators for a model of e-service quality. J Mark Theory Pract 2014 Dec 8;17(3):283-293. [doi: 10.2753/mtp1069-6679170306]
31. Carter M, Grover V. Me, my self, and I(T): conceptualizing information technology identity and its implications. MIS Q 2015 Apr 4;39(4):931-957. [doi: 10.25300/misq/2015/39.4.9]
32. Carter M, Grover V, Thatcher JB. Mobile devices and the self: Developing the concept of mobile phone identity. In: Lee I, editor. Strategy, adoption, and competitive advantage of mobile services in the global economy. Hershey, Pennsylvania: IGI Global; 2013.
33. Utz S. The function of self-disclosure on social network sites: not only intimate, but also positive and entertaining self-disclosures increase the feeling of connection. Comput Hum Behav 2015 Apr;45:1-10. [doi: 10.1016/j.chb.2014.11.076]
34. Card SK. The Psychology of Human-Computer Interaction. Boca Raton, FL: CRC Press; 2018.
35. Mcmillan SJ, Morrison M. Coming of age with the internet. New Media Soc 2016 Jun 30;8(1):73-95. [doi: 10.1177/1461444806059871]
36. 36 H. J. Eysenck, The structure of human personality (Psychology Revivals), Routledge, New York 2013. [doi: 10.4324/9780203753439]
37. Carter M, Petter S, Compeau D. Identifying with IT in a Digital World. In: Proceedings of the Fortieth International Conference on Information Systems. 2019 Presented at: ICIS 2019; December 15–18, 2019; Munich, Germany.
38. Balapour A, Reychav I, Sabherwal R, Azuri J. Mobile technology identity and self-efficacy: implications for the adoption of clinically supported mobile health apps. Int J Info Manag 2019 Dec;49:58-68. [doi: 10.1016/j.ijinfomgt.2019.03.005]
39. Bansal G, Zahedi F, Gefen D. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. Decis Support Syst 2010 May;49(2):138-150. [doi: 10.1016/j.dss.2010.01.010]
40. Kim JA, Yang SJ, Chee YK, Kwon KJ, An J. Effects of health status and health behaviors on depression among married female immigrants in south Korea. Asian Nurs Res (Korean Soc Nurs Sci) 2015 Jun;9(2):125-131 [FREE Full text] [doi: 10.1016/j.anr.2015.01.003] [Medline: 26160241]
41. Esmaeilzadeh P. Healthcare consumers' opt-in intentions to health information exchanges (HIEs): an empirical study. Comput Hum Behav 2018 Jul;84:114-129. [doi: 10.1016/j.chb.2018.02.029]
42. Hollis-Walker L, Colosimo K. Mindfulness, self-compassion, and happiness in non-meditators: a theoretical and empirical examination. Pers Individ Differ 2011 Jan;50(2):222-227. [doi: 10.1016/j.paid.2010.09.033]
43. Ryan RM, Brown KW. Why we don't need self-esteem: on fundamental needs, contingent love, and mindfulness. Pyschol Inq 2003;14(1):71-76 [FREE Full text]
44. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003 Apr;84(4):822-848. [doi: 10.1037/0022-3514.84.4.822] [Medline: 12703651]
45. Cash M, Whittingham K. What facets of mindfulness contribute to psychological well-being and depressive, anxious, and stress-related symptomatology? Mindfulness 2010 Jul 21;1(3):177-182. [doi: 10.1007/s12671-010-0023-4]
46. Raes F, Dewulf D, Van Heeringen C, Williams JM. Mindfulness and reduced cognitive reactivity to sad mood: evidence from a correlational study and a non-randomized waiting list controlled study. Behav Res Ther 2009 Jul;47(7):623-627. [doi: 10.1016/j.brat.2009.03.007] [Medline: 19362705]
47. Keng S, Smoski MJ, Robins CJ. Effects of mindfulness on psychological health: a review of empirical studies. Clin Psychol Rev 2011 Aug;31(6):1041-1056 [FREE Full text] [doi: 10.1016/j.cpr.2011.04.006] [Medline: 21802619]
48. Esmaeilzadeh P. The effects of public concern for information privacy on the adoption of health information exchanges (HIES) by healthcare entities. Health Commun 2019 Sep;34(10):1202-1211. [doi: 10.1080/10410236.2018.1471336] [Medline: 29737872]

XSL•FO
RenderX

49. Sun H. Understanding user revisions when using information system features: adaptive system use and triggers. MIS Q 2012;36(2):453. [doi: 10.2307/41703463]
50. Limayem M, Hirt SG, Cheung CM. How habit limits the predictive power of intention: the case of information systems continuance. MIS Q 2007;31(4):705. [doi: 10.2307/25148817]
51. Burton-Jones A, Gallivan MJ. Toward a deeper understanding of system usage in organizations: a multilevel perspective. MIS Q 2007;31(4):657. [doi: 10.2307/25148815]
52. Bagayogo F, Lapointe L, Bassellier G. Enhanced use of IT: a new perspective on post-adoption. J Assoc Inf Syst 2014 Jul;15(7):361-387. [doi: 10.17705/1jais.00367]
53. Wong CW, Lai K, Teo TS. Institutional pressures and mindful IT management: the case of a container terminal in China. Info Manag 2009 Dec;46(8):434-441. [doi: 10.1016/j.im.2009.08.004]
54. Baer RA, Smith GT, Allen KB. Assessment of Mindfulness by Self-Report. Assessment 2016 Jul 26;11(3):191-206. [doi: 10.1177/1073191104268029]
55. Maxham JG. Service recovery's influence on consumer satisfaction, positive word-of-mouth, and purchase intentions. J Bus Res 2001 Oct;54(1):11-24. [doi: 10.1016/s0148-2963(00)00114-4]
56. Bhattacherjee A. Understanding information systems continuance: an expectation-confirmation model. MIS Q 2001 Sep;25(3):351. [doi: 10.2307/3250921]
57. Lucas HC, Spitler V. Technology use and performance: a field study of broker workstations. Decis Sci 1999 Mar;30(2):291-311. [doi: 10.1111/j.1540-5915.1999.tb01611.x]
58. Carter M. IT Identity: Developing Valid Measures Through CFA-Based MTMM Analysis. In: Proceedings of the Thirty Fourth International Conference on Information Systems. 2013 Presented at: ICIS'13; December 15-18, 2013; Milano, Italy URL: https://experts.nebraska.edu/en/publications/it-identity-developing-valid-measures-through-cfa-based-mtmm-anal
59. Kayhan VO. The nature, dimensionality, and effects of perceptions of community governance. Info Manag 2015 Jan;52(1):18-29. [doi: 10.1016/j.im.2014.10.004]
60. langer EJ, Ngnoumen CT. Mindfulness in Positive Psychology: The Science of Meditation and Wellbeing. New York, USA: Routledge; 2017.
61. Venkatesh V, Goyal S. Expectation disconfirmation and technology adoption: polynomial modeling and response surface analysis. MIS Q 2010;34(2):281. [doi: 10.2307/20721428]
62. Hoehle H, Venkatesh V. Mobile application usability: conceptualization and instrument development. MIS Q 2015 Feb 2;39(2):435-472. [doi: 10.25300/misq/2015/39.2.08]
63. Huang JL, Curran PG, Keeney J, Poposki EM, DeShon RP. Detecting and deterring insufficient effort responding to surveys. J Bus Psychol 2011 May 31;27(1):99-114. [doi: 10.1007/s10869-011-9231-8]
64. Mason SW, Suri W. Conducting behavioral research on Amazon's Mechanical Turk. Behav Res Methods 2012 Mar;44(1):1-23. [doi: 10.3758/s13428-011-0124-6] [Medline: 21717266]
65. Boyer O'Leary M, Wilson JM, Metiu A. Beyond being there: the symbolic role of communication and identification in perceptions of proximity to geographically dispersed colleagues. MIS Q 2014 Apr 4;38(4):1219-1243. [doi: 10.25300/misq/2014/38.4.13]
66. Muthén LK, Muthén BO. How to use a Monte Carlo study to decide on sample size and determine power. Struct Equ Modeling 2002 Oct;9(4):599-620. [doi: 10.1207/s15328007sem0904_8]
67. Westland JC. Erratum to 'Lower bounds on sample size in structural equation modeling' [electron commerce res appl 9 (6) (2010) 476–487]. Electron Commer Res Appl 2012 Jul;11(4):445. [doi: 10.1016/j.elerap.2012.06.001]
68. Andersen PH, Kumar R. Emotions, trust and relationship development in business relationships: a conceptual model for buyer–seller dyads. Ind Mark Manag 2006 May;35(4):522-535. [doi: 10.1016/j.indmarman.2004.10.010]
69. West SG, Finch JF, Curran PJ. Structural Equation Models With Nonnormal Variables: Problems and Remedies. Thousand Oaks, CA: Sage; 1995.
70. Hair Jr JF, Hult GTM, Ringle C, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Thousand Oaks, CA: Sage; 2016.
71. Gefen D, Straub D, Boudreau M. Structural equation modeling and regression: guidelines for research practice. Commun Assoc Info Syst 2000;4:-. [doi: 10.17705/1cais.00407]
72. Chin WW. The partial least squares approach to structural equation modeling. Modern Method Bus Res 2020:295 [FREE Full text]
73. Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate Data Analysis. New Jersey, USA: Pearson Prentice Hall; 2006.
74. Segars A. Assessing the unidimensionality of measurement: a paradigm and illustration within the context of information systems research. Omega 1997 Feb;25(1):107-121. [doi: 10.1016/s0305-0483(96)00051-5]
75. Fornell G. Validity Assessment: A Structural Equations Approach Using Partial Least Squares. In: Proceedings of the American Marketing Association Educators' Conference. 1982 Presented at: AMA'82; April 6-8, 1982; Chicago, IL.
76. Ho R. Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS. Boca Raton, FL, USA: CRC Press Taylor & Francis Group; 2006.

XSL•FO
RenderX

77. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
78. Byrne BM. Structural equation modeling: perspectives on the present and the future. Int J Test 2001 Sep;1(3-4):327-334. [doi: 10.1080/15305058.2001.9669479]
79. Kline R. Principles and Practice of Structural Equation Modeling. Fourth Edition. New York, USA: Guilford Publications; 2015.
80. Mpinganjira M. Precursors of trust in virtual health communities: a hierarchical investigation. Info Manag 2018 Sep;55(6):686-694. [doi: 10.1016/j.im.2018.02.001]
81. Coltman T, Devinney TM, Midgley DF, Venaik S. Formative versus reflective measurement models: two applications of formative measurement. J Bus Res 2008 Dec;61(12):1250-1262. [doi: 10.1016/j.jbusres.2008.01.013]
82. Henseler J, Hubona G, Ray PA. Using PLS path modeling in new technology research: updated guidelines. Indust Manag Data Syst 2016 Feb;116(1):2-20. [doi: 10.1108/imds-09-2015-0382]
83. Henseler J, Chin WW. A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. Struct Equ Modeling 2010 Jan 8;17(1):82-109. [doi: 10.1080/10705510903439003]
84. Chin WW, Marcolin BL, Newsted PR. A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. Inf Syst Res 2003 Jun;14(2):189-217. [doi: 10.1287/isre.14.2.189.16018]
85. Timmerman CE. The moderating effect of mindlessness/mindfulness upon media richness and social influence explanations of organizational media use. Commun Monogr 2002 Jun;69(2):111-131. [doi: 10.1080/714041708]
86. Junglas I, Goel L, Abraham C, Ives B. The social component of information systems—how sociability contributes to technology acceptance. J Assoc Inf Syst 2013 Oct;14(10):585-616. [doi: 10.17705/1jais.00344]
87. Valorinta M. Information technology and mindfulness in organizations. Indust Corporate Change 2009 May 28;18(5):963-997. [doi: 10.1093/icc/dtp027]
88. Makanyeza C. Consumer awareness, ethnocentrism and loyalty: an integrative model. J Int Consum Mark 2015 Feb 25;27(2):167-183. [doi: 10.1080/08961530.2014.980927]
89. Tallon PP, Queiroz M, Coltman T, Sharma R. Information technology and the search for organizational agility: a systematic review with future research possibilities. J Strategic Inf Syst 2019 Jun;28(2):218-237. [doi: 10.1016/j.jsis.2018.12.002]

Abbreviations

AVE: average variance extracted
GFI: goodness-of-fit index
IS: information system
IT: information technology
mHealth: mobile health
MTurk: Mechanical Turk
PHD: personal health device
PLS: partial least squares
WOM: word-of-mouth

Edited by G Eysenbach; submitted 04.02.20; peer-reviewed by M Prietula, S Bidmon; comments to author 12.06.20; revised version received 19.06.20; accepted 01.09.20; published 05.10.20.

Please cite as:
Esmaeilzadeh P
The Role of Information Technology Mindfulness in the Postadoption Stage of Using Personal Health Devices: Cross-Sectional Questionnaire Study in Mobile Health
JMIR Mhealth Uhealth 2020;8(10):e18122
URL: http://mhealth.jmir.org/2020/10/e18122/
doi:10.2196/18122
PMID:33016884

©Pouyan Esmaeilzadeh. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 05.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User Perspectives of Mood-Monitoring Apps Available to Young People: Qualitative Content Analysis

Emily Widnall[1], MSc; Claire Ellen Grant[1], MSc; Tao Wang[1], PhD; Lauren Cross[1], MPhil; Sumithra Velupillai[1], PhD; Angus Roberts[1], PhD; Robert Stewart[1], PhD; Emily Simonoff[1], PhD; Johnny Downs[1], PhD

Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, United Kingdom

Corresponding Author:
Emily Widnall, MSc
Institute of Psychiatry, Psychology and Neuroscience
Kings College London
16 De Crespigny Park
London
United Kingdom
Phone: 44 7730683269
Email: emily.widnall@kcl.ac.uk

Abstract

Background: Mobile health apps are increasingly available and used in a clinical context to monitor young people's mood and mental health. Despite the benefits of accessibility and cost-effectiveness, consumer engagement remains a hurdle for uptake and continued use. Hundreds of mood-monitoring apps are publicly available to young people on app stores; however, few studies have examined consumer perspectives. App store reviews held on Google and Apple platforms provide a large, rich source of naturally generated, publicly available user reviews. Although commercial developers use these data to modify and improve their apps, to date, there has been very little in-depth evaluation of app store user reviews within scientific research, and our current understanding of what makes apps engaging and valuable to young people is limited.

Objective: This study aims to gain a better understanding of what app users consider useful to encourage frequent and prolonged use of mood-monitoring apps appropriate for young people.

Methods: A systematic approach was applied to the selection of apps and reviews. We identified mood-monitoring apps (n=53) by a combination of automated application programming interface (API) methods. We only included apps appropriate for young people based on app store age categories (apps available to those younger than 18 years). We subsequently downloaded all available user reviews via API data scraping methods and selected a representative subsample of reviews (n=1803) for manual qualitative content analysis.

Results: The qualitative content analysis revealed 8 main themes: accessibility (34%), flexibility (21%), recording and representation of mood (18%), user requests (17%), reflecting on mood (16%), technical features (16%), design (13%), and health promotion (11%). A total of 6 minor themes were also identified: notification and reminders; recommendation; privacy, security, and transparency; developer; adverts; and social/community.

Conclusions: Users value mood-monitoring apps that can be personalized to their needs, have a simple and intuitive design, and allow accurate representation and review of complex and fluctuating moods. App store reviews are a valuable repository of user engagement feedback and provide a wealth of information about what users value in an app and what user needs are not being met. Users perceive mood-monitoring apps positively, but over 20% of reviews identified the need for improvement.

(JMIR Mhealth Uhealth 2020;8(10):e18140) doi:10.2196/18140

KEYWORDS

mood monitoring; engagement; mobile applications; mHealth; mental health; smartphone; qualitative research; mobile phone

Introduction

Young people are leaders in adopting new technology, with recent statistics highlighting that 96% of those aged 16-24 years own a smartphone [1], and mobile phone usage among teenagers is increasing more than any other age group [2]. The smartphone revolution has not only changed the way young people communicate and acquire new information [3] but also

XSL•FO
RenderX

encouraged a rapid increase of mobile apps with varying functions. This growth is notable in the field of mobile health (mHealth) apps that deliver health and wellness technologies; as of 2017, consumer app stores had >325,000 mHealth apps available for download [4].

There has been increasing interest in using digital technology to administer interventions and monitor mental health symptoms in young people [5]. A particular area of mHealth growth lies in mood-monitoring apps. Mood monitoring is a widely used technique within nonclinical populations, provides insight into the development and trajectory of common mental health difficulties [6-8], and is an embedded technique in existing self-management techniques and evidence-based mental health treatments [9]. Self-tracking mood encourages users to actively engage in their health care management, provides a sense of autonomy [10,11], and increases awareness and self-regulation of emotional well-being [12,13]. There are several reviews exploring mood monitoring in adult populations [14-17]; however, much less is known about their use in child and adolescent populations. Although mood-monitoring apps are potentially cost-effective, accessible, and convenient, there remains a lack of evidence on how acceptable existing mood-monitoring apps are and particularly what features and functions engage younger populations.

This lack of understanding is further compounded by a limited consensus on how to measure user engagement. It is widely acknowledged within the literature that app engagement metrics and reporting remain unstandardized and heterogeneous [18,19]. The term *acceptability* often ranges from proxy markers, that is, adherence rates and utilization data [20], to participants' experience of burden [16], rather than understanding the features and functions that motivate and satisfy users. A recent systematic review evaluating mobile mood-monitoring apps in young people further demonstrates these inconsistencies [21]. A total of 9 studies of the 25 reviewed considered participants' perception of the apps, with only 3 studies specifically referring to *acceptability*, which was not explicitly defined; these used utilization and completion data as a proxy, which were interpreted by the authors as demonstrating broad acceptability [16,20,22]. The review demonstrated that young people generally positively perceive mood-monitoring apps and view them as user friendly, convenient, noninvasive, and useful; however, technological difficulties were reported to negatively affect user experience [16,23-25]. The review concluded that very few high-quality studies were available for inclusion and there is a need for more qualitative research to broaden our understanding of factors pertinent to the uptake of mood-monitoring apps.

The adoption of digital tools can also be evaluated by theories of *technology acceptance*, which argue that a person's intent to use and actual use of a technology is predicated by the person's perceptions of the technology's usefulness and ease of use [26,27]. The technology acceptance model (TAM) [26] has been used more recently to explore mHealth whereby *perceived helpfulness*, *perceived ease of use*, *perceived trust*, and *perceived security* were all found to directly influence user intention to use mHealth services [28]. However, the use of the TAM has been criticized for its weaknesses in explaining users' behavior and oversimplification of user perceptions to *usefulness* and *ease of use* [29].

It is important to gain a deeper understanding of user engagement with mHealth apps, particularly if young people are going to be using publicly available products unaccompanied and on a large scale. It is also crucial that young people are not set up to fail through poorly designed health apps or engagement with well-designed but ineffective digital treatments. Although small-scale qualitative studies have explored young people's views of mental health apps [30], more extensive research is needed to understand the nuances of user engagement. Written user reviews on mobile app stores contain a wealth of information about user experience and expectations and are a potentially untapped source of information in research, despite being used by smartphone owners to consider whether to download and engage with a given app [31]. We can, therefore, explore rich user reviews to understand what makes a mood-monitoring app acceptable to end users and what features are most prominent in positively reviewed apps.

The analysis of publicly available app reviews has been successfully used in recent literature to investigate user attitudes toward existing apps and their feature requests [32-34]. However, only a handful of studies have analyzed mHealth app reviews. These investigations have tended to assess app content quality for specific disorders [35], app functionality and user experience of a specific intervention [36], or apps targeting medication adherence [37]. Consumer perspectives of apps for bipolar disorder [35] found mostly positive feedback but also a large number of requests for desired functions. Users valued apps that were helpful, supportive, and easy to use and often integrated them into their health management and clinical care. Interestingly, users often did not consider the evidence base or clinical effectiveness of the app. User experience of cognitive behavioral therapy apps for depression [36] found that users valued the app in supporting their mental well-being and used the app as an adjunct to treatment. Concerns were also highlighted, particularly surrounding the importance of privacy, security, and trust. User experience of medication adherence apps [37] again found that users valued customization, the ability to monitor health information, and the ability for apps to support health care visits. Negative user experiences included technical difficulties, confusing app navigation, and inflexibility in the reminder setup.

In this paper, we perform a qualitative content analysis of user reviews to explore what app users consider useful to encourage frequent and prolonged app usage of mood-monitoring apps appropriate for young people.

Methods

Data Collection

A systematic review framework was applied to the search, screening, and assessment of apps. We searched the 2 major commercially available app stores: Google Play and Apple iPhone Operating System (iOS) store, by using the application programming interfaces (APIs) on these platforms [38,39]. We first used manual keyword searches to create a set of seed apps,

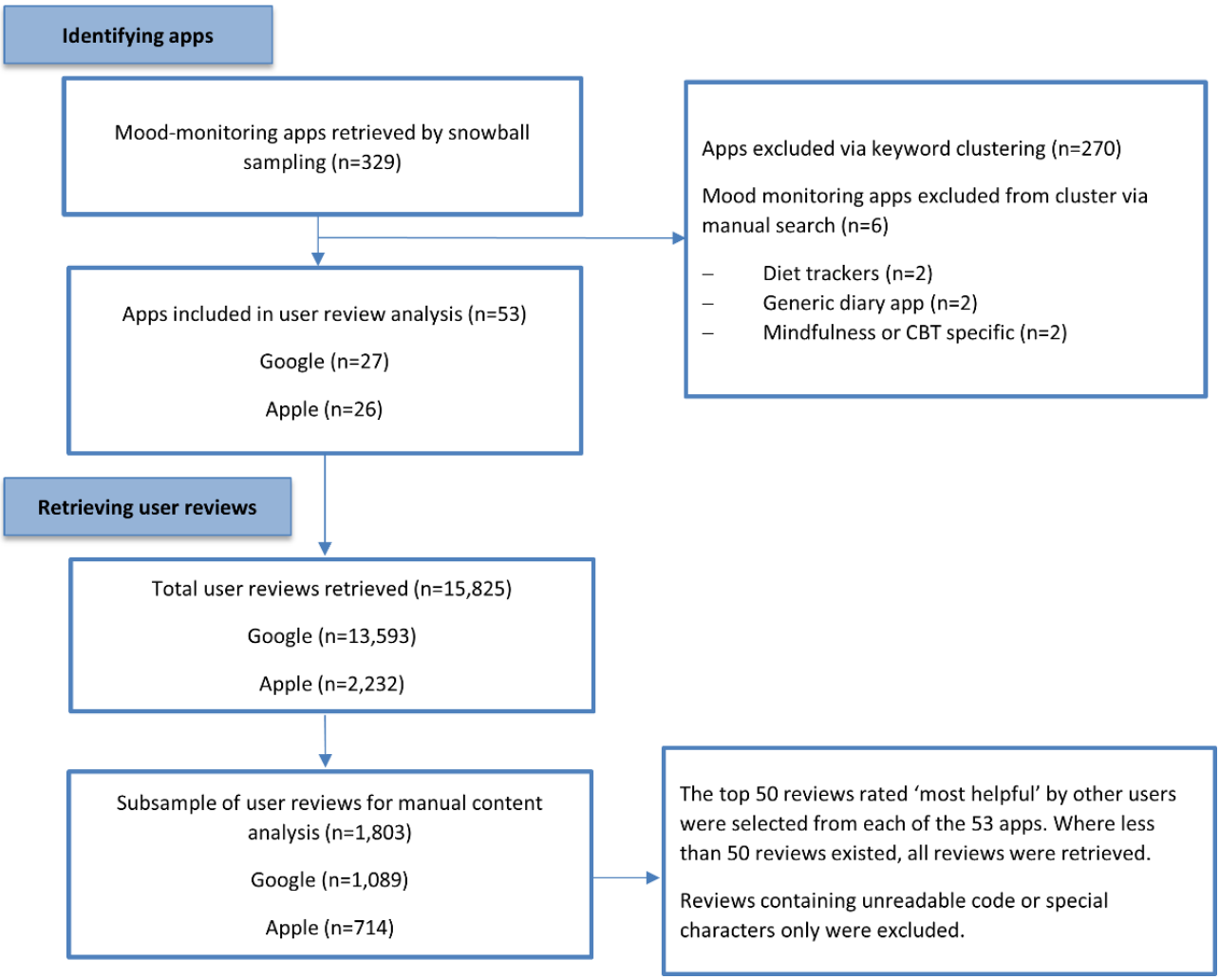
XSL•FO
RenderX

and then, we used a combination of API methodologies to identify and refine a set of relevant mood-monitoring apps on each platform. We then used API data scraping methods to collect all available user-generated reviews for the relevant apps identified. We used a combination of keyword searches to define a set of seed apps, followed by a snowball sampling technique to collect a series of similar apps. A full description of the API data collection methods can be found in Multimedia Appendix 1 [36,40-44].

Apps that met the following criteria were included: (1) self-reported mood-monitoring was the app's primary purpose, (2) the app was suitable for young people (aged less than 18 years) as described via the app store age rating, and (3) the app was available in English. Apps designed for health conditions other than mood disorders were excluded.

In total, 53 apps had 15,825 reviews. To gain a feasible number of reviews to manually appraise, we systematically selected a subsample of reviews. Both app stores provide different ways of sorting reviews, such as by date (most recent) or by the helpfulness of a review rated by other users (most helpful). After establishing that there was no significant difference between star rating distributions between the two ranking systems (see Multimedia Appendix 1 for full description), we ranked the reviews by *most helpful* and retrieved the top 50 reviews from each app. Where apps had less than 50 reviews, all reviews were collected. This resulted in a subset of 11.39% (1803/15,825) reviews: 1092 Google and 716 iOS. Figure 1 shows a flow diagram detailing the review process and results at each stage.

Figure 1. Flow diagram illustrating the selection of apps and user reviews. CBT: cognitive behavioral therapy.



Qualitative Content Analysis

Given the exploratory nature of this study, we conducted a qualitative content analysis to interpret themes in the app user reviews. Content analysis techniques have been widely used to understand user-generated review data, particularly in fields of research where existing theory is limited [45-47]. Given the rarity of content analyses of consumer perspectives on mobile apps within the literature, the majority of the analysis used an inductive approach to developing a coding framework [48]. Although the existing literature on analyzing mHealth app user reviews is limited, a smaller-scale deductive approach was carried out by using existing themes drawn from what published research was available to further inform our content analysis framework [11,21,35,49]. Following the guidelines on inductive analysis approaches in previous studies [21,35], we developed a database of coded user reviews. Our approach to the analysis

XSL•FO
RenderX

followed established coding techniques [50] across 3 phases: (1) data immersion, (2) data reduction, and (3) interpretation.

A preliminary sample of 500 reviews was randomly selected to ensure adequate coverage of a range of apps. Three coders (EW, CG, and LC) each coded the reviews for positive, negative, or neutral sentiment and noted whether the review highlighted a specific feature of engagement. This allowed the coders to become familiar with the informational content and to generate first-stage concepts.

After initial sentiment and feature identification, the coders developed a preliminary framework to organize codes (code name, description guidelines, and example quotes). Each researcher then used the codebook on an additional 100 randomly selected reviews. All researchers then met to revise, refine, and finalize the codes.

Interpretative notes were made and discussed, particularly around exploring word usage and the range of meanings. A set of themes and subthemes was subsequently revised and reordered during the interpretation phase. A coding framework was finalized, and 2 researchers (EW and CG) then independently coded the subset of user reviews (n=1803).

Results

Research Rigor

A substantial level of intercoder reliability across all codes was observed (κ=0.68), with high agreement for the themes adverts (κ=0.85), reminders (κ=0.86), and transparency (κ=0.89). Substantial intercoder agreement was observed for reflecting on mood (κ=0.72), technical (κ=0.72) accessibility (κ=0.71), recommendation (κ=0.79), and recording/representation of mood (κ=0.71). There were moderate levels of intercoder agreement for design (κ=0.61), developer (κ=0.65), health promotion (κ=0.61), flexibility (κ=0.64), social support (κ=0.51), and user requests (κ=0.59). A poor level of intercoder agreement was observed for games/gamification (κ=0.29). As this theme occurred in less than 1% of the reviews, it was subsequently removed from the coding framework.

Review Sentiment

Of the subset of 1803 user reviews, 1474 (81.7%) had a positive sentiment, that is, featured positive commentary on the app. However, positive reviews often included a contrasting statement, most commonly a user request for an additional feature. A total of 20% of positive reviews were general in nature and did not provide specific details on which features of the app were valued. A total of 8.9% (162/1803) of user reviews had a negative sentiment, and 9.2% (167/1803) of the reviews had a neutral sentiment. Over a third of the reviews with a negative sentiment included user feedback surrounding technical difficulties.

Description of Apps and Reviews

A total of 14 themes were identified in the data. Eight themes were prevalent in over 10% of the coded reviews (ranging from 34% prevalence to 11%), and 7 themes were present in less than 10% of the reviews (ranging from 7% to 1%). Codes with >10% prevalence were named major themes, whereas codes with <10% prevalence were named minor themes. Multimedia Appendix 2 shows the full coding framework describing all themes, subthemes, and illustrative quotes. Table 1 shows the frequency and percentage presence of all 14 themes within the user reviews. Multimedia Appendix 3 shows all 53 apps included in the analysis as well as app metadata. The included apps spanned 4 different categories: lifestyle (29/53, 55%), health and fitness (20/53, 37%), medical (2/53, 4%), and productivity (2/53, 4%). Across all 53 apps, there was an average star rating of 4.35, with 33 out of 53 apps rated 4.5 stars or above. Out of 53 apps, 47 were free of charge. The 6 apps that charged users ranged from $0.99 to $4.99.

XSL•FO
RenderX

Table 1. Prevalence of major and minor themes identified in user reviews (N=1803, categories not exclusive)

Themes[a]	Prevalence of theme in reviews, n (%)
Major themes	
1. Accessibility	614 (34.05)
2. Flexibility	370 (20.52)
3. Recording/representation of mood	322 (17.86)
4. User requests	302 (16.78)
5. Reflecting on mood	291 (16.14)
6. Technical feature	284 (15.75)
7. Design	225 (12.48)
8. Health promotion	186 (10.32)
Minor themes	
1. Notifications/reminders	130 (7.21)
2. Recommendation	115 (6.37)
3. Privacy, security, and transparency	102 (5.66)
4. Developer	69 (3.82)
5. Adverts	55 (3.05)
6. Social/community	38 (2)

[a]In 366 (20.30%) reviews, no engagement feature was coded.

Major Theme 1: Accessibility

Over a third of reviews centered around accessibility of the app. Users valued simplicity and frequently praised a simple and straightforward design that was easy to use. Users also frequently praised apps they perceived as fast and efficient; however, they expressed frustration with inefficient or slow apps.

Cost was also an important aspect; overall, users appreciated apps that were free of charge, but users often seemed happy to pay for premium versions if the app met their needs. However, users were frustrated when there were hidden subscription costs or when they had to pay for an app that did not meet their needs. Other users disagreed with developers charging at all for mental health–related apps.

Major Theme 2: Flexibility

The second most prevalent theme was flexibility. Users frequently referenced the need for the app to offer personalized and customizable features to suit individual user needs. This theme largely centered on 4 main features, the first of which was the ability for users to create their own personalized emotions or mood descriptions. Second was the ability for users to enter as many mood entries as they wished to in 1 day. Third was the ability for users to edit/modify/delete a previous entry, and the fourth one was that users preferred no restrictions/character limits being placed on free text entries.

Major Theme 3: Recording/Representation of Mood

How users record and represent their mood within the app was the third major theme. Of particular importance to the user within this theme was the variety of options available to represent mood, linking in with the theme of flexibility. Users often highlighted the complexity of moods and the need for multiple mood entries as well as custom scales. Similarly, users often described how a predetermined list of moods or emotions did not allow them to accurately represent their feelings and often requested the ability to elaborate on their mood using free text descriptions in their own words. Interestingly, some users also indicated the need for a balance between choice and specificity, for example, finding it helpful having a list to choose from when feeling confused over their own emotions but also the need to be able to name a mood of their choice.

Major Theme 4: User Requests

Approximately 1 in 5 reviews contained a request to the app developer. These requests were often a user wish list and requested features to improve their app experience. The most common requests centered on the theme of flexibility and personalization, such as customizable emotions, multiple entries per day, and editing entries.

Major Theme 5: Reflecting on Mood

The ability of users to reflect on their mood and mood entries over time was another main feature of engagement. Users particularly valued seeing their mood entries in the form of a graph or diagram. Users described how a visual display of mood over time allowed them to reflect on their good and bad days and value the ability to observe patterns and link moods to particular activities. Users frequently described the positive effect of reviewing their moods and experiences.

Major Theme 6: Technical Feature

Technical features largely referred to technical issues within the apps, such as data loss, inability to share mood entries across devices, or difficulties accessing or using the app. The technical

XSL•FO
RenderX

features theme, therefore, often referred to barriers to engagement with the app.

Due to the personal nature of data entered into mood-monitoring apps, users had frequent concerns surrounding loss of data and difficulties backing up or saving data. Many users reported experiencing significant amounts of data loss.

The ability to export or share data with different devices as well as with friends, family, and medical professionals was valued by users. When a mood-monitoring app did not include an export or share feature, this was frequently requested by users.

Major Theme 7: Design

The design of the app was important to users particularly in terms of the user interface being visually appealing, described by users with terms such as "beautiful," "pretty," and "sleek." Design preferences included a clean, simple design, which was intuitive and easy to navigate. Users valued simplicity and minimal designs over a cluttered screen.

Major Theme 8: Health Promotion

Users valued the ability of mood-monitoring apps to facilitate health promotion. Health promotion had 4 categories: (1) the mood-monitoring app itself being therapeutic for users and aiding self-awareness; (2) the ability to share mood entries with health care professionals to aid clinical appointments and facilitate discussions around their mood; (3) the ability for apps to provide psychoeducation, for example, understanding components of cognitive behavioral therapy; and (4) apps including signposting materials to available support services.

Minor Theme 1: Notifications/Reminders

Overall, users were positive around the use of notifications and reminders in apps and found this a helpful way of keeping on track with their mood monitoring. Some users even mentioned that a notification would promote a positive thought. It was also important that users were given the option to tailor their notifications/reminders to suit them.

Minor Theme 2: Recommendation

Written reviews also consisted of a number of recommendations to other users. These recommendations were indicative of their appreciation and positive experience with mood-monitoring apps. Users would often make recommendations to friends or family members as well as the wider app community. Users would also sometimes mention that their health care professional had recommended the app to them, which was typically followed by a positive review.

Minor Theme 3: Security, Privacy, and Transparency

Security and privacy mechanisms within mood-monitoring apps were important to users, and mistrust became an important issue, particularly surrounding the use of Facebook. The lack of openness regarding how and where data were stored was also a concern for some users. Transparency was rarely explicitly mentioned by users (<1% reviews), but it became a significant issue within individual apps, for example. Reviewers sometimes implicitly discussed themes of transparency, although this often conflated with trust, security, and privacy. Although transparency does not appear to be a crucial theme for engagement, knowledge of breeches, although rare, is key for rapid disengagement.

Minor Theme 4: Developer

App developers were an important factor for users, and comments to developers included praise and thanks, particularly commenting on timely responses from app developers. Users also demonstrated frustrations when app developers were not responsive to technical issues within the app, which led to users leaving a negative review surrounding developer communication itself rather than a specific feature of the app.

Minor Theme 5: Adverts

The use and frequency of adverts was important to many users who typically preferred apps with no adverts. Users generally disliked intrusive adverts, particularly those that interrupted the design or visual display of the app. Users were happy with being presented with optional adverts, particularly if wanting to support app developers.

Minor Theme 6: Social/Community

A number of reviews referred to the *community* aspect. Some mood-monitoring apps provided a peer support network feature, which was generally positively reviewed by users with a sense of listening to others as well as being listened to. Reviews also often included requests from users for app developers to include a support network to be built into apps that did not have one. Where there was a peer network available, some users described feeling limited in the way they were able to offer support. Users wanted to offer encouraging words but felt unable to, for example, some apps would limit user communication to emojis, which some users felt was not encouraging enough.

Discussion

Principal Findings

The aim of this study was to summarize and evaluate the main features of engagement within publicly available mood-monitoring apps appropriate for young people (aged less than 18 years) using app store user reviews. To our knowledge, this is the first exploration of consumer perspectives on mood-monitoring apps appropriate for young people using publicly available review data.

User feedback on mood-monitoring apps could generally be summarized by 8 main themes and 6 minor themes. Reviews varied in length, sentiment, and specificity, with many providing detailed and informative feedback about what engages and disengages users in mood-monitoring apps. Although 1 in 5 reviews did not contain a specific feature of engagement, the majority of reviews that did contain a feature of engagement contained multiple themes, demonstrating the complex and multifaceted nature of user needs.

The proportion of reviews containing positive and negative sentiments was similar to previous results in both general and mHealth apps whereby the majority of reviews contained positive sentiment [21,35]. The central positive features of engagement consisted of accessibility and personalization/customization of app content, which are in line

XSL•FO
RenderX

with previous findings of user reviews [30,35,36,49]. The main content of negative reviews in this study also supports previous findings, which cite functionality issues, lack of features, and crashing/data loss as the most common complaints [16,21,24,35,36]. These findings are broadly consistent with TAMs, in that users are more likely to adopt apps with high-quality design that is usable (easy to use, simple, and efficient) and useful (ability to reflect on mood and therapeutic features).

Although the majority of reviews had positive sentiments, the number of user requests (1 in 5 reviews) suggests room for improvement in currently available mood-monitoring apps to adequately address user needs. The frequency of user requests is in line with recent findings from user reviews of apps for bipolar disorder [35] and indicates that users have evolving needs and requirements when engaging with health apps. Users also hold an expectation that developers should address their needs and requests, which has also been found in previous literature [35,51,52]. This expectation also supports the findings that app stores serve as a communication channel between users and developers [52]. There appears to be a user-developer community, which highlights the potential for engaging end users throughout the app development process to ensure that the apps meet user needs before being made publicly accessible. User co-design poses obvious advantages for app function, uptake, and use by the target community [53-55].

Themes of accessibility (free and easy use), design (appearance and content), and social support (peers) show similarity with a recent study exploring adolescents' needs from mental health mobile apps as well as the importance of young people being in control, which is reflected in our theme of recording/representing mood and the significance of users having ownership over how they record their complex and changing moods [51]. Interestingly, social support, however, was much less prominent within this study compared with previous adult and adolescent studies [35,49,56]. This discrepancy could be due to our focus on generic mood-monitoring apps rather than clinical intervention apps or those designed for specific mental health conditions. This finding could also indicate the facilitation of self-management within mood-monitoring apps, which has been demonstrated in previous research, as well as creating a sense of greater control and autonomy around health management [11,49].

Although there may be a sense of self-management and autonomy within personal mood-monitoring apps, several reviews mentioned the benefit of being able to share their mood data with general practitioners, therapists, or counselors. This demonstrates joint partnerships and facilitation of communication between mHealth apps and health care providers. Users mentioned the ease of using app data within their clinical appointments to better communicate their mood over time as well as how different events had been affecting their mood. This two-way communication with health care providers perhaps demonstrates mHealth apps as a complementary tool to facilitate patient-provider relationships, which is in line with previous findings [37,49].

As found in user reviews of bipolar apps [35], there was a clear absence of discussion of scientific quality within user reviews. This also relates to previous qualitative research whereby users were motivated by information about whether or not an app would help them, but this information was not necessarily evidence based [30]. Again, this may be due to the study focusing on generic mood-monitoring apps, but it could also represent a disinterest in the scientific basis of mental health apps or a lack of health app regulation knowledge among users. It could also suggest that users implicitly trust apps that are publicly available on the app store, which highlights the importance of mHealth app literacy among users regarding evidence and data privacy.

The results of this study demonstrate a range of features that engage users in mood-monitoring apps but also highlight existing barriers that may prevent successful engagement. The positive features of engagement found in this paper include personalization and customization, a simple and intuitive design, features allowing users to reflect on their mood, and the facilitation of both self-management and communication with health care providers. The main barriers to engagement include concerns around privacy and security and technical difficulties surrounding data loss and app bugs/errors.

Limitations

Our results should be considered in light of the following limitations. First, the data used in this study were publicly available reviews. Our sampling frame for contributors and the representativeness of the views expressed are unknown. User reviews on app stores do not provide demographic data; therefore, we are not aware of the age of users submitting reviews. Although we based this study on apps available to those aged less than 18 years, the user reviews analyzed may be from a wide variety of age ranges, including adults. This is a limitation of using publicly available app data; hence, the engagement features we reviewed cannot be generalized to youth populations specifically.

Another limitation of looking at written user reviews is that they lack data surrounding user retention rates and periods of user engagement. It is, therefore, not possible to determine if reviews have been written after limited or extensive use of an app. Further research is needed to explore additional variables such as level of usage, understanding intent to start using mood-monitoring apps, and social influences.

This was an exploratory study of a relatively new area; therefore, specific research questions or hypotheses were not defined before the study. It is possible that important user attitudes may have been omitted from the publicly available reviews. It is also likely that there may be an element of bias in publicly available user reviews, for example, iOS does not make all written reviews publicly available [57]. However, given the intricacy of reviews and the number of reviews analyzed, we are confident that our findings represent users' likes and dislikes of mood-monitoring apps.

As app stores are very dynamic and frequently changing, the apps available, their features, and user review feedback are subject to change. Therefore, it is important for future research

XSL•FO
RenderX

to develop effective methodologies that can rapidly evaluate user feedback within this field. We were able to automate the identification and extraction of apps and reviews; however, developing an automated analysis of user reviews would be a valuable advancement in future research.

Conclusions

In this study, a content analysis framework was applied to a subsample of 1803 publicly available user reviews from 53 mood-monitoring apps appropriate for children and young people (based on app store age ratings). App store user reviews provide a valuable repository of anonymous, self-driven, and unstructured feedback. This paper provides a unique perspective on user attitudes and expectations toward mood-monitoring apps and allows an in-depth evaluation of the main features of engagement and potential barriers to adoption. Users value apps that can be personalized to their needs, have a simple and intuitive design, and allow accurate representation and review of complex and fluctuating moods.

Future studies should explore qualitative feedback from specifically recruited samples of children and adolescents using publicly available apps to explore whether the main features of engagement discovered in this study generalize to a defined child- and adolescent-only group and whether further details might be obtained from more reflective content. We hope these findings can support future guidelines on how apps are developed for end users, and we highlight the importance of including young people within the app design process to address disparities between end user perspectives and actual provisions within mHealth apps.

Acknowledgments

EW, CG, and LC were supported by the Medical Research Council (MRC) Mental Health Data Pathfinder award (MC_PC_17214) at Kings College London. JD received support from a National Institute for Health Research (NIHR) Clinician Scientist Fellowship (CS-2018-18-ST2-014) and Psychiatry Research Trust Peggy Pollak Research Fellowship in Developmental Psychiatry. SV, AR, RS, ES, and JD are partially funded by the NIHR Biomedical Research Centre at the South London and Maudsley NHS Foundation Trust and King's College London. The MRC was not involved in the study design, data collection, data analysis, report writing, and decision to submit the manuscript for publication.

Authors' Contributions

EW, TW, LC, JD, SV, and CG conceived the study design. EW, CG, and LC developed the content analysis coding framework; TW developed the API methodology with guidance from SV and AR. EW wrote the first draft of this manuscript, and all authors critically revised, edited, and approved the final version.

Conflicts of Interest

RS has received research support in the last 3 years from Janssen, Roche, and Takeda.

Multimedia Appendix 1
Description of application programming interface data collection methodology.
[DOCX File , 45 KB - mhealth_v8i10e18140_app1.docx]

Multimedia Appendix 2
Full content analysis coding framework and illustrative quotes of all themes and subthemes identified.
[DOCX File , 26 KB - mhealth_v8i10e18140_app2.docx]

Multimedia Appendix 3
List and meta data of all apps included in the analysis (n=53).
[DOCX File , 24 KB - mhealth_v8i10e18140_app3.docx]

References

1. Ofcom. Smartphone Usage in the United Kingdom (UK) 2012-2019, by Age. Statista. 2018. URL: https://www.statista.com/statistics/300402/smartphone-usage-in-the-uk-by-age/ [accessed 2020-09-28]
2. Fowler J, Noyes J. Are There Health Risks for Teenagers Using Mobile Phones? A Study of Phone Use Amongst 14–18 Year Olds. In: Selected Contributions From the International Symposium Occupational Safety and Hygiene (SHO 2017). Guimarães, Portugal (2017): CRC Press; Apr 10, 2017:59.
3. Matthews M, Doherty G, Sharry J, Fitzpatrick C. Mobile phone mood charting for adolescents. Br J Guid Couns 2008 May;36(2):113-129. [doi: 10.1080/03069880801926400]
4. The Market for Mhealth App Services Will Reach $26 Billion by 2017. Research2Guidance. 2017. URL: http://www.research2guidance.com/the-market-for-mhealth-app-services-will-reach-26-billion-by-2017/ [accessed 2019-09-25]

XSL•FO
RenderX

5. Kenny R, Dooley B, Fitzgerald A. Ecological momentary assessment of adolescent problems, coping efficacy, and mood states using a mobile phone app: an exploratory study. JMIR Ment Health 2016 Nov 29;3(4):e51 [FREE Full text] [doi: 10.2196/mental.6361] [Medline: 27899340]
6. Paris J. Borderline or bipolar? Distinguishing borderline personality disorder from bipolar spectrum disorders. Harv Rev Psychiatry 2004;12(3):140-145. [doi: 10.1080/10673220490472373] [Medline: 15371068]
7. Crowell S, Beauchaine T, Linehan M. A biosocial developmental model of borderline personality: elaborating and extending Linehan's theory. Psychol Bull 2009 May;135(3):495-510 [FREE Full text] [doi: 10.1037/a0015616] [Medline: 19379027]
8. Marwaha S, Balbuena L, Winsper C, Bowen R. Mood instability as a precursor to depressive illness: a prospective and mediational analysis. Aust N Z J Psychiatry 2015 Jun;49(6):557-565. [doi: 10.1177/0004867415579920] [Medline: 25850428]
9. Bastiaansen J, Meurs M, Stelwagen R, Wunderink L, Schoevers R, Wichers M, et al. Self-monitoring and personalized feedback based on the experiencing sampling method as a tool to boost depression treatment: a protocol of a pragmatic randomized controlled trial (ZELF-i). BMC Psychiatry 2018 Sep 3;18(1):276 [FREE Full text] [doi: 10.1186/s12888-018-1847-z] [Medline: 30176845]
10. Boydell K, Volpe T, Pignatiello A. A qualitative study of young people's perspectives on receiving psychiatric services via televideo. J Can Acad Child Adolesc Psychiatry 2010 Feb;19(1):5-11 [FREE Full text] [Medline: 20119561]
11. Wisniewski H, Liu G, Henson P, Vaidyam A, Hajratalli N, Onnela J, et al. Understanding the quality, effectiveness and attributes of top-rated smartphone health apps. Evid Based Ment Health 2019 Feb;22(1):4-9 [FREE Full text] [doi: 10.1136/ebmental-2018-300069] [Medline: 30635262]
12. Church K, Hoggan E, Oliver N. A Study of Mobile Mood Awareness and Communication Through MobiMood. In: 6th Nordic Conference on Human-Computer Interaction. 2010 Presented at: HCI'10; October 16-20, 2010; Reykjavik, Iceland. [doi: 10.1145/1868914.1868933]
13. Gay G, Pollak J, Adams P, Leonard J. Pilot study of Aurora, a social, mobile-phone-based emotion sharing and recording system. J Diabetes Sci Technol 2011 Mar 01;5(2):325-332 [FREE Full text] [doi: 10.1177/193229681100500219] [Medline: 21527101]
14. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
15. Naslund J, Marsch L, McHugo G, Bartels S. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015 May;24(5):321-332 [FREE Full text] [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
16. Dennis ML, Scott CK, Funk RR, Nicholson L. A pilot study to examine the feasibility and potential effectiveness of using smartphones to provide recovery support for adolescents. Subst Abus 2015;36(4):486-492 [FREE Full text] [doi: 10.1080/08897077.2014.970323] [Medline: 25310057]
17. Torous J, Powell A. Current research and trends in the use of smartphone applications for mood disorders. Internet Interv 2015 May;2(2):169-173 [FREE Full text] [doi: 10.1016/j.invent.2015.03.002]
18. Torous J, Luo J, Chan S. Mental health apps: what to tell patients. Curr Psychiatr. 2018. URL: https://mdedge-files-live.s3.us-east-2.amazonaws.com/files/s3fs-public/Document/February-2018/CP01703021.PDF
19. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: a review of measurement, reporting, and validity. Psychiatr Serv 2019 Jul 1;70(7):538-544 [FREE Full text] [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
20. Tregarthen JP, Lock J, Darcy AM. Development of a smartphone application for eating disorder self-monitoring. Int J Eat Disord 2015 Nov;48(7):972-982. [doi: 10.1002/eat.22386] [Medline: 26213130]
21. Dubad M, Winsper C, Meyer C, Livanou M, Marwaha S. A systematic review of the psychometric properties, usability and clinical impacts of mobile mood-monitoring applications in young people. Psychol Med 2018 Jan;48(2):208-228. [doi: 10.1017/S0033291717001659] [Medline: 28641609]
22. Reid SC, Kauer SD, Dudgeon P, Sanci LA, Shrier LA, Patton GC. A mobile phone program to track young people's experiences of mood, stress and coping. Development and testing of the mobiletype program. Soc Psychiatry Psychiatr Epidemiol 2009 Jun;44(6):501-507. [doi: 10.1007/s00127-008-0455-5] [Medline: 19011723]
23. Huh J, Shin H, Leventhal A, Spruijt-Metz D, Abramova Z, Cerrada C, et al. Momentary negative moods and being with friends precede cigarette use among Korean American emerging adults. Nicotine Tob Res 2014 Sep;16(9):1248-1254. [doi: 10.1093/ntr/ntu063] [Medline: 24755398]
24. Sacco G. Developing a Smartphone Application for Depression: Tracking Risk and Wellness Factors. University of Sydney Library. URL: https://udspace.udel.edu/handle/19716/17364 [accessed 2016-06-01]
25. Løventoft P, Nørregaard L, Frøkjær E. Designing Daybuilder: an Experimental App to Support People With Depression. In: Proceedings of the 12th Participatory Design Conference: Exploratory Papers, Workshop Descriptions, Industry Cases - Volume 2. 2012 Presented at: PDC'12; August 12-16, 2012; Roskilde, Denmark URL: https://www.researchgate.net/publication/235687221 [doi: 10.1145/2348144.2348146]
26. Davis F. A Technology Acceptance Model for Empirically Testing New End-user Information Systems : Theory and Results. DSpace@MIT Home. 1985. URL: https://dspace.mit.edu/handle/1721.1/15192 [accessed 2020-09-28]

XSL•FO
RenderX

27. Venkatesh V, Morris M, Davis G, Davis F. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425 [FREE Full text] [doi: 10.2307/30036540]
28. Alloghani M, Hussain A, Al-Jumeily D, Abuelma'atti O. Technology Acceptance Model for the Use of M-Health Services Among Health Related Users in UAE. In: International Conference on Developments of E-Systems Engineering. 2015 Presented at: DeSE'15; December 13-14, 2015; Duai, United Arab Emirates. [doi: 10.1109/dese.2015.58]
29. Ajibade P. Technology acceptance model limitations and criticisms: Exploring the practical applications and use in technology-related studies, mixed-method, and qualitative researches. Libr Philos Pract. 2019. URL: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=5294&context=libphilprac
30. Schueller SM, Neary M, O'Loughlin K, Adkins EC. Discovery of and interest in health apps among those with mental health needs: survey and focus group study. J Med Internet Res 2018 Jun 11;20(6):e10141 [FREE Full text] [doi: 10.2196/10141] [Medline: 29891468]
31. Genc-Nayebi N, Abran A. A systematic literature review: Opinion mining studies from mobile app store user reviews. Journal of Systems and Software 2017 Mar;125:207-219. [doi: 10.1016/j.jss.2016.11.027]
32. Iacob C, Harrison R. Retrieving and Analyzing Mobile Apps Feature Requests From Online Reviews. In: 10th Working Conference on Mining Software Repositories. 2013 Presented at: MSR'13; May 18-19, 2013; San Francisco, CA, USA. [doi: 10.1109/msr.2013.6624001]
33. Iacob C, Veerappa V, Harrison R. What Are You Complaining About?: A Study of Online Reviews of Mobile Applications. The 27th International BCS Human Computer Interaction Conference 2013;27. [doi: 10.5555/2578048.2578086]
34. Caldeira C, Chen Y, Chan L, Pham V, Chen Y, Zheng K. Mobile apps for mood tracking: an analysis of features and user reviews. AMIA Annu Symp Proc 2017;2017:495-504 [FREE Full text] [Medline: 29854114]
35. Nicholas J, Fogarty A, Boydell K, Christensen H. The Reviews Are in: A Qualitative Content Analysis of Consumer Perspectives on Apps for Bipolar Disorder. J Med Internet Res 2017 Apr 07;19(4):e105 [FREE Full text] [doi: 10.2196/jmir.7273] [Medline: 28389420]
36. Stawarz K, Preist C, Tallon D, Wiles N, Coyle D. User Experience of Cognitive Behavioral Therapy Apps for Depression: An Analysis of App Functionality and User Reviews. J Med Internet Res 2018 Jun 06;20(6):e10120 [FREE Full text] [doi: 10.2196/10120] [Medline: 29875087]
37. Park JY, Li J, Howren A, Tsao NW, De Vera M. Mobile Phone Apps Targeting Medication Adherence: Quality Assessment and Content Analysis of User Reviews. JMIR Mhealth Uhealth 2019 Jan 31;7(1):e11919 [FREE Full text] [doi: 10.2196/11919] [Medline: 30702435]
38. Olano F. Scrape Data From the iTunes App Store. GitHub. URL: https://github.com/facundoolano/app-store-scraper [accessed 2019-08-13]
39. Olano F. The Search Return Only 100 Records. GitHub. URL: https://github.com/facundoolano/google-play-scraper/issues/190 [accessed 2019-08-13]
40. Linares-Vasquez M, Holtzhauer A, Poshyvanyk D. On automatically detecting similar Android apps. In: IEEE 24th International Conference on Program Comprehension (ICPC).: IEEE; 2016 Presented at: International Conference on Program Comprehension; May 16-17, 2016; Austin, TX, USA. [doi: 10.1109/ICPC.2016.7503721]
41. Shah FA, Sabanin Y, Pfahl D. Feature-Based evaluation of competing apps. In: Proceedings of the International Workshop on App Market Analytics. 2016 Presented at: International Workshop on App Market Analytics; November, 2016; Seattle, WA, USA. [doi: 10.1145/2993259.2993267]
42. Shen N, Levitan MJ, Johnson A, Bender JL, Hamilton-Page M, Jadad AA, et al. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR Mhealth Uhealth 2015 Mar 16;3(1):e16 [FREE Full text] [doi: 10.2196/mhealth.3713] [Medline: 25689790]
43. Spärck Jones K. IDF term weighting and IR research lessons. J Doc 2004 Oct;60(5):521-523. [doi: 10.1108/00220410410560591]
44. Ward J. Hierarchical grouping to optimize an objective function. J Am Stat Assoc 1963 Mar;58(301):236. [doi: 10.2307/2282967]
45. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
46. Neuendorf K. Defining Content Analysis. Content Anal Guideb Internet . [doi: 10.4135/9781071802878.n1]
47. Tesch R. Qualitative Research: Analysis Types and Software. New York: Routledge; 2013.
48. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
49. Vo V, Auroy L, Sarradon-Eck A. Patients' Perceptions of mHealth Apps: Meta-Ethnographic Review of Qualitative Studies. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: 10.2196/13817] [Medline: 31293246]
50. Forman J, Damschroder L. Qualitative content analysis. In: Jacoby L, Siminoff L, editors. Empirical Methods for Bioethics: A Primer (Advances in Bioethics, Vol. 11). Bingley: Emerald Group Publishing Limited; 2007:62.
51. Mcilroy S, Shang W, Ali N, Hassan A. User reviews of top mobile apps in Apple and Google app stores. Commun. ACM 2017 Oct 24;60(11):62-67 [FREE Full text] [doi: 10.1145/3141771]

XSL•FO
RenderX

52. Pagano D, Maalej W. User Feedback in the Appstore: an Empirical Study. In: 21st IEEE International Requirements Engineering Conference. 2013 Presented at: RE'13; July 15-19, 2013; Rio de Janeiro, Brazil URL: https://ieeexplore.ieee.org/abstract/document/6636712/ [doi: 10.1109/re.2013.6636712]
53. Ellis LA, McCabe K, Davenport T, Burns JM, Rahilly K, Nicholas M, et al. Development and evaluation of an Internet-based program to improve the mental health and wellbeing of young men. 2015 Apr 20;12(1):2-13. [doi: 10.1108/itse-05-2014-0009]
54. Ellis L, McCabe K, Rahilly K, Nicholas M, Davenport T, Burns J, et al. Encouraging young men's participation in mental health research and treatment: perspectives in our technological age. Clinical Investigation 2014 Oct;4(10):881-888 [FREE Full text] [doi: 10.4155/cli.14.61]
55. Grant C, Widnall E, Cross L, Simonoff E, Downs J. Informing the development of an E-platform for monitoring wellbeing in schools: involving young people in a co-design process. Res Involv Engagem 2020;6:51 [FREE Full text] [doi: 10.1186/s40900-020-00219-0] [Medline: 32908677]
56. Kenny R, Dooley B, Fitzgerald A. Developing mental health mobile apps: exploring adolescents' perspectives. Health Informatics J 2016 Jun;22(2):265-275. [doi: 10.1177/1460458214555041] [Medline: 25385165]
57. Martin W, Harman M, Jia Y, Sarro F, Zhang Y. The App Sampling Problem for App Store Mining. In: 12th Working Conference on Mining Software Repositories. 2015 Presented at: ACM'15; May 16-17, 2015; Florence, Italy. [doi: 10.1109/msr.2015.19]

Abbreviations

API: application programming interface
iOS: iPhone operating system
mHealth: mobile health
MRC: Medical Research Council
NIHR: National Institute for Health Research
TAM: technology acceptance model

Edited by G Eysenbach; submitted 23.03.20; peer-reviewed by V Stamenova, B Loo Gee, C Reis; comments to author 04.05.20; revised version received 04.06.20; accepted 11.08.20; published 10.10.20.

Please cite as:
Widnall E, Grant CE, Wang T, Cross L, Velupillai S, Roberts A, Stewart R, Simonoff E, Downs J
User Perspectives of Mood-Monitoring Apps Available to Young People: Qualitative Content Analysis
JMIR Mhealth Uhealth 2020;8(10):e18140
URL: http://mhealth.jmir.org/2020/10/e18140/
doi:10.2196/18140
PMID:33037875

©Emily Widnall, Claire Ellen Grant, Tao Wang, Lauren Cross, Sumithra Velupillai, Angus Roberts, Robert Stewart, Emily Simonoff, Johnny Downs. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 10.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Comparison of Older and Younger Adults' Attitudes Toward the Adoption and Use of Activity Trackers

Sunyoung Kim[1], PhD; Abhishek Choudhury[1], MI

School of Communication and Information, Rutgers University, New Brunswick, NJ, United States

Corresponding Author:
Sunyoung Kim, PhD
School of Communication and Information
Rutgers University
4 Huntington Street
New Brunswick, NJ, 08901
United States
Phone: 1 8489327585
Email: sunyoung.kim@rutgers.edu

Abstract

Background: Activity tracking devices have significant potential in assisting older adults' health care and quality of life, but this population lags behind in the adoption of these devices. While theoretical frameworks have been introduced to explain and increase the adoption of this technology by older adults, little effort has been made to validate the frameworks with people in other age groups.

Objective: The goal of this study was to validate the theoretical framework of technology acceptance by older adults that we previously proposed through a direct comparison of the attitudes to and experiences of activity trackers in older and younger users.

Methods: Semistructured interviews were conducted with 2 groups of 15 participants to investigate their experiences of using activity trackers. The recruitment criteria included age (between 18 years and 24 years for the younger participant group or 65 years and older for the older participant group) and prior experiences of using mobile devices or apps for activity tracking for 2 months and longer.

Results: Our findings showed that the phase of *perceived ease of learning* as a significant influencer of the acceptance of activity trackers existed only in the older participant group, but this phase never emerged in the younger participant group. In addition, this study confirmed that other phases exist in both age groups, but 2 distinct patterns emerged according to age groups: (1) the *social influence* construct influenced the older participants positively but the younger participants negatively and (2) older participants' exploration in the *system experiment* phase was purpose-driven by particular needs or benefits but for younger participants, it was a phase to explore a new technology.

Conclusions: This study confirms the validity of the proposed theoretical framework to account for the unique aspect of older adults' technology adoption. This framework can provide theoretical guidelines when designing technology for older adults as well as when generating new investigations and experiments for older adults and technology use.

(JMIR Mhealth Uhealth 2020;8(10):e18312) doi:10.2196/18312

KEYWORDS

older adults; technology acceptance; activity tracker; fitness tracker; mHealth; health care; quality of life

Introduction

Background

Activity tracking devices that enable continuous monitoring of physical activities and physiological parameters have become widely available, allowing people to monitor their daily activity and overall health. People are now able to track their steps, heart rates, sleep patterns, and even engage in social forms of health tracking by using activity trackers [1]. With the collected data and presented information on these devices, not only can people gain insight into their daily activities but also be empowered to proactively manage and monitor health concerns, as physical activity helps reduce the risk of chronic diseases such as cardiovascular diseases, obesity, and diabetes [2].

XSL•FO
RenderX

The use of activity trackers by older adults is an area of particular research interest, since monitoring physical activity is a valuable parameter to define if persons are performing enough physical activities to prevent age-related chronic diseases or if they are manifesting early symptoms of those diseases [3]. However, there remains a notable digital divide between young adults and older adults. While over half of the Americans reported using a wearable fitness tracker at least once a day, only a little over 20% of the older adults owned an activity tracker in the United States as of 2016 [4]. In fact, low adoption of technology by older adults is not specific to activity trackers but is common with regard to any personal computing devices. While the adoption rates of computers and the internet by older adults are steadily increasing (from 12% in 2000 to 67% in 2016), these rates are significantly lower when compared with 90% of the general adult population using web-based services regularly [5]. Therefore, it is important to understand how older adults perceive and use new technology to meet their needs and to increase the adoption of new technology among older adults. To achieve this goal, studies have sought to understand how and why older adults maintain the use of new technology such as activity trackers and why they choose not to use or stop using this technology [6,7]. However, little comparative evidence exists with regard to the usage patterns and perspectives of older adults on new technology in a direct comparison with those of persons of other age groups.

Over the decades, technology acceptance models have been developed and refined to theoretically conceptualize the factors that influence the decision of whether to adopt new technology [8-13]. Within the context of technology adoption and the aging population, researchers have attempted to conceptualize older adults' technology acceptance [14-16]. As part of this effort, we proposed a new framework to account for older adults' acceptance of mobile technology for health care in our previous work [7], wherein *perceived ease of learning* had a significant influence on older adults' technology acceptance behavior, which did not appear in the existing frameworks. This study aimed to validate this framework by directly comparing the attitudes to and experiences of activity trackers in older and younger users.

Literature Review

Mobile Technology and Older Adults

Mobile technology is increasingly focused on the development of apps and tools to support health care, healthy living, and quality of life [17]. Wearable devices and other mobile technology for health care allow users to continuously track and manage health data without having to see their health care provider, such as diabetes management [18] and weight loss [19]. There is also a plethora of mobile apps for health care; as of 2019, there were over 45,000 apps for health care available for download from Apple app stores [20].

With regard to older adults and mobile technology, older adults are increasingly becoming savvy consumers of smartphone-based health solutions and information. With the increased desirability for aging in place, numerous technologies have emerged with the aim of supporting aging-related health concerns, including Alzheimer and dementia care [21], palliative care [22], monitoring fall risks [23,24], and osteoarthritis [25]. Moreover, research has shown that older adults hold positive views toward technology and have taken the steps for technology adoption [26,27]. For instance, Puri et al [28] showed that older adults were mostly accepting wearable activity trackers once they had a clear understanding of its value for their lives, and Preusse [29] showed that the adoption of activity trackers can be increased by addressing the barriers to acceptance. Despite the potential benefits and the increasing interests in mobile technology for health care, their adoption rates among older adults are still low [30]. Thus, researchers have extensively investigated how and why older adults decide to adopt and use mobile technology and why they choose not to use or stop using it. For instance, Lee and Coughlin [31] reviewed studies of older adults' technology acceptance and identified factors that are critical for older adults' acceptance of technology, including value, usability, affordability, accessibility, technical support, social support, emotions, independence, experience, and confidence. However, relatively little effort has been put to directly compare older adults' adoption of a new technology with those of the younger populations, with few exceptions [32,33].

Technology Acceptance Models for Older Adults

Technology acceptance models have been developed and refined over last couple of decades to explain technology adoption practices of different user groups [34] in various contexts [35] since the advent of foundational models, that is, Technology Acceptance Model [8] and the Unified Theory of Acceptance and Use of Technology [13].

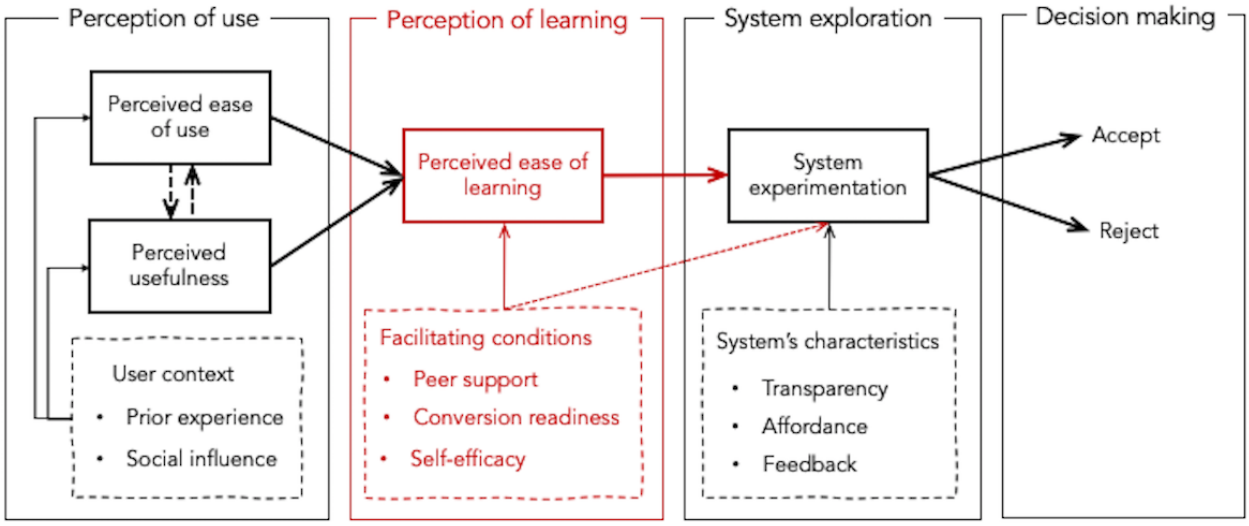
Extending these models, researchers have sought to conceptualize older adults' technology acceptance practice, though there are only few [14,36,37]. As part of this effort, we previously proposed a new theoretical framework to explain older adults' acceptance of mobile technology for health care as an extension of the predecessor theories by investigating the experiences and perspectives of 2 groups of older adults who were aged 60 years or older: technology adopters and nonadopters [7]. This framework introduced the perceived effort of *learning* as a significant obstacle for older adults' technology acceptance, which has been noted in prior research but has never been incorporated into any prior models of technology acceptance. For instance, Heart and Kalderon [38] suggest that special attention needs to be paid to teaching and training senior citizens to use new technology, and Yusif et al [39] pointed out lack of training as an area of concern in older adults' technology adoption. In Klimova and Poulova's literature review [15], they found that the existing technology acceptance models are suitable as the foundational theoretical basis for empirical studies, but more attention should be paid to forms of training for older adults. While these empirically grounded works are critical, there is no theoretical model that includes learning as an important phase of older adults' technology acceptance.

The theoretical framework we proposed comprises 4 phases, that is, (1) *perception of use*, the phase in which a user forms the intention to use a system; (2) *perception of learning*, the phase in which a user forms the intention to learn a system; (3) *system experimentation and exploration*, the phase in which a

XSL•FO
RenderX

user explores and experiments with a system, and (4) *decision making*, the phase in which a user decides whether to accept or reject a system (Figure 1). This framework suggests that availability of facilitating conditions, including *peer support, conversion readiness,* and *self-efficacy*, is critical for older adults to take the first step into the digital world of *learning a new technology*, thereby echoing prior work [36,40]. While Davis [9] previously highlighted learning as an important construct to account for technology acceptance, he regarded "ease of learning" as a substratum of the ease of use construct, while our framework proposed that perceptions about use and learning are not necessarily related. Our previous finding demonstrated a clear distinction between "perceived ease of learning" and "perceived ease of use" among older participants; older adults tend to think that the device might be easy to use for young people but not necessarily for them. Thus, they tend to give up learning new technology regardless of it being perceived as useful [7]. The limitation of our proposed framework is that it has not been validated with the young population to assure its unique application in older adults. Further, there has been no research, to the best of our knowledge, that has evaluated the validity of any existing models with people in other age groups. Therefore, this study aims to validate our framework for older adults' acceptance of mobile technology by conducting a comparison study with people in different age groups.

Figure 1. The proposed procedural model for older adults' acceptance of mobile technology for health care. The red-boxed section is a new phase with accompanying constructs that is proposed to be crucial for older adults.



Methods

Data Collection

This study employed semistructured interviews and short questionnaires with 2 groups of 15 participants. The recruitment criteria included age (between 18 years and 24 years for the younger participant group or 65 years and older for the older participant group) and prior experiences of using mobile devices or apps for activity tracking for 2 months and longer. Semistructured interviews explored topics related to their everyday experiences of using activity trackers, including how they acquired the device, experiences of learning and using, resources for support, and when applicable, reasons for attrition. A questionnaire was administered prior to the interviews to record demographic information and the type of device the participant was using or has used before. These questionnaires were used as prompts to supplement the interview questions.

The first group, that is, "younger participant group" consisted of 15 college students and the second group, that is, "older participant group" consisted of 15 older adults. The younger participant group consisted of 10 males and 5 females with a mean (SD) age of 20.1 (1.5) years (age range, 18-24 years). Their mean (SD) duration of use of the device was 19.3 (17.6) months (range, 2 months to 6 years). They were recruited through flyers posted at the university's student centers and the mailing lists. They were using a variety of activity tracking devices, including smartwatches, wristband-type activity trackers, and mobile apps for health care. The older participant group consisted of 4 males and 11 females aged 65 years or older with a mean (SD) age of 71.5 (6.3) years. Their mean (SD) duration of the use of the device was 38 (19.3) months (range, 2 months to 6 years). They were recruited through flyers posted at local libraries and senior centers. Device usage in this group, unlike that in the younger participant group, gravitated toward wristband-type activity trackers (eg, Fitbit) and mobile apps for health care (eg, the iPhone's Health app) and nobody used smartwatches (see Table 1 for demographic information on the participants). Each interview lasted approximately an hour and participants were compensated for their participation. All interactions were audio recorded and transcribed. The study was reviewed and approved by the institutional review board.

XSL•FO
RenderX

Table 1. Demographics of the participants and the devices they used.

Younger (Y) participant group					Older (O) participant group				
ID	Gender[a]	Age (years)	Type of device	Duration of use	ID	Gender	Age (years)	Type of device	Duration of use
Y1	M	20	Fossil smartwatch	1 year	O1	F	89	Fitbit	3 years
Y2	F	19	Fitbit	2 years	O2	F	69	Fitbit	4 years
Y3	M	21	Frontier smartwatch	2 years	O3	F	69	iPhone Health app	4 years
Y4	M	19	Fitbit	2 years	O4	F	72	Fitbit	5 years
Y5	F	21	Fitbit	7 months	O5	F	69	Fitbit	6 months
Y6	M	22	Fitbit	1 year	O6	M	66	Fitbit	2 years
Y7	M	19	Apple watch	3 years	O7	M	78	Fitbit, iPhone Health app	3 years
Y8	M	24	Garmin smartwatch	2 months	O8	F	80	Fitbit	1 year
Y9	F	20	Fitbit	18 months	O9	F	71	iPhone Health app	3 years
Y10	F	21	Fitbit	2 years	O10	F	70	Fitbit, H-Band	5 years
Y11	M	18	Fitbit	3 months	O11	F	72	Fitbit	6 years
Y12	M	20	Fitbit, MyKronoz	1 year	O12	M	67	iPhone Health app	5 years
Y13	M	20	Fitbit, iPhone Health app	2 years	O13	M	69	iPhone Health app	2 months
Y14	M	19	Apple watch	18 months	O14	F	65	Fitbit	1 year
Y15	F	19	iPhone Health app, Apple watch, Lose it app	6 years	O15	F	67	Fitbit, Weight Watchers app, My Fitness Pal app	3 years

[a]Gender: male (M), female (F).

Data Analysis

The interview data were analyzed using inductive and deductive approaches informed by grounded theory and other thematic analysis methods [41,42]. The themes and categories were identified deductively based on our proposed framework. Because the aspects and perspectives relating to *perceived learning* emerged as unique components that were critical to older adults' technology acceptance but are not presented in existing models, these were the focus of the validation. Then, the interview transcripts were open-coded and analyzed both inductively to identify new themes that emerged from the data and deductively to validate the themes related to *learning*. Both authors read and discussed the interview transcripts and developed codes to describe important concepts that emerged directly from the data. We coded independently with frequent discussions to reach consensus. We then analyzed the data to verify the themes and to ensure we had reached data saturation until no new themes or concepts emerged.

Results

Definitions

We described our findings of the attitudes to and experiences of activity trackers in different participant groups by the first 3 phases of our framework, that is, *perception of use*, *perception of learning*, and *system experimentation*. The term *learning* used in this section refers to the acquisition of knowledge or skills by being taught from external resources and *tinkering* refers to a self-guided, hands-on, trial-and-error–based process to acquire knowledge or skills.

Perception of Use

The first phase toward technology adoption in our framework is to formulate the perception about its use, which is influenced by its *perceived usefulness* and *perceived ease of use*. There were significant evidences to support the existence of these constructs from both participant groups.

Perceived Usefulness

Both participant groups acknowledged the perceived usefulness of activity trackers. All participants agreed on the potential utility of activity trackers to manage and improve health concerns (In the excerpts, "Participant O#" refers to the #th interviewee in the older participant group and "Participant Y#" refers to #th interviewee in the younger participant group).

> *…I think it's a good thing because it helps you to understand how healthy you are and what you're doing with yourself during the day to keep yourself healthy as you get older.* [Participant O12]

> *…In general, I think Fitbit is very useful...It really is a great device for tracking for people getting into shape and steps and anyone who is calorie counting.* [Participant Y2]

While the perceived "general" usefulness of activity trackers was unanimous across the groups, its perceived "personal" usefulness reflecting on one's own potential benefits was

XSL•FO
RenderX

divergent. Older participants described the perceived usefulness as a potential personal benefit to fulfill their own needs and deed, whereas younger participants perceived the devices to be useful for people other than themselves. This is not surprising since people generally become more vigilant about health concerns as they age, and young populations tend not to attend health care unless they have particular health problems. Prior work has shown that young adults have significantly lower rates of health care system utilization compared to older adults [43].

> *…I was excited because I could see that it was going to measure how many steps I was walking, how many times I was going up and down my steps, because I have about 14 steps in my house. So it was going to measure how many times I'm going up and down. I thought that was fascinating. You don't realize how many times you go up and down a step and how many steps you take every day when you walk… But this makes me conscious of all of that.* [Participant O8]

> *…I don't need all the fancy stuff about my health data, like how much I sleep, my nutrition, my laboratory results, and my reproductive health. Also, I definitely don't need to have my health records on my phone… That's not something I would want to do.* [Participant Y6]

While both *prior experience and social influence* emerged as significant constructs that influence the perceived usefulness in both groups, *social influence* was found to influence the perceived usefulness in different ways in different groups. For older participants, *social influence* positively impacted their perceived usefulness, as peers and other people in their close social network helped them discover and understand the potential utility of activity trackers. However, for some younger participants, *social influence* played a negative role, as an activity tracker was stereotyped as a tool for those with health concerns or weight management issues and thus using it was perceived to break a social norm of being healthy and active adolescents.

> *…My grandson gave it to me as a gift and I've had it about 2 or 3 years. He explained to me how it was working because he had one already and he thought that it would be a good idea for his old grandmother and his aunts to have one, so we all have one.* [Participant O1]

> *…You don't want to be seen with a Fitbit in high school. It would make you look 30. A suburban soccer mom trying to get into shape, I assume, would love the Fitbit. Kids don't want to do stuff like that. I didn't really see anyone else with a Fitbit because people are going to be like oh, why is he tracking his steps? Is he like a soccer mom who just got it to get active?* [Participant Y14]

Perceived Ease of Use

The theme of *perceived ease of use* emerged as a significant factor to distinguish the adoption of activity trackers, though directions and perspectives were different in each group. All younger participants said that they would never expect any difficulty in interacting with new devices, while many older participants expressed a general fear of interacting with new technology, which is not an exception for the case of activity trackers.

> *…I know with computers you can lose everything. I mean if I lose something, I have no idea how to find it. Or, if I change a setting and I can't find or go to where I want it to go anymore so that's why it is intimidating for me.* [Participant O9]

> *…The general idea of it (Fitbit) and the main features that I would be working with all of them are very easy to understand. It's very intuitive. And, it was pretty well organized. There wasn't much I had to work on to use it.* [Participant Y3]

Perception of Learning

Our findings confirmed that the *perceived ease of learning* phase exists only among older participants as a primary negative influencer in their adoption of activity trackers. It was evident that learning was perceived as a significant challenge for older participants. Older participants were hesitant to learn about using a new technology because they perceived that a new technology might be too difficult for them to learn, and some participants even thought that they are not capable of learning at all. Consequently, they refused to learn a new technology, regardless of its perceived usefulness. While a prior work by Renaud and Biljan [14] proposed *ease of learning* as an important construct in their model of seniors' technology acceptance, our finding is different in that their notion of ease of learning occurs as part of the actual system use phase, whereas ours is *"perceived"* ease of learning that is formulated prior to the actual system experimentation and exploration. Meanwhile, there was no comment related to learning throughout the entire transcripts of the younger participants.

> *…At this age, to learn everything is not possible. I do emailing and certain things by myself, but I don't want to learn everything because I may not be able to remember all that. But, certain things, if it is required for my Fitbit, I try to learn that. I will have to catch up with my grandson.* [Participant O8]

> *…I know technology is useful, but I don't make an effort to learn it. If you go to the phone company, they'll help you… I'm afraid to touch buttons because I might throw the whole thing out of whack. I just feel like I can't do it.* [Participant O11]

Our framework has 3 constructs that were proposed to influence older adults' *perception of learning* (and using) a new technology, that is, *peer support*, *conversion readiness*, and *self-efficacy*. The findings from this study confirmed that all 3 constructs have a significant influence on older participants' perception of learning how to use activity trackers but none of these emerged in the data of the younger participant group. The first construct, *peer support*, refers to support from people in a close social network. Our findings confirmed that older participants rely primarily on *peer support* when interacting with activity trackers for the first time.

XSL•FO
RenderX

> *…My husband shows some functions (of Fitbit) to me. Then, I understand, I operate, and I work with that. After a few hours or few days, however, I forget that, and I will again ask him if he shows me again. It was not that easy to be familiar with that.* [Participant O2]

> *…I had to take help of my grandson to figure out how to set this up. Because this Fitbit is installed by my grandson and he knows when he goes to fix. I don't know.* [Participant O6]

However, receiving support from other people was not something that older participants were always in favor of. In fact, they wanted to avoid seeking help from other people, if possible, which echoes the findings of prior work [7]. Researchers have found several reasons for older adults to be hesitant in receiving support from other people; older people are unwilling to reveal their lack of knowledge [44], a generational attitude of self-sufficiency exists, and the older adults prefer keeping their problems to themselves [45], or they do not want to bother people or interrupt people at what they consider to be crucial times [46]. While our findings did not demonstrate all these reasons, at least it was obvious that older participants tried to minimize seeking help from other people as much as possible, but most of the times, support from others was inevitable for them at least in the first few interactions with an activity tracker.

> *…Unfortunately, as much as I hate to admit, I will ask my friend for help. I really don't like asking him because I want to know how to do it on my own. But if I'm really stuck, then I'll ask him and then I can continue.* [Participant O5]

> *…I don't like to bother my son too often because he's very busy at work. So, when we see him maybe on Sunday morning at brunch where we get together for breakfast or something, I'll ask him.* [Participant O7]

The second construct, *conversion readiness*, refers to the degree to which a person is ready to accept a new thing. Prior work demonstrated that older adults are resistant to changing their current practices regardless of how useful a new technology is because they are set to their own ways of doing things without the use of technology [7]. Our findings confirmed that this construct exists among older participants, negatively influencing the intention to learn and adopt a new technology. Since they were satisfied with the current way of doing things, they did not even attempt to find out about the capabilities or benefits of new technologies, all of which did not appear among younger participants.

> *…I think the people and my friends at this age are more satisfied with what they have. I think, on average, the youngsters are enthusiastic with having more and more and more. That is the difference between those youngsters and we the people in the age ranges of 60 and 65 years. We are happy with what we have and what is needed, that's it.* [Participant O11]

The third construct, *self-efficacy*, refers to the degree to which a person believes to be capable of accomplishing a task. Our findings confirmed prior work that older participants lack *self-efficacy* in a new technology, which negatively influences their intention to learn how to use a new technology [29]. When a technology did not operate properly, older participants blamed themselves for the problem, which resulted in feeling "scared" or "afraid" of using a new technology. Again, these tendencies never emerged in the data of the younger participant group.

> *…I'm afraid to set it up myself because then I might mess up something else. I'm afraid if I enter something and everything gets messed up. So, I would not try it on my own. That's why it's always good to watch my son set some of the stuff up for me.* [Participant O15]

While *perceived ease of learning* was found to be a significant challenge for older participants' adoption of a new technology, they quickly became its active users once they successfully overcame this barrier. Several older participants who experienced difficulty in using activity trackers in their first acquisition reported that they now "feel comfortable" with using activity trackers because they "know what to do now."

> *…At first, it was difficult to navigate through it because there is like, you press this, you get this menu and then you get this menu and then you get if you wanted to enter information, then you have to do all these things. But now I learned all and feel comfortable with using it.* [Participant O4]

> *…At first, I was paranoid, scared, whatever but after doing it and asking questions a couple of times, maybe 3 or 4 times, and it's the same thing over and over. But now, I don't have to keep bothering anybody what to do anymore. I know what to do now.* [Participant O9]

System Experimentation

The *system experimentation* phase was confirmed to exist in both groups, though its pattern was different. Younger participants expressed a strong propensity to explore or tinker with a new technology rather than learning it when they first interacted with it as part of their efforts to figure out the features and functionalities of a device. A few younger participants searched information on the internet about how to use the device, but most of them jumped right into exploring and experimenting the features in their first interaction with it. Such explorations led to serendipitous discoveries of new functions and how to operate the features. Exploration and tinkering played as a key theme in which younger participants deepen their knowledge of the device. Numerous comments were received that demonstrated younger participants' practices of tinkering with or exploring a device when they first acquired it throughout the entire transcript.

> *….I didn't read the manual. I just synced it (Fitbit) up with my phone and started using it. When I was looking for something, I could figure out myself or looked it up online.* [Participant Y2]

> *…I just messed around with it. I started play around with the features and see what other stuff it did by just pressing the buttons on the app, like the different*

XSL•FO
RenderX

> *icons, to see what I can do. Like I found out about that alarm menu in there.* [Participant Y6]
>
> *…I was toying around with all the features and such. I just press around the different icons to see what I can do. Like I found out about that alarm menu and there were a bunch of other options...I fiddled around with it for about 15 minutes, but I wouldn't say there are any difficulties or complications with the device. So that's what I did the first couple times I used the Fitbit.* [Participant Y8]

Older participants also commented on their practices of exploring a new technology as part of an attempt to find new technologies. However, their exploration patterns were distinct from those of the younger participants in that the older participants' exploration was purpose-driven by particular needs or identified benefits, while younger participants' exploration was more of serendipitous and random experimentation.

> *…There's a lot of functions on it that I don't even understand. At my age, I don't explore that much but whatever I want to do, I try to investigate and find out like as I said the email, the—the phone, the texting, and the health part of it—that's mainly what I use all the time. And that's the benefit from having it, why I really wanted to have another one when the first one broke.* [Participant O12]

Discussion

Principal Findings

Our findings validated significant differences in the process through which people in different age groups accept or reject a new technology, using an activity tracking device as an exemplar. The phase of *perception of learning* existed among older participants as a significant influencer of their technology adoption but did not exist in younger participants. While older participants exhibited needs for some form of support for learning after acknowledging the *perception of use* but before *system exploration*, younger participants started to explore and tinker with activity trackers in their first acquisition of an activity tracker. This tendency can be explained by the fact that today's generations of older adults have not grown up using the contemporary personal technologies since their childhood; therefore, they are not familiar with the technologies [7]. Thus, there might be a natural confounding factor associated with age and experience, since "today's older adults are exposed to these technologies at a different point in their lives than today's young adults" [47]. Our findings suggest that this natural confounder results in the emergence of *perception of learning* a new technology as a unique phase to facilitate older adults' technology acceptance. The learning point of view is important because there will always be new technologies and new generations of older adults who have to learn how to use these. Therefore, in the development of new technologies, the learning perspective should be considered crucial to avoid exclusion of users of older groups.

Even though the phase of *perception of learning* was a significant challenge for older participants, it was not difficult to overcome. During the learning phase, older participants exhibited a strong reliance on *peer support* to either learn or get away from learning how to use an activity tracker, most of which successfully turned their final decision making into acceptance. This implies that it is crucial to provide older adults with easy access to facilitating conditions to lessen their tension and concern about learning a new technology [48]. While our data included only family members and friends as a resource of *peer support*, prior research has demonstrated a wide variety of resources within neighborhood groups and community groups that older adults can use to overcome learning-related difficulties, such as senior centers, local libraries, and local retail stores [49]. Offering classes lectured by older adult peers or peer-collaborative workshops through local community centers would be a way to lower older adults' perceived effort of learning as well as helping them reduce the burden of asking for help to other people. Fostering older adults' participation in such events will help technically isolated older adults find potential peer support for technology adoption.

Lastly, this study showed that all phases of technology adoption except *learning* exist in both age groups, but the patterns of how some phases and constructs influence technology adoption were different in different age groups. First, our findings demonstrated that the *social influence* construct had a significant influence on technology adoption but varied in different age groups; social influence positively influenced older users' technology, but it negatively influenced younger participants because an activity tracking device was negatively positioned for its use among some young generations. Second, the *system experiment* phase existed in both age groups, but the purpose was different; older user's exploration was driven by particular needs or benefits after learning it, whereas younger users explored a new technology to tour available features and functionalities and to figure out how to use it as a first step into its use. This illustrates that more in-depth investigation and discussion of how a theoretical framework of technology adoption applies to different age groups, since the same factors can have a different (or even opposite) influence on technology adoption in different age groups.

Limitations

The analyses presented in this paper are of a qualitative and explorative nature, providing in-depth insights into the issues older adults experience when using and learning to use activity trackers, in comparison to those experienced by college-age users. Small-scale qualitative studies have the advantage that they provide a rich picture of the ideas and experiences of the participants, but they are not able to provide a complete and representative picture of all the issues that are involved. Therefore, our results must be evaluated within the context of several study limitations. First, our sample size of patients was small (n=15 per group), and thus our participant pool may not be representative of a general population. In particular, all the younger participants were college students; therefore, they may not be representative of the entire young population. However, it is common in the sociology literature to regard college students as a representative of young adults when investigating age-related technology use practices since they are the major users of information and communication technologies [50,51]. In addition, other factors that might have influenced the results

XSL•FO
RenderX

were not investigated, such as gender difference [52], difference by levels of technology expertise [53], or difference by the duration of device use [54]. In particular, further research would be helpful to explore the perspectives by usage durations since our participants had varying durations of activity tracker use; our data relied on participants' memory, and memory could change over time. Lastly, all participants were recruited from an eastern metropolitan area of the United States. Therefore, our results may not generalize to the larger population of participants.

Conclusion

In an aging society, technological advances can have a positive impact on promoting the quality of later life. An activity tracking device is a type of electronic wearable device that holds significant potential in assisting older adults' health care by allowing to monitor and track health-related metrics. However, this population still shows slow rates of its adoption. While theoretical frameworks have been introduced to explain and promote the adoption of technology for older adults, little effort has been made to validate the frameworks with people in other age groups. Thus, we previously proposed a theoretical framework that sought to explicate technology acceptance for older adults [7], and this study aimed to validate this framework by directly comparing the attitudes to and experiences of activity trackers in older and younger users.

Our findings confirmed that the phase of *perceived ease of learning* as a significant influencer on the acceptance of activity trackers existed only among older users, but it never emerged among younger users. In addition, this study confirmed that other phases exist in both age groups, but 2 distinct patterns emerged by age groups: (1) the *social influence* construct influenced older participants positively but the younger participants negatively, and (2) older participants' exploration in the *system experiment* phase was purpose-driven by particular needs or benefits, but for younger participants, it was a phase to explore a new technology's features and functionalities. Based on these findings, we confirmed the validity of our proposed theoretical framework to account for the unique aspect of older adults' technology adoption. This framework can provide theoretical guidelines when designing a technology for older adults as well as when generating new ideas for investigations and experiments about older adults and technology use. We are hopeful that our findings will be useful toward expanding the knowledge and practices for leveraging emerging personal technologies to support the aging society.

Authors' Contributions

SK designed the study, performed data analysis and interpretation, and wrote the manuscript. AC undertook the data collection. All authors reviewed the final manuscript.

Conflicts of Interest

None declared.

References

1. Tedesco S, Barton J, O'Flynn B. A Review of Activity Trackers for Senior Citizens: Research Perspectives, Commercial Landscape and the Role of the Insurance Industry. Sensors (Basel) 2017 Jun 03;17(6). [doi: 10.3390/s17061277] [Medline: 28587188]
2. Steinert A, Haesner M, Steinhagen-Thiessen E. Activity-tracking devices for older adults: comparison and preferences. Univ Access Inf Soc 2017 Apr 8;17(2):411-419. [doi: 10.1007/s10209-017-0539-7]
3. Walsh M, Barton J, O'Flynn B, O'Mathuna C, Hickey A, Kellett J. On the relationship between cummulative movement, clinical scores and clinical outcomes. SENSORS, IEEE 2012:1-4. [doi: 10.1109/ICSENS.2012.6411587]
4. Elflein J. Percentage of the global population that used a mobile app or fitness tracking device to track their health as of 2016, by age. statista. 2019. URL: https://www.statista.com/statistics/742448/global-fitness-tracking-and-technology-by-age/ [accessed 2019-12-05]
5. Anderson M, Perrin A. Tech Adoption Climbs Among Older Adults. Pew research center. 2017 May 17. URL: https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2019-12-05]
6. Kononova A, Li L, Kamp K, Bowen M, Rikard R, Cotten S, et al. The Use of Wearable Activity Trackers Among Older Adults: Focus Group Study of Tracker Perceptions, Motivators, and Barriers in the Maintenance Stage of Behavior Change. JMIR Mhealth Uhealth 2019 Apr 05;7(4):e9832 [FREE Full text] [doi: 10.2196/mhealth.9832] [Medline: 30950807]
7. Kim S, Gajos KZ, Muller M, Grosz BJ. Acceptance of mobile technology by older adults: a preliminary study. 2016 Presented at: the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services; 2016; Florence, Italy p. 147-157. [doi: 10.1145/2935334.2935380]
8. Davis FD. A technology acceptance model for empirically testing new end-user information systems: Theory and results. PhD dissertation, Massachusetts Institute of Technology. 1985. URL: https://dspace.mit.edu/handle/1721.1/15192 [accessed 2019-12-05]
9. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989:319-340. [doi: 10.2307/249008]
10. Venkatesh V. Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. Information Systems Research 2000 Dec;11(4):342-365. [doi: 10.1287/isre.11.4.342.11872]

XSL•FO
RenderX

11. Venkatesh V. Technology acceptance model and the unified theory of acceptance and use of technology. Wiley Encyclopedia of Management 2015:1-9. [doi: 10.1002/9781118785317.weom070047]
12. Venkatesh V, Bala H. Technology Acceptance Model 3 and a Research Agenda on Interventions. Decision Sciences 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
13. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
14. Renaud K, Biljon JV. Predicting technology acceptance and adoption by the elderly: a qualitative study. 2008 Presented at: the annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology; 2008; Wilderness, South Africa p. 210-219. [doi: 10.1145/1456659.1456684]
15. Klimova B, Poulova P. Older people and technology acceptance. 2018 Presented at: International Conference on Human Aspects of IT for the Aged Population; 2018; Las Vegas, NV, USA p. 85-94. [doi: 10.1007/978-3-319-92034-4_7]
16. Macedo IM. Predicting the acceptance and use of information and communication technology by older adults: An empirical examination of the revised UTAUT2. Computers in Human Behavior 2017 Oct;75:935-948. [doi: 10.1016/j.chb.2017.06.013]
17. Wildenbos GA, Peute LWP, Jaspers MWM. A framework for evaluating mHealth tools for Older Patients on Usability. European Federation for Medical Informatics 2015:783-787. [doi: 10.3233/978-1-61499-512-8-783]
18. Hartz J, Yingling L, Powell-Wiley TM. Use of Mobile Health Technology in the Prevention and Management of Diabetes Mellitus. Curr Cardiol Rep 2016 Dec;18(12):130. [doi: 10.1007/s11886-016-0796-8] [Medline: 27826901]
19. Burke LE, Styn MA, Sereika SM, Conroy MB, Ye L, Glanz K, et al. Using mHealth technology to enhance self-monitoring for weight loss: a randomized trial. Am J Prev Med 2012 Jul;43(1):20-26 [FREE Full text] [doi: 10.1016/j.amepre.2012.03.016] [Medline: 22704741]
20. Matej M. Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 3rd quarter 2019. statista. 2020 May 13. URL: https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/ [accessed 2020-05-30]
21. Faucounau V, Riguet M, Orvoen G, Lacombe A, Rialle V, Extra J, et al. Electronic tracking system and wandering in Alzheimer's disease: a case study. Ann Phys Rehabil Med 2009;52(7-8):579-587 [FREE Full text] [doi: 10.1016/j.rehab.2009.07.034] [Medline: 19744906]
22. McCall K, Keen J, Farrer K, Maguire R, McCann L, Johnston B, et al. Perceptions of the use of a remote monitoring system in patients receiving palliative care at home. Int J Palliat Nurs 2008 Sep;14(9):426-431. [doi: 10.12968/ijpn.2008.14.9.31121] [Medline: 19060793]
23. Tamrat T, Griffin M, Rupcic S, Kachnowski S, Taylor T, Barfield J. Operationalizing a wireless wearable fall detection sensor for older adults. 2012 Presented at: International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops; 2012; San Diego, CA p. 297-302. [doi: 10.4108/icst.pervasivehealth.2012.248643]
24. Yamada M, Aoyama T, Okamoto K, Nagai K, Tanaka B, Takemura T. Using a Smartphone while walking: a measure of dual-tasking ability as a falls risk assessment tool. Age Ageing 2011 Jul;40(4):516-519. [doi: 10.1093/ageing/afr039] [Medline: 21593058]
25. Bellamy N, Wilson C, Hendrikz J, Whitehouse SL, Patel B, Dennison S, et al. Osteoarthritis Index delivered by mobile phone (m-WOMAC) is valid, reliable, and responsive. J Clin Epidemiol 2011 Feb;64(2):182-190. [doi: 10.1016/j.jclinepi.2010.03.013] [Medline: 20609562]
26. Carpenter BD, Buday S. Computer use among older adults in a naturally occurring retirement community. Computers in Human Behavior 2007 Nov;23(6):3012-3024. [doi: 10.1016/j.chb.2006.08.015]
27. Kuerbis A, Mulliken A, Muench F, Moore AA, Gardner D. Older adults and mobile technology: Factors that enhance and inhibit utilization in the context of behavioral health. Ment Health Addict Res 2017;2(2). [doi: 10.15761/mhar.1000136]
28. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User Acceptance of Wrist-Worn Activity Trackers Among Community-Dwelling Older Adults: Mixed Method Study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]
29. Preusse KC, Mitzner TL, Fausset CB, Rogers WA. Older Adults' Acceptance of Activity Trackers. J Appl Gerontol 2017 Feb;36(2):127-155 [FREE Full text] [doi: 10.1177/0733464815624151] [Medline: 26753803]
30. Friemel TN. The digital divide has grown old: Determinants of a digital divide among seniors. New Media and Society 2014 Jun 12;18(2):313-331. [doi: 10.1177/1461444814538648]
31. Lee C, Coughlin JF. PERSPECTIVE: Older Adults' Adoption of Technology: An Integrated Approach to Identifying Determinants and Barriers. J Prod Innov Manag 2014 Jun 03;32(5):747-759. [doi: 10.1111/jpim.12176]
32. Broady T, Chan A, Caputi P. Comparison of older and younger adults' attitudes towards and abilities with computers: Implications for training and learning. British Journal of Educational Technology 2010;41(3):473-485. [doi: 10.1111/j.1467-8535.2008.00914.x]
33. Deng Z, Mo X, Liu S. Comparison of the middle-aged and older users' adoption of mobile health services in China. Int J Med Inform 2014 Mar;83(3):210-224. [doi: 10.1016/j.ijmedinf.2013.12.002] [Medline: 24388129]
34. Scherer R, Siddiq F, Tondeur J. The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. Computers & Education 2019 Jan;128:13-35. [doi: 10.1016/j.compedu.2018.09.009]

XSL•FO
RenderX

35. Dutot V, Bhatiasevi V, Bellallahom N. Applying the technology acceptance model in a three-countries study of smartwatch adoption. The Journal of High Technology Management Research 2019 May;30(1):1-14. [doi: 10.1016/j.hitech.2019.02.001]
36. Pan S, Jordan-Marsh M. Internet use intention and adoption among Chinese older adults: From the expanded technology acceptance model perspective. Computers in Human Behavior 2010 Sep;26(5):1111-1119. [doi: 10.1016/j.chb.2010.03.015]
37. Chung JE, Park N, Wang H, Fulk J, McLaughlin M. Age differences in perceptions of online community participation among non-users: An extension of the Technology Acceptance Model. Computers in Human Behavior 2010 Nov;26(6):1674-1684. [doi: 10.1016/j.chb.2010.06.016]
38. Heart T, Kalderon E. Older adults: are they ready to adopt health-related ICT? Int J Med Inform 2013 Nov;82(11):e209-e231. [doi: 10.1016/j.ijmedinf.2011.03.002] [Medline: 21481631]
39. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: A systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
40. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. Computers in Human Behavior 2013 Jul;29(4):1715-1724. [doi: 10.1016/j.chb.2013.02.006]
41. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
42. Strauss A, Corbin J. Grounded theory methodology. In: Denzin NK, Lincoln YS, editors. Handbook of qualitative research. NY, USA: Sage Publications Inc; 1994:273-285.
43. Bonnie RJ. Investing in the health and well-being of young adults. In: Stroud C, Breiner H, editors. The National Academies of Science, Engineering, and Medicine. Washington DC: The National Academies Press; 2014.
44. Kang HG, Mahoney DF, Hoenig H, Hirth VA, Bonato P, Hajjar I, Center for Integration of Medicine and Innovative Technology Working Group on Advanced Approaches to Physiologic Monitoring for the Aged. In situ monitoring of health in older adults: technologies and issues. J Am Geriatr Soc 2010 Aug;58(8):1579-1586. [doi: 10.1111/j.1532-5415.2010.02959.x] [Medline: 20646105]
45. Knight BG, Winterbotham S. Rural and urban older adults' perceptions of mental health services accessibility. Aging Ment Health 2020 Jun;24(6):978-984. [doi: 10.1080/13607863.2019.1576159] [Medline: 30761911]
46. Coleman GW, Gibson L, Hanson VL, Bobrowicz A, McKay A. Engaging the disengaged: How do we design technology for digitally excluded older adults? 2010 Presented at: the 8th ACM Conference on Designing Interactive Systems; 2010; Aarhus, Denmark p. 175-178. [doi: 10.1145/1858171.1858202]
47. Hanson VL. Technology skill and age: what will be the same 20 years from now? Univ Access Inf Soc 2011 Apr 20;10(4):443-452. [doi: 10.1007/s10209-011-0224-1]
48. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
49. Delello JA, McWhorter RR. Reducing the Digital Divide: Connecting Older Adults to iPad Technology. J Appl Gerontol 2017 Jan;36(1):3-28. [doi: 10.1177/0733464815589985] [Medline: 26084479]
50. Pak R, Price MM, Thatcher J. Age-sensitive design of online health information: comparative usability study. J Med Internet Res 2009 Nov 16;11(4):e45 [FREE Full text] [doi: 10.2196/jmir.1220] [Medline: 19917549]
51. Lian J, Yen DC. Online shopping drivers and barriers for older adults: Age and gender differences. Computers in Human Behavior 2014 Aug;37:133-143. [doi: 10.1016/j.chb.2014.04.028]
52. Whitley BE. Gender differences in computer-related attitudes and behavior: A meta-analysis. Computers in Human Behavior 1997 Jan;13(1):1-22. [doi: 10.1016/s0747-5632(96)00026-x]
53. Ziefle M, Schaar AK. Technical expertise and its influence on the acceptance of future medical technologies: what is influencing what to which extent? : Springer; 2010 Presented at: Symposium of the Austrian HCI and Usability Engineering Group; 2010; Berlin, Heidelberg p. 513-529. [doi: 10.1007/978-3-642-16607-5_40]
54. Zhou J, Rau PP, Salvendy G. Older adults' use of smart phones: an investigation of the factors influencing the acceptance of new functions. Behavior and Information Technology 2013 Apr 09;33(6):552-560. [doi: 10.1080/0144929x.2013.780637]

Edited by G Eysenbach; submitted 18.02.20; peer-reviewed by S Jones, M Jordan-Marsh; comments to author 10.03.20; revised version received 27.03.20; accepted 14.05.20; published 22.10.20.

Please cite as:
Kim S, Choudhury A
Comparison of Older and Younger Adults' Attitudes Toward the Adoption and Use of Activity Trackers
JMIR Mhealth Uhealth 2020;8(10):e18312
URL: https://mhealth.jmir.org/2020/10/e18312
doi:10.2196/18312
PMID:33090116

XSL•FO
RenderX

©Sunyoung Kim, Abhishek Choudhury. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 22.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

New Checklist for the Heuristic Evaluation of mHealth Apps (HE4EH): Development and Usability Study

Kamran Khowaja[1], PhD; Dena Al-Thani[1], PhD

Information & Computing Technology, College of Science & Engineering, Hamad Bin Khalifa University, Doha, Qatar

Corresponding Author:
Kamran Khowaja, PhD
Information & Computing Technology
College of Science & Engineering
Hamad Bin Khalifa University
LAS Building
Education City
Doha,
Qatar
Phone: 974 66314481
Email: kamran.khowaja@gmail.com

Abstract

Background: Diabetes is one of the leading causes of death in developing countries. Existing mobile health (mHealth) app design guidelines lack a description of the support of continuous self-monitoring of health status, behavior change to improve and adopt a healthy lifestyle, and communication with health educators and health care professionals in case of any need.

Objective: This paper presents the development of a specialized set of heuristics called heuristic evaluation for mHealth apps (HE4EH) as an all-in-one tool and its applicability by performing a heuristic evaluation of an mHealth app.

Methods: An extensive review of heuristics and checklists was used to develop the HE4EH. The HE4EH was evaluated by domain experts for heuristics, checklist items, severity ratings, and overall satisfaction. The OneTouch app, which helps individuals with diabetes manage their blood glucose levels, was evaluated using HE4EH to identify usability problems that need to be fixed in the app.

Results: The expert evaluation of HE4EH revealed that the heuristics were important, relevant, and clear. The checklist items across the heuristics were clear, relevant, and acceptably grouped. In terms of evaluating the OneTouch app using the HE4EH, the most frequently violated heuristics included Content, Visibility, Match, and Self-monitoring. Most of the usability problems found were minor. The system usability scale score indicated that the OneTouch app is marginally acceptable.

Conclusions: This heuristic evaluation using the OneTouch app shows that the HE4EH can play a vital role for designers, researchers, and practitioners to use HE4EH heuristics and checklist items as a tool to design a new or evaluate and improve an existing mHealth app.

(JMIR Mhealth Uhealth 2020;8(10):e20353) doi:10.2196/20353

KEYWORDS

mHealth; eHealth; heuristic evaluation; expert evaluation; self-monitoring; behavior change; design guidelines; framework

Introduction

The leading reason for morbidity and mortality, especially in developing countries, is chronic disease [1]. About 80% of these deaths in developing countries are due to cardiovascular disease and diabetes mellitus, a proportion that is higher than in developed countries. It is expected that this number will increase further to 85% by 2030. The World Health Organization has reported that the number of people with these diseases increased from 108 million in 1980 to 422 million in 2014 [2]. At least 17% of the population in all the Gulf countries, including Qatar (location of the workplace of both authors of this manuscript), has diabetes. Researchers at Weill Cornell Medicine-Qatar have predicted that the prevalence rate of type II diabetes will increase from 12% in 2012 to at least 24% by 2050 [3]. One main reported finding is that most cases are due to obesity. This shows a need to support affected individuals with health care education and lifestyle changes such as a healthy diet and increasing physically active to improve their quality of life. Beratarrechea

XSL•FO
RenderX

and colleagues [1] conducted a systematic review on mobile health (mHealth) apps in developing countries and highlighted that mHealth apps address health care system constraints related to increased demand and limited resources. Developing countries have barriers, including limited health care human resources, insufficient finances, increasing populations, and an inability to reach the portion of the population living in remote areas, among others. Proper adherence to chronic disease management is essential to improve an individual's quality of life and health-related outcomes [4]. Hamine and colleagues [4] also conducted a systematic review of the impact of mHealth on interventions and found that mHealth apps are increasingly accessible and highly acceptable tools for patient communication, monitoring, and education. They are also effective means to facilitate self-management. Other barriers, including language, literacy, access to a smartphone, cost, access to the internet, or mobile network, are typically present everywhere. This shows that there is a need to design mHealth tools with these barriers in mind so that targeted patients can easily use the developed mHealth solution with minimal training or support. To keep patients motivated in using mHealth solutions on a regular basis throughout the treatment or intervention, gamification is increasingly adopted, as it facilitates self-management of chronic conditions [5]. Thus, an app needs to support an individual to continuously self-monitor his or her health status, change behavior to improve his or her lifestyle, and communicate with health educators and health care professionals in case of any need.

Designers typically use an existing set of guidelines or heuristics as a base to design a new app; they also evaluate and improve the usability aspects of a developed app in light of an existing set of guidelines or heuristics. A search on guidelines and heuristics did not reveal an all-in-one package that supports self-monitoring, behavior change, and communication with health care professionals, as already mentioned, but it did reveal related guidelines and heuristics. These include mHealth apps [6], privacy of mHealth apps for self-tracking [7], electronic medical records [8,9], personal health records [10], eHealth [11], patient safety [12], and electronic health records [12]. To the best of our knowledge, there is no existing set of guidelines or heuristics that can be used to serve the purpose of evaluating an app that supports self-monitoring, behavior change, and communication with health care professionals.

Once an app is developed, a potential end user of the app or a domain expert can be recruited to evaluate the app. End users are typically exposed to the app towards the last stage of app development; therefore, the expertise of domain experts is utilized to quickly evaluate the app on behalf of end users. In this research, domain experts evaluated the OneTouch app. Methods that can be used for an expert evaluation include heuristic evaluation [13]; cognitive walkthrough [14,15]; goals, operators, methods, and selection [16]; keystroke-level model [17]; or using results of a previous study as the basis to prove or disprove different aspects of the design. Among these methods, heuristic evaluation is not only quick but also inexpensive and easy to perform in comparison to other methods.

This paper presents the development of a specialized set of heuristics based on existing heuristics, called the heuristic evaluation for mHealth apps (HE4EH), that can be used as a tool to design or evaluate an mHealth app. This paper is structured as follows. The Methods section presents the mHealth app design framework. The Results section presents the evaluation of the HE4EH by domain experts and its applicability to evaluate the OneTouch app. Last, the article is concluded in the Discussion section.

Methods

Compilation of Heuristics

To compile the heuristics, we searched for existing sets of mobile heuristics, mHealth heuristics, behavioral change, and self-monitoring of blood glucose in electronic databases, including Scopus, Web of Science, PubMed, and Google Scholar. The research revealed 8 sets, by Shneiderman [18], Nielsen [19], Gómez et al [6], Lacerda et al [20], Dourado and Canedo [21], Monkman and Kushniruk [22], Abraham and Michie [23], and the International Diabetes Federation [24]. Figure 1 presents the compilation of heuristics retrieved from the results. The single word from each heuristic that is written in square brackets is used as the shorter name of the heuristic in the following sections.

XSL•FO
RenderX

Figure 1. Compilation of heuristics.

S#	Heuristics	[18]	[19]	[6]	[20]	[21]	[22]	[23]	[24]
1	[Visibility] of system status	√	√	√	√	√			
2	User [control] and freedom	√	√	√	√	√			
3	[Match] between system and real world		√	√	√	√			
4	[Consistency] and Standards	√	√	√	√	√			
5	[Error] Prevention	√	√	√	√	√			
6	Help users recognize, diagnose, and [recover] from errors		√	√	√	√			
7	[Recognition] rather than recall		√	√		√			
8	[Flexibility] and efficiency of use	√	√	√	√	√			
9	Aesthetic and [minimalist] design		√	√	√	√			
10	Help and [documentation]		√	√	√	√			
11	[Privacy]			√		√			
12	[Skills]			√					
13	[Pleasurable] interaction			√		√			
14	[Accessibility]	√					√		
15	[Compatibility] between different platforms				√				
16	Minimized human/device [interaction]				√				
17	Physical interaction and [ergonomics]				√				
18	[Readability] and layout				√				
19	Non-interruptive app information [visualization]				√				
20	[Content]						√		
21	[Display]						√		
22	[Navigation]						√		
23	[Interactivity]						√		
24	[Behavior] change							√	
25	[Self-monitoring]								√

Intraset Heuristics

There are similar heuristics across the sets; for instance, consistency and standards, strive for consistency, and consistency mean the same thing. Similarly, flexibility and efficiency of use, efficient interaction, customizability, and efficiency also mean the same thing. There are also other examples. Therefore, all such heuristics were grouped together using a concept similar to an affinity diagram.

Interset Checklist Items

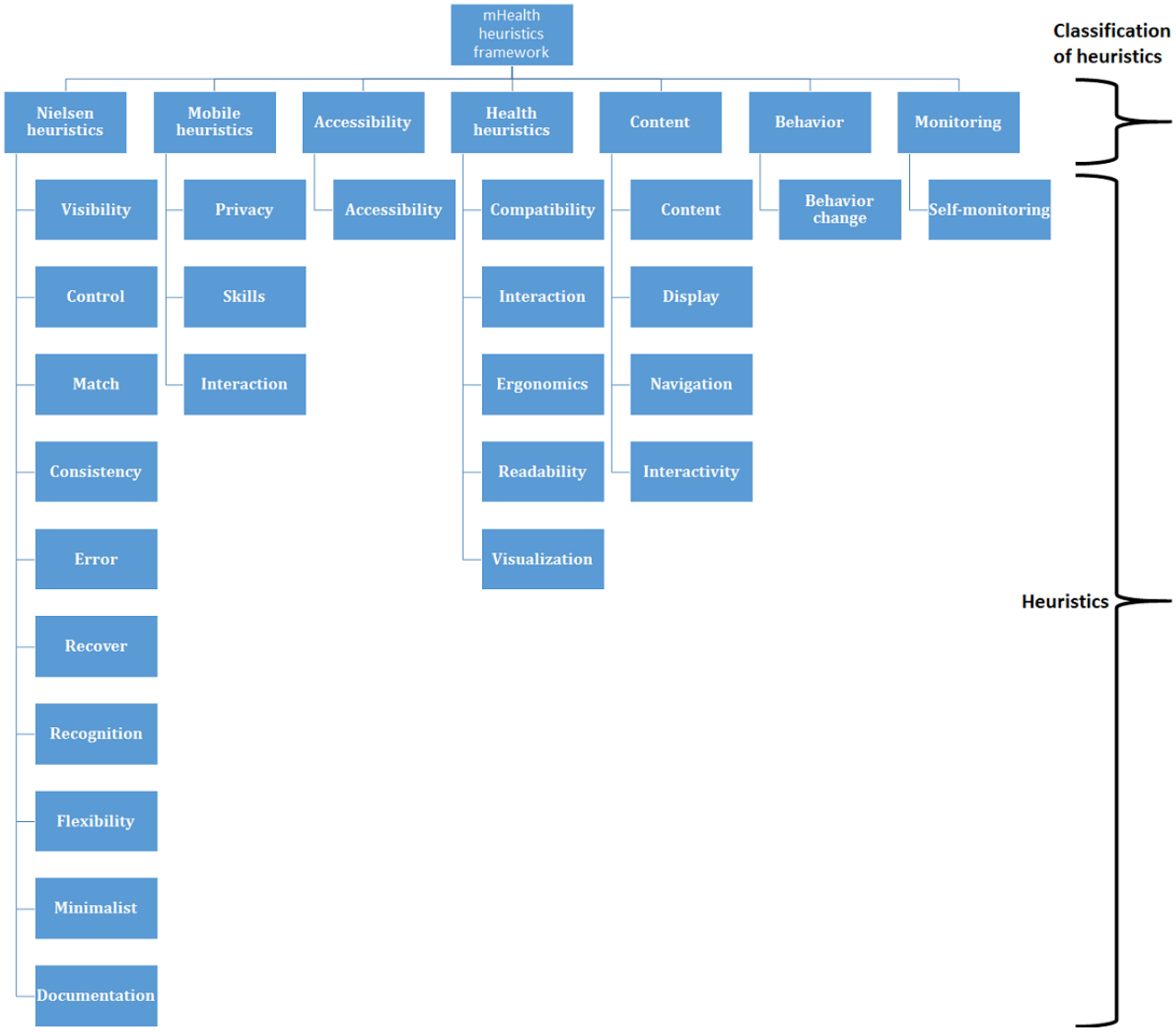
Once all the related heuristics were grouped, it became obvious that there would be multiple related checklist items. The next step was to identify all those related checklist items and group, merge, and expand the checklist items wherever needed. The checklist items were mostly in the form of statements. Each statement was converted into a question, so the checklist items could be used to identify which items satisfied, did not satisfy, and were not applicable for the app to support the researchers and practitioners in evaluating the mHealth app.

mHealth Heuristics Framework

Figure 2 shows the framework of an mHealth heuristics. The 25 heuristics in the set were classified into 7 types of heuristics. The framework consists of a set of components where each component contains specific details and the interconnection between the components form a structure of the framework. Similarly, the heuristics and their checklist items can be associated with the components and their details, respectively, and the interconnection between a group of heuristics to a type of heuristics forms a structure of the framework. Furthermore, the checklist items associated with each heuristic can be adapted to meet more or specific needs. Multimedia Appendix 1 presents the distribution of the heuristics and checklist items. An accessibility heuristic contains one checklist item only. However, there are various accessibility guidelines in the literature that can be used; the most used accessibility guidelines in recent years are the Web Content Accessibility Guidelines 1.0, 2.0, and 2.1. Similarly, in the context of this research, the checklist items of the self-monitoring heuristics are specific to diabetes; however, they can be adapted with a set of guidelines specific to any other chronic disease such as cancer, HIV, or cardiovascular disease or acute, but equally life-threatening, conditions such as ongoing coronavirus disease.

XSL•FO
RenderX

Figure 2. Mobile health (mHealth) heuristics framework.



Expert Evaluation Study 1

This subsection presents the overall study design of the expert evaluation study 1, while the results are presented in the Results section.

Participants and Recruitment

Experts are widely used in research studies to mainly gather their opinion and improve the instrument based on their feedback. Experts were also used in this research to improve the instrument (ie, a set of heuristics and checklist items). All the participants were recruited through convenience sampling and snowball sampling.

Instruments Used

The HE4EH is a proposed set of heuristics that contains 25 heuristics (see Figure 1) and a total of 436 checklist items (See Multimedia Appendix 1).

The prestudy questionnaire included questions related to the demographic information of the experts. The demographic questions included gender; occupation; industry; experience (in years); familiarity with types of diabetes; level of human-computer interaction (HCI), usability, user interface, or user experience design experience; experience working with patients with diabetes; and involvement in a number of projects for patients with diabetes.

The expert review questionnaire consisted of 4 parts, namely heuristics, checklist items, severity ratings, and satisfaction. A brief description of each follows.

The heuristics part of the expert review questionnaire presents all 25 heuristics provided to the experts; for each heuristic, the expert had to answer one question each for the importance of the heuristic, the relevancy of the heuristic, and clarity of the heuristic in the set. Each question was rated using a 5-point Likert scale. The items corresponding to an odd-numbered value for importance were “Not Important,” “Moderately Important,” and “Very Important”; similarly, the items corresponding to an odd-numbered value for relevancy included “Not Relevant,” “Moderately Relevant,” and “Very Relevant.” Last, the items corresponding with odd-numbered values for clarity were “Very Poor,” “Barely Acceptable,” and “Very Good.”

A total of 436 checklist items were presented to the experts in the expert review questionnaire. Similar to the heuristics, for

XSL•FO
RenderX

each checklist item, the expert had to answer one question each for the clarity, grouping, and relevance of the checklist item. Each question for clarity, grouping, and relevance was rating using a 5-point Likert scale. The items associated with odd-numbered values for clarity included "Very Poor," "Barely Acceptable," and "Very Good"; similarly, the items associated with odd-numbered values for grouping were "Strongly Disagree," "Undecided," and "Strongly Agree." Last, the items associated with odd-numbered values for relevance were "Not Relevant," "Moderately Relevant," and "Very Relevant."

Researchers in this space typically use a 5-point severity rating scale where 0 means "No problem," 1 means "Cosmetic," 2 means "Minor," 3 means "Major," and 4 means "Catastrophe." We decided to investigate 3 alternate severity rating scales. For the individual items in the 3 alternative scales, experts were asked to choose a suitable option on a 5-point Likert scale from "Very Difficult" to "Very Easy." Then, we used descriptive statistics to identify the most suitable of the 3 scales.

The first alternative scale was scored as follows: 0, "No Violation"; 1, "Low"; 2, "Moderate"; 3, "High, Severe." The second alternative scale was scored as follows: 0, "Not Applicable"; 1, "No Violation"; 2, "Minor"; 3, "Major" 4, "Catastrophe." The third alternative scale was scored as follows: 0, "No Violation"; 1, "Low"; 2, "Medium"; 3, "High."

The last part of the expert review evaluation asked the expert about their level of satisfaction in terms of using HE4EH for the evaluation of an mHealth app. They were asked to answer 13 questions (see Multimedia Appendix 2) using a 5-point Likert scale from 1 for "Strongly Disagree" to 5 for "Strongly Agree."

Study Protocol

Each expert received a separate email to provide consent for the evaluation of a set of heuristics and the checklist items. In the email, they were briefly informed about the background of the study and its objectives and were introduced to the set of heuristics and its checklist items. The experts who agreed to participate in the study were sent another email with 3 instruments: (1) prestudy questionnaire, (2) heuristics and checklist items, and (3) expert review questionnaire. They were asked to complete the prestudy questionnaire first and then review all the heuristics and associated checklist items. Last, they were asked to complete a comprehensive expert review questionnaire and return the completed questionnaires via email.

Data Analysis

Microsoft Excel software was used to analyze the data collected in this study.

For each heuristic, checklist item, severity rating, and satisfaction rating, the expert was asked to answer one or more questions using a 5-point or 4-point Likert scale. The frequency was calculated as a sum of responses for each Likert-scale score for each question.

Expert Evaluation Study 2

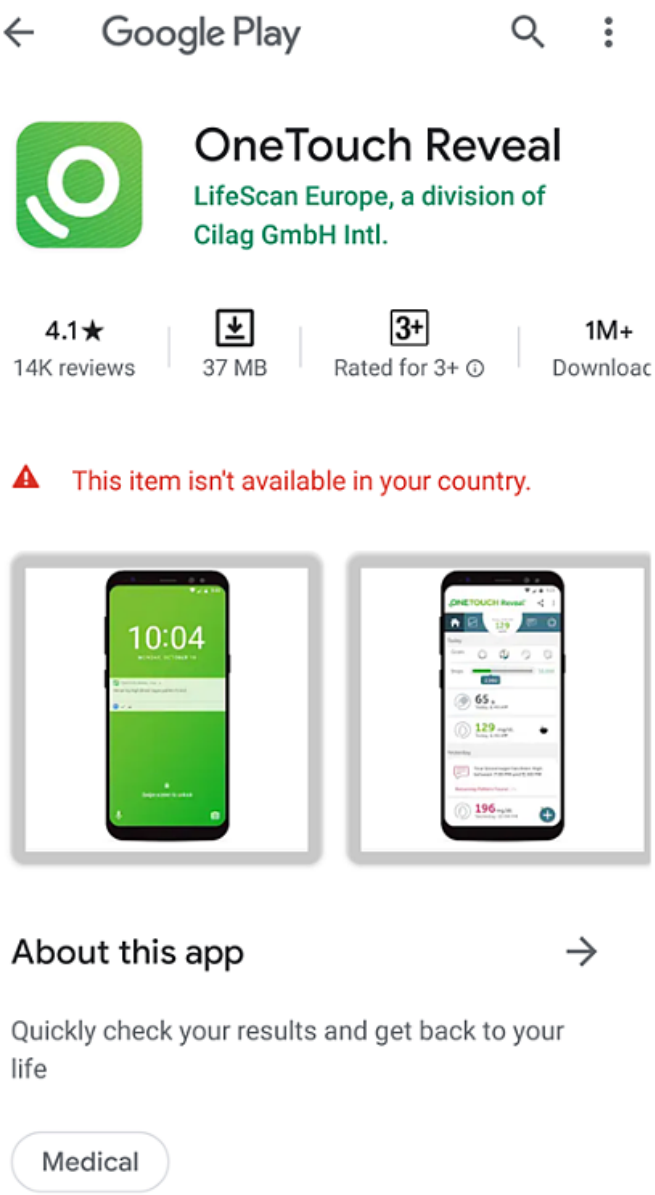
This subsection presents the overall study design of expert evaluation study 2, while the results are presented in the Results section.

Participants and Recruitment

In studies evaluating heuristics, at least 3 participants are recommended, as they can contribute to identifying ≥75% of the usability problems in the app or the system used for the evaluation [13]. In this research, a total of 7 participants took part in the evaluation. The approach used to recruit participants was like that of expert evaluation study 1. Initially, 10 participants were recruited and divided into 2 sets of 5 groups, and only 1 participant was in the group. Three participants using an Android smartphone informed us they were unable to download the app. The Google Play Store was giving the following error message: "This item isn't available in your country" as shown in Figure 3. This shows that access to the app was possibly restricted for certain countries; therefore, the participants from Malaysia and Pakistan could not download the app. It was then decided to proceed with the 7 participants from Qatar. The group-related details are described in Results: Expert Evaluation Study 2.

XSL•FO
RenderX

Figure 3. Error message for the app in the Google Play Store.



App Selection

Websites with rankings of apps to manage diabetes were searched. The selection criteria used to choose a webpage and the app were as follows: (1) webpage provided the names and descriptions of apps to manage diabetes, (2) app was available for both Android and iOS devices, (3) app was free to download, and (4) app did not have in-app purchases.

One of the pages entitled “The best diabetes apps of 2019” [25] was also part of the results. This webpage provided details for 13 apps, of which 3 apps met the aforementioned criteria, namely the Sugar Sense Diabetes app, OneTouch Reveal, and BeatO. The first app was excluded because it was not found in the Google Play Store. The iOS and Android ratings for the remaining 2 apps were 4.7 and 4.0, respectively, and 2.7 and 4.5, respectively. Based on the ratings for both platforms, the OneTouch app was selected for the evaluation in this research. The details of the OneTouch app are described in the next subsection.

Instruments Used

The proposed set of heuristic HE4ED and its checklist were used as the instruments in this research. The details of heuristics and checklists are discussed in detail in the earlier sections.

The OneTouch app helps manage blood glucose levels; this app complements the OneTouch Verio Flex meter. Once the blood glucose readings are added, the app automatically searches for and highlights trends and provides push notifications so the user can take the necessary actions. Physicians have access to their patients’ data; physicians can change their patients’ diet plans based on their history. The app also provides the opportunity to share data about progress with family members and friends. Furthermore, users can log in to their accounts online from their desktop or laptop. This app was carefully chosen because it may be useful for an individual with diabetes and obesity in the Middle East and North Africa region.

The prestudy questionnaire contained questions related to demographic information. The questions included gender, age,

XSL•FO
RenderX

highest degree or level of education, designation (if working), HCI-related courses taken, previously evaluated apps, and set of heuristics or guidelines used.

The usability problem reporting form allowed participants to report identified usability problems with the OneTouch app. For each identified problem, the participant needed to report the broken heuristic and checklist, description of the problem, possible solution from their perspective, and severity rating. The severity rating ranged from 0 to 4 inclusive: "not a problem," "cosmetic problem only," "minor usability problem," "major usability problem," and "usability catastrophe." The typical severity rating scale was chosen for 2 reasons. First, the experts' opinions on the severity rating scales were almost the same. Second, the new ratings may not have given a better picture as each expert only received about 25% of the checklist items. Researchers conducting heuristic evaluation studies typically use the same severity scale mentioned in this subsection rather than the one proposed in this study.

The system usability scale (SUS) by Brooke [26] was used as the poststudy questionnaire. The usability measurements included in the SUS cover effectiveness, efficiency, and user satisfaction. The questionnaire includes 10 statements, and for each statement, the respondent needs to select the best possible choice based on a 5-point Likert scale from "strongly disagree" to "strongly agree."

Study Protocol

We sent an email to the participants to brief them about the set of heuristics developed and the evaluation of the OneTouch app. They were informed that they would be divided into 5 groups, with 1 participant in each group, and groups would be randomly formed. They were asked to give consent by replying to the email. The participants consenting to participant were sent another email with the following:

- Set of instructions to download and install the OneTouch app on their Android or iOS device and to register an account for themselves. They were informed that they would evaluate the app on their own.
- Set of heuristics and either 107 or 108 checklist items (ie, about 25% of the total checklist items). We provided about one-fourth of the checklist items to limit workload and, at the same time, define a minimal set of checklist items that could be useful to evaluate the app. These checklist items were equally distributed in a way that each heuristic has one or more checklist item. There were some similarities in the checklists between groups.
- Pre-study questionnaire, usability problems reporting form for the OneTouch app, and poststudy questionnaire. They were informed to complete these in the following order: prestudy questionnaire, the evaluation of the OneTouch app and reporting of usability problems in the reporting form, and poststudy questionnaire. We requested that they return all completed files via email.

They were also informed that, in case of any confusion or query, they could always communicate via email, and a Skype call could be initiated for discussion, if needed. They were also asked to nominate a potential list of participants who would be interested in evaluating the app.

Data Analysis

Microsoft Excel software was also used to analyze the data collected during expert evaluation study 2.

For each heuristic, the usability problems represent the sum of usability problems found for all the severity ratings.

For each heuristic, the average severity rating represents the average of usability problems found for all the severity ratings.

The standard calculation method for the SUS score was used in this research.

Results

Expert Evaluation Study 1

Table 1 shows the demographic information of all the experts who participated in the study.

Table 1. Demographic information of the experts who participated in expert evaluation study 1.

Participant	Occupation	Industry	Experience (years)	HCI[a] expertise	Projects[b]
Expert 1	Assistant professor	Education	18	Intermediate	N/A[c]
Expert 2	Senior software engineer	IT[d]	10	Intermediate	1
Expert 3	Assistant professor	Education	17	Intermediate	0
Expert 4	Associate professor	Education	20	Expert	2
Expert 5	Associate professor	Education	23	Expert	2

[a]HCI: human-computer interaction.

[b]Number of projects with patients with diabetes.

[c]N/A: not applicable.

[d]IT: information technology.

Heuristic

Multimedia Appendix 3 shows the expert opinions for the 3 questions (importance, relevance, and clarity of the heuristic) asked for all 25 heuristics of the set. The first column shows the serial number, while the second column shows the full name of the heuristic. The remaining 9 columns show the sum of the

XSL•FO
RenderX

opinions for the aforementioned 3 questions. The sum of the expert opinions is shown in 3 columns for each question.

The numeric values in the 3 columns (from left to right) for the importance of a heuristic represent the number of responses for "Very Important," "Important," and "Moderately Important," respectively. Similarly, the numeric values in the next 3 columns (from left to right) for the relevance of a heuristic represent the number of responses for "Very Relevant," "Relevant," and "Moderately Relevant." The numeric values in the last 3 columns (from left to right) for the clarity of a heuristic represent the number of responses for "Very Good," "Good," and "Barely Acceptable." There are only 3 negative responses for the "Little Importance" item of the Likert scale for a question related to the importance of the heuristic; the 3 heuristics marked with an asterisk "*" had 1 negative response each. Therefore, the 2 columns for the negative items on the Likert scale are not shown in the table. Similarly, 1 expert did not answer a question related to the importance and relevance of the heuristic for 3 heuristics; these heuristics are marked with a hash "#."

Of the responses, 90.4% (113/125) showed that the heuristics are important, 4.8% (6/125) showed that they are moderately important, and 2.4% (3/125) showed that heuristics are of little importance, while the remaining 2.4% (3/125) were unanswered.

Of the responses, 93.6% (117/125) showed that the heuristics are relevant, and 4.0% (5/125) of the responses showed that they are moderately relevant, while the remaining 2.4% (3/125) of the responses were unanswered.

Of the responses, 74.4% (93/125) showed that the heuristics are clear, while the remaining 25.6% (93/125) showed that they are barely acceptable.

Checklist Items

Multimedia Appendix 4 shows the mean scores of all 3 questions (ie, clarity, grouping, and relevance) for each checklist item. The means are presented in ranges for each question. The analysis of the mean scores revealed that the minimum mean score for each question of the checklist was 2.8. Therefore, the ranges are presented in Multimedia Appendix 4 in increments of 0.2 (except the last range, which is an increment of 0.21) from 2.8 to 5.0. The decimal value of the checklist items represents the heuristic number in the set (whole number) and the checklist item for that heuristic (fractional number). The results of the questions are discussed in the following paragraphs.

The results show that the experts had a neutral opinion about the clarity of 2 of the 436 (0.5%) checklist items (ie, 3.26 and 17.4). The experts agreed for 173 of the 436 (39.6%) checklist items, while they strongly agreed for the remaining 262 of the 436 (60.0%) checklist items.

In terms of the grouping of checklists for each heuristic, the results show that the experts agreed for 149 of the 436 (34.1%) checklist items, while the experts strongly agreed for the remaining 288 of the 436 (66.0%) checklist items.

The results show that the experts had a neutral opinion about the relevance of 2 of the 436 (0.5%) checklist items (ie, 12.3 and 17.4). The experts agreed for 141 of the 436 (32.3%) checklist items, while they strongly agreed for the relevance of the remaining 294 of the 436 (67.4%) checklist items.

Severity Rating Scales

Multimedia Appendix 5, Multimedia Appendix 6, and Multimedia Appendix 7 present the results of 3 different severity rating scales in terms of the mean individual rating or score and the overall mean of all the ratings or scores. The results show that there is a subtle difference between all 3 rating scales.

Satisfaction

Multimedia Appendix 2 shows the level of satisfaction with using the HE4EH to evaluate an mHealth app. The numeric values in the 5 columns (from left to right) indicate the sum of the responses for "Strongly Agree," "Agree," "Neutral," "Disagree," and "Strongly Disagree" for each question.

According to the experts, sufficient or enough material was used to develop te HE4EH; it is not complex to use and provides useful information to design or evaluate an mHealth app. The terminologies used in the HE4EH are clear and easy to understand, learn, and use; the length of the HE4EH is suitable for the evaluation of mHealth apps that aim to improve the health of patients, especially those with diabetes in the context of this research, and allow them to self-monitor and change their behaviors. They were satisfied with the number of checklist items and their categorization, and they could use them to evaluate an app.

Expert Evaluation Study 2

The demographic information of the recruited participants is presented in Table 2.

The results of the heuristic evaluation study are shown in Figure 4. The stacked columns show the number of usability problems found for each heuristic under zero, one, or more severity levels, while the line with markers shows the average severity rating of all the usability problems found for each heuristic. The horizontal axis shows the 25 heuristics of the set, and there are 2 vertical axes. The vertical axis on the left is the primary axis and represents the number of usability problems found, while the vertical axis on the right side shows the severity rating.

XSL•FO
RenderX

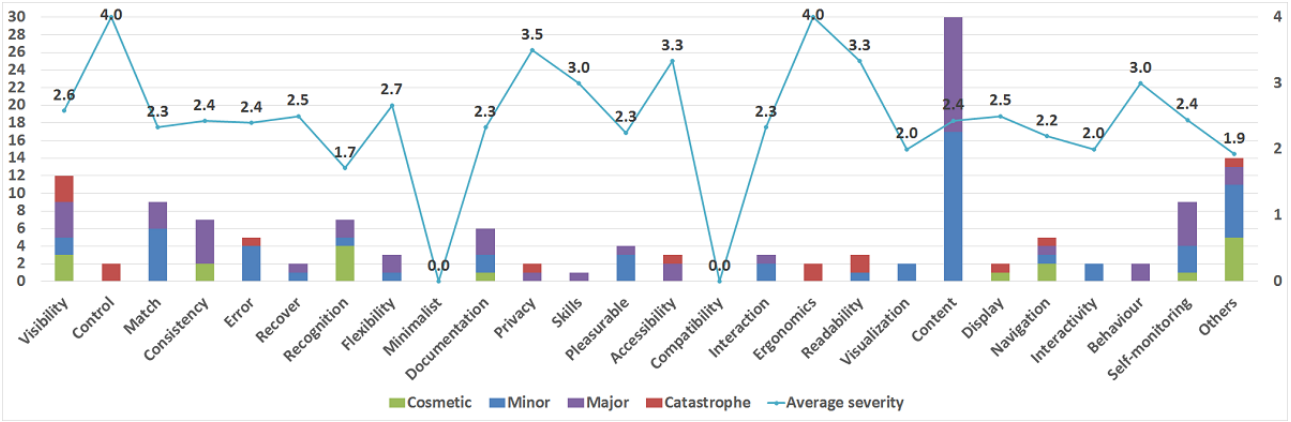
Table 2. Details of the participants in the expert evaluation study 2.

Parameter	Results, n (%)
Gender	
Male	2 (29)
Female	5 (71)
Age group (years)	
21-25	3 (43)
26-30	1 (14)
31-35	1 (14)
36-40	2 (29)
Highest degree	
Bachelor's degree	1 (14)
Master's degree	4 (57)
Doctorate	2 (29)
Employment status	
Full-time student	5 (71)
Full-time employment and part-time student	2 (29)
HCI[a] course[b]	7 (100)
Previous app evaluation using a set of heuristics or guidelines	
1-4	3 (43)
≥5	4 (57)

[a]HCI: human-computer interaction.

[b]Courses included HCI, interactive design, health care interaction design, and information visualization.

Figure 4. Usability problems identified using a proposed set of heuristics.



The cumulative number of usability problems found by all the groups was 137: severity rating 4: 15/137, 11.0%; severity rating 3: 49/137, 35.8%; severity rating 2: 54/137, 39.4%; severity rating 1: 19/137, 13.9%.

Usability Problems

The results show that one or more usability problems were found for most of the heuristics (23/25, 92%). The most frequently broken heuristics included content (30/137), visibility (12/137), and match and self-monitoring (9/137 each).

During the process of dividing the checklists into the groups, it was expected that some groups might find a usability problem in the app but that they may not find suitable checklist items to which they could map the problem. This was expected because each group received about 25% of the entire checklist. During the analysis of the usability problems, the groups found 14 usability problems for which they could not find a matching checklist item. In Figure 3, these usability problems are presented in the last stacked column labeled Others.

Multimedia Appendix 8 presents the details of some of the usability problems found by the participants. The details presented include heuristic number, problem description, the potential solution from their perspective, and severity rating. An "OT" in the first column of the table means the participants

XSL•FO
RenderX

were unable to find a matching heuristic for the mentioned usability problems.

Average Severity Ratings

The average severity rating of all usability problems found was 2.4; this shows that most of the usability problems found were minor.

System Usability Scale (SUS)

Table 3 shows the SUS scores of each participant for the OneTouch app. The scores ranged from 42.5 through 75, with an average score of 60.

Table 3. System usability scale (SUS) scores.

Participant ID	SUS score
1	50.0
2	70.0
3	47.5
4	72.5
5	62.5
6	75.0
7	42.5

There are different ways to interpret the SUS scores [27]; these include percentiles, grades, adjectives, acceptability, and promoters and detractors. In this research, SUS scores were interpreted using acceptability, which defines the score in terms of what is "acceptable" or "not acceptable."

According to the grading scale interpretation of SUS scores by Bangor and colleagues [28], the OneTouch app is Marginally acceptable. This shows that the OneTouch app needs enhancements to improve the usability of the app.

Discussion

This research presents a modified set of heuristics called the HE4EH and its applicability by evaluating the OneTouch app. The HE4EH consists of Nilsen's heuristic, mobile heuristics, health heuristics, and other heuristics (behavior change and self-monitoring of health). The HE4EH is a tool to assist HCI experts and mobile app developers when designing or evaluating an mHealth app for individuals with diabetes, by measuring the usability problems that can influence the user experience with the app. First, an expert evaluation study was conducted to improve heuristics and associated checklist items. Then, another expert evaluation study was conducted to evaluate a popular app called OneTouch intended to help with blood glucose management. The findings of the usability evaluation are as follows:

- Although the participants had a partial list of checklist items, they were able to identify the usability problems and link them to the given checklist items.
- The participants were also able to identify usability problems that could not be mapped to the partial list of given checklist items. This shows that, if given a complete list of checklist items, more usability problems could be identified.
- The top 3 frequently violated heuristics included Content, followed by Visibility, Match, and Self-monitoring.
- Although minimal, the participants were able to identify usability problems in the OneTouch app associated with the heuristics incorporated explicitly for behavior change and self-monitoring of blood glucose. The lack of incorporation of these and related heuristics and the associated checklist items means the identified usability problems had remained undetected.
- The average severity rating of all the usability problems found was minor.
- The mean SUS score showed that the OneTouch app is marginally acceptable and needs enhancements to improve the overall usability of the app.

Limitations

As with all studies, this research has a few limitations. First, we used only selected bibliographic databases and search terms to identify a set of related heuristics. Second, a limited set of checklist items was given to each expert in the study. Last, the severity rating scale typically used in heuristic evaluation studies was also used in this research due to a subtle difference between the severity rating scales proposed in this research.

Future Work

The current research can be extended in different ways. First, we intend to evaluate the mHealth apps by giving a complete set of checklist items to the HCI experts and investigate the impact on identifying usability problems in the apps, especially in terms of the specialized heuristics added to the HE4EH. They can consider using one of the severity rating scales discussed in this research. Second, future research can evaluate the effectiveness of different severity ratings proposed in this research. Third, the accessibility heuristic contains one generic checklist item; future research can, therefore, enrich it with a specialized set of guidelines like the Web Content Accessibility Guidelines [29] and then investigate the accessibility of mHealth apps. Similarly, checklist items for heuristics can also be adapted (eg, for privacy [7]). Fourth, future research can adapt the HE4EH and incorporate specific types of heuristics with more checklist items not covered in the set, especially in terms of controlling blood glucose monitoring. Fifth, future research can also adapt the HE4EH for behavioral change and self-monitoring

XSL•FO
RenderX

of other health-related issues such as obesity and depression, among others. Last, future research can also conduct a qualitative study with domain experts to identify the heuristics or checklist items that can be combined to reduce the number of heuristics or checklist items.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Set of heuristics and checklist items.
[DOCX File , 45 KB - mhealth_v8i10e20353_app1.docx]

Multimedia Appendix 2
Responses to the experts' satisfaction questionnaire.
[DOCX File , 25 KB - mhealth_v8i10e20353_app2.docx]

Multimedia Appendix 3
Expert opinion on the heuristics.
[DOCX File , 28 KB - mhealth_v8i10e20353_app3.docx]

Multimedia Appendix 4
Mean score of the checklist items, as scored by the experts.
[DOCX File , 27 KB - mhealth_v8i10e20353_app4.docx]

Multimedia Appendix 5
Mean scores for severity rating scale 1.
[DOCX File , 23 KB - mhealth_v8i10e20353_app5.docx]

Multimedia Appendix 6
Mean scores for severity rating scale 2.
[DOCX File , 23 KB - mhealth_v8i10e20353_app6.docx]

Multimedia Appendix 7
Mean scores for severity rating scale 3.
[DOCX File , 23 KB - mhealth_v8i10e20353_app7.docx]

Multimedia Appendix 8
Participants' comments about issues with the heuristics.
[DOCX File , 25 KB - mhealth_v8i10e20353_app8.docx]

References

1. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. Telemed J E Health 2014 Jan;20(1):75-82 [FREE Full text] [doi: 10.1089/tmj.2012.0328] [Medline: 24205809]
2. Diabetes. World Health Organization. 2020 Jun 08. URL: https://www.who.int/news-room/fact-sheets/detail/diabetes [accessed 2020-07-05]
3. Awad SF, O'Flaherty M, Critchley J, Abu-Raddad LJ. Forecasting the burden of type 2 diabetes mellitus in Qatar to 2050: A novel modeling approach. Diabetes Res Clin Pract 2018 Mar;137:100-108 [FREE Full text] [doi: 10.1016/j.diabres.2017.11.015] [Medline: 29175341]
4. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
5. Miller AS, Cafazzo JA, Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2016 Jun 10;22(2):184-193. [doi: 10.1177/1460458214537511] [Medline: 24986104]
6. Yáñez Gómez R, Cascado Caballero D, Sevillano J. Heuristic evaluation on mobile interfaces: a new checklist. Scientific World Journal 2014;2014:434326 [FREE Full text] [doi: 10.1155/2014/434326] [Medline: 25295300]

XSL•FO
RenderX

7. Hutton L, Price BA, Kelly R, McCormick C, Bandara AK, Hatzakis T, et al. Assessing the Privacy of mHealth Apps for Self-Tracking: Heuristic Evaluation Approach. JMIR Mhealth Uhealth 2018 Oct 22;6(10):e185 [FREE Full text] [doi: 10.2196/mhealth.9217] [Medline: 30348623]
8. HIMSS EHR Usability Task Force. Defining and Testing EMR Usability: Principles and Proposed Methods of EMR Usability Evaluation and Rating. Healthcare Information Management Systems Society 2009 Jun:1 [FREE Full text]
9. Kuqi K, Eveleigh T, Holzer T, Sarkani S, Levin J, Crowley R. Design of electronic medical record user interfaces: a matrix-based method for improving usability. J Healthc Eng 2013 Sep;4(3):427-452 [FREE Full text] [doi: 10.1260/2040-2295.4.3.427] [Medline: 23965597]
10. Tzeng SW, Zhou Y. Design Guidelines for an Integrated PHR System: An Approach for UI Designers to Break Down Individual-Level Barriers to PHR Adoption. In: Stephanidis C, Antona M, editors. Universal Access in Human-Computer Interaction. Applications and Services for Quality of Life. UAHCI 2013. Lecture Notes in Computer Science, vol 8011. Berlin, Heidelberg: Springer Publishing Company; 2013.
11. Baumel A, Muench F. Heuristic Evaluation of Ehealth Interventions: Establishing Standards That Relate to the Therapeutic Process Perspective. JMIR Ment Health 2016 Jan 13;3(1):e5 [FREE Full text] [doi: 10.2196/mental.4563] [Medline: 26764209]
12. Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. Journal of Biomedical Informatics 2003 Feb;36(1-2):23-30. [doi: 10.1016/s1532-0464(03)00060-1]
13. Nielsen J, Molich R. Heuristic evaluation of user interfaces. 1990 Apr Presented at: SIGCHI Conference on Human Factors in Computing Systems; March 1990; Seattle, WA p. 249-256. [doi: 10.1145/97243.97281]
14. Wharton C, Rieman J, Lewis C, Polson P. The cognitive walkthrough: a practitioner's guide. In: Nielsen J, Mack RL, editors. Usability Inspection methods. New York, NY: John Wiley & Sons, Inc; Jun 1994.
15. Polson PG, Lewis C, Rieman J, Wharton C. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. International Journal of Man-Machine Studies 1992 May;36(5):741-773. [doi: 10.1016/0020-7373(92)90039-n]
16. Card SK, Moran TP, Newell A. The psychology of human-computer interaction. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 1983.
17. Card SK, Moran TP, Newell A. The keystroke-level model for user performance time with interactive systems. Commun. ACM 1980 Jul;23(7):396-410. [doi: 10.1145/358886.358895]
18. Shneiderman B, Plaisant C, Cohen M, Jacobs S, Elmqvist N, Diakopoulos N. Designing the User Interface: Strategies for Effective Human-Computer Interaction. London, England: Pearson; 2016.
19. Nielsen J. 10 Usability Heuristics for User Interface Design. 1994 Apr 24. URL: http://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2020-02-25]
20. Lacerda TC, Nunes JV, Gresse von Wangenheim C. Usability Heuristics for Mobile Phone Applications: A Literature Review. In: Lumsden J, editor. Emerging Perspectives on the Design, Use, and Evaluation of Mobile and Handheld Devices. Hershey, PA: IGI Global; Jul 15, 2015:143-157.
21. Dourado MAD, Canedo ED. Usability Heuristics for Mobile Applications: A Systematic Review. Proceedings of the 20th International Conference on Enterprise Information Systems (ICEIS 2018) 2019:483-494. [doi: 10.5220/0006781404830494]
22. Monkman H, Kushniruk A. A health literacy and usability heuristic evaluation of a mobile consumer health application. Stud Health Technol Inform 2013;192:724-728. [Medline: 23920652]
23. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
24. Self-monitoring of blood glucose in non-insulin treated type 2 diabetes. International Diabetes Federation. 2017 Nov 04. URL: https://www.idf.org/e-library/guidelines/85-self-monitoring-of-blood-glucose-in-non-insulin-treated-type-2-diabetes.html [accessed 2020-10-05]
25. Doyle A. The Best Diabetes Apps of 2020. Healthline. 2020. URL: https://www.healthline.com/health/diabetes/top-iphone-android-apps [accessed 2020-10-05]
26. Brooke J. SUS - a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press; Jun 10, 1996:189-194.
27. Sauro J. 5 Ways to Interpret a SUS Score. Measuring U. 2018 Sep 19. URL: https://measuringu.com/interpret-sus-score/ [accessed 2020-02-27]
28. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
29. Web Content Accessibility Guidelines (WCAG) Overview. Web Accessibility Initiative. 2020 Sep 22. URL: https://www.w3.org/WAI/standards-guidelines/wcag/ [accessed 2020-10-05]

Abbreviations

HCI: human-computer interaction
HE4EH: heuristic evaluation for mHealth apps
mHealth: mobile health

XSL•FO
RenderX

SUS: system usability scale

Edited by G Eysenbach; submitted 18.05.20; peer-reviewed by RJ Yusof, D Szinay; comments to author 08.06.20; revised version received 31.08.20; accepted 22.09.20; published 28.10.20.

Please cite as:
Khowaja K, Al-Thani D
New Checklist for the Heuristic Evaluation of mHealth Apps (HE4EH): Development and Usability Study
JMIR Mhealth Uhealth 2020;8(10):e20353
URL: http://mhealth.jmir.org/2020/10/e20353/
doi:10.2196/20353
PMID:33112252

©Kamran Khowaja, Dena Al-Thani. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 28.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Engagement, Acceptability, Usability, and Preliminary Efficacy of a Self-Monitoring Mobile Health Intervention to Reduce Sedentary Behavior in Belgian Older Adults: Mixed Methods Study

Sofie Compernolle[1], PhD; Greet Cardon[1], PhD; Hidde P van der Ploeg[2], PhD; Femke Van Nassau[2], PhD; Ilse De Bourdeaudhuij[1], PhD; Judith J Jelsma[2], PhD; Ruben Brondeel[1,3], PhD; Delfien Van Dyck[1], PhD

[1]Department of Movement and Sport Sciences, Ghent University, Ghent, Belgium

[2]Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[3]Research Foundation Flanders, Brussels, Belgium

Corresponding Author:
Sofie Compernolle, PhD
Department of Movement and Sport Sciences
Ghent University
Watersportlaan 2
Ghent, 9000
Belgium
Phone: 32 092646323
Email: sofie.compernolle@ugent.be

Abstract

Background: Although healthy aging can be stimulated by the reduction of sedentary behavior, few interventions are available for older adults. Previous studies suggest that self-monitoring might be a promising behavior change technique to reduce older adults' sedentary behavior. However, little is known about older adults' experiences with a self-monitoring–based intervention aimed at the reduction of sedentary behavior.

Objective: The aim of this study is to evaluate engagement, acceptability, usability, and preliminary efficacy of a self-monitoring–based mHealth intervention developed to reduce older adults' sedentary behavior.

Methods: A mixed methods study was performed among 28 community-dwelling older adults living in Flanders, Belgium. The 3-week intervention consisted of general sedentary behavior information as well as visual and tactile feedback on participants' sedentary behavior. Semistructured interviews were conducted to explore engagement with, and acceptability and usability of, the intervention. Sitting time was measured using the thigh-worn activPAL (PAL Technologies) accelerometer before and after the intervention. System usage data of the app were recorded. Quantitative data were analyzed using descriptive statistics and paired-samples *t* tests; qualitative data were thematically analyzed and presented using pen profiles.

Results: Participants mainly reported positive feelings regarding the intervention, referring to it as motivating, surprising, and interesting. They commonly reported that the intervention changed their thinking (ie, they became more aware of their sedentary behavior) but not their actual behavior. There were mixed opinions on the kind of feedback (ie, tactile vs visual) that they preferred. The intervention was considered easy to use, and the design was described as clear. Some problems were noticed regarding attaching and wearing the self-monitoring device. System usage data showed that the median frequency of consulting the app widely differed among participants, ranging from 0 to 20 times a day. No significant reductions were found in objectively measured sitting time.

Conclusions: Although the intervention was well perceived by the majority of older adults, no reductions in sitting time were found. Possible explanations for the lack of reductions might be the short intervention duration or the fact that only bringing the habitual sedentary behavior into conscious awareness might not be sufficient to achieve behavior change.

Trial Registration: ClinicalTrials.gov NCT04003324; https://tinyurl.com/y2p4g8hx

(JMIR Mhealth Uhealth 2020;8(10):e18653) doi:10.2196/18653

KEYWORDS

mHealth; older adults; self-monitoring; perceptions; engagement; acceptability; mixed methods

XSL•FO
RenderX

Introduction

The aging population continues to expand rapidly. Estimates indicate that the global number of adults over the age of 65 years will nearly double from a current population of about 800 million to approximately 1.5 billion in 2050 [1]. This unprecedented population boom poses a major public health challenge. Aging will present an economic burden on society because of increased needs resulting from age-related decline of physical, mental, and cognitive health [2]. To maintain the quality of life of older adults while living independently, healthy aging has become a main priority in the field of public health.

Up until now, the majority of efforts to facilitate healthy aging have been focused on increasing moderate- to vigorous-intensity physical activity [3] but have neglected sedentary behavior. However, both physical inactivity and high levels of sedentary time have been shown to be significantly related to detrimental health effects, like an increased risk for all-cause mortality, noncommunicable diseases [4], and geriatric syndromes, such as physical and cognitive impairments [5,6]. Research has indicated that an increase in moderate- to vigorous-intensity physical activity is often not sufficient to offset the negative health consequences of high levels of sedentary behavior [7]. Given the negative health consequences and the high prevalence of sedentary behavior in older adults (ie, 60 years and over) [8], creating interventions specifically focusing on the reduction of sedentary behavior is recommended to promote healthy aging.

Existing sedentary behavior interventions have mainly focused on social-cognitive models of behavioral change (eg, theory of planned behavior) [9,10]. However, most of these models are based on an expectancy-value framework in which behavior is determined by expected outcomes and the value that is placed on them [11]. As such, these models do not adequately capture processes underlying unintentional and habit-like behavior. Given that a large part of older adults' sedentary behavior is habitual, specific strategies are needed to better control sedentary behavior. One might, for example, change the circumstances, so that habit cueing does not occur anymore [12], or alter external cues that lead to habit execution [13]. These strategies are rather manipulative and often impossible; therefore, they are not always ethical [14,15]. Another way to disrupt undesired habits is preferred, namely by bringing habitual behavior and its context into conscious awareness. This might be achieved by means of self-monitoring [16].

Self-monitoring, which is defined as keeping a record of a specified behavior as a method for changing behavior [16], has been identified as a promising behavior change technique to reduce sedentary behavior in adults [10,17]. A recent meta-analysis, in which interventions including self-monitoring were summarized that aimed to reduce sedentary behavior, showed a significant reduction in total sedentary time [17]. Specifically, an overall mean difference of 34.37 min/day (95% CI 14.48-54.25) was found for total sedentary time between intervention and control groups. It is important to note, however, that the majority of the included interventions targeted young and middle-aged adults. Only four studies targeted older adults with a mean age above 60 years. Of these four studies, only one used an electronic self-monitoring device to provide information on older adults' sedentary behavior, namely the Fitbit One. As the Fitbit One is worn on the wrist, the validity of the sedentary behavior information can be questioned.

Given the limited quantity and quality of existing research on this topic, it remains unclear how older adults experience and use self-monitoring–based mobile health (mHealth) interventions specifically developed to reduce sedentary behavior. However, this information is essential to inform decisions on the development of future interventions. The conceptual model by Perski et al has indicated that user engagement (ie, the combination of subjective experiences characterized by attention, interest, and affect, and objectively measured intervention usage) is assumed to moderate the influence of the mHealth intervention on the mechanisms of action [18]. Next to user engagement, other aspects of acceptability (ie, how well older adults perceived the intervention and the extent to which the intervention met their needs), such as perceived relevance, satisfaction, and perceived usefulness, as well as usability (ie, the extent to which the intervention could be used by other older adults to reduce their sedentary behavior), also contribute to an individual's motivation to continue using the app [19].

Therefore, the overall aim of this study is to gain insight into older adults' experiences with, and the use of, a self-monitoring–based mHealth intervention specifically developed to reduce sedentary behavior. As both qualitative and quantitative data are essential to fully understand concepts such as user engagement, a mixed methods study is used. Moreover, preliminary efficacy of the intervention on older adults' objectively assessed sedentary behavior is examined to get a first indication of the effect size.

Methods

Participants and Design

A convergence model with a triangulated mixed methods approach was conducted to gain in-depth comprehension in the user engagement, acceptability, and usability of an mHealth intervention aimed at the reduction of sedentary behavior. This methodology allowed us to compare, corroborate, or relate quantitative data (ie, system usage and activity monitor data) and qualitative data (ie, interview data). Quantitative and qualitative data were analyzed separately, followed by an integrated interpretation of the results. Participants in the mixed methods study were recruited in Flanders, Belgium, from February to May 2019 using convenience sampling. Recruitment continued until data were saturated (ie, until no new themes emerged in additional interviews). Firstly, an advertisement was distributed via Facebook, and secondly, the advertisement was electronically sent to older adults who were included in a previous study by our research group and who had expressed interest in future studies. To be eligible for this study, participants needed to (1) be at least 60 years old, (2) be Dutch speaking, (3) be able to walk 100 meters without severe difficulties, and (4) have a smartphone. The study was registered at ClinicalTrials.gov (Identifier: NCT04003324) and was approved by the Committee of Medical Ethics of the Ghent

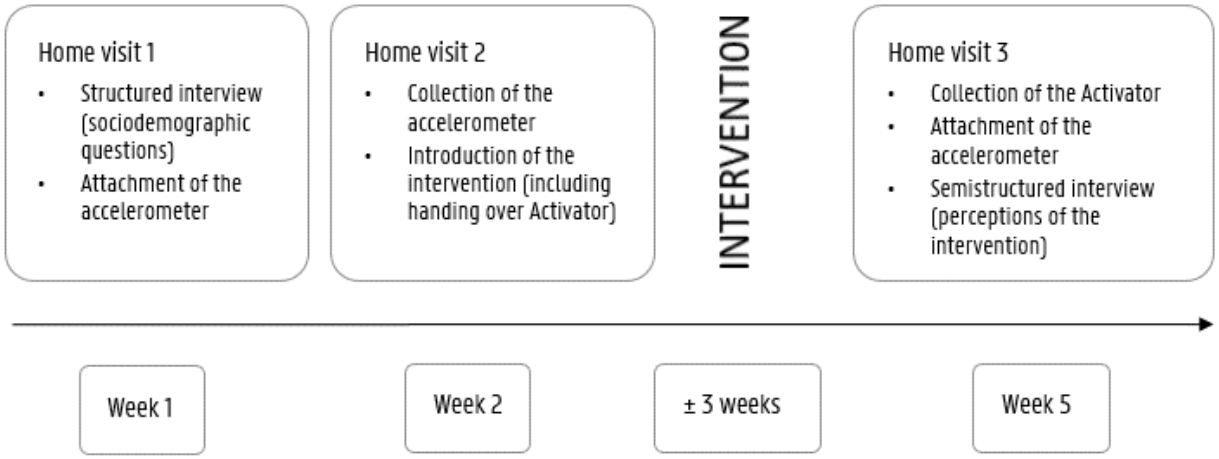
XSL•FO
RenderX

University Hospital (Belgian registration number: 2019/0398). All participants provided written informed consent.

Procedure

The study procedure is explained in Figure 1. Concretely, older adults who agreed to participate were contacted by phone to make an appointment for a first home visit. During this home visit, they received an information letter explaining the purpose of the study and an informed consent form. After signing the informed consent form, baseline measures were collected. Specifically, a structured interview was conducted to assess participants' sociodemographic characteristics. Moreover, an accelerometer—activPAL (PAL Technologies)—was attached to the participants' thighs to objectively measure their sedentary behavior. Participants were instructed to wear the accelerometer for 1 week and to fill out the accompanying diary. After 1 week, a researcher visited the participants again at their homes to collect the accelerometers. After baseline measurements, the self-monitoring–based mHealth intervention was introduced to the participants (see Self-Monitoring mHealth Intervention section). At the end of the intervention (ie, after 3 weeks), participants were asked to complete a semistructured interview that included questions on user engagement with the intervention and perceptions regarding usability and acceptability. At the end of this last home visit, participants were instructed to wear the accelerometer for another week (ie, postmeasurements). Participants were given a prestamped envelope and were asked to send the accelerometer back by postal mail.

Figure 1. Study procedure.

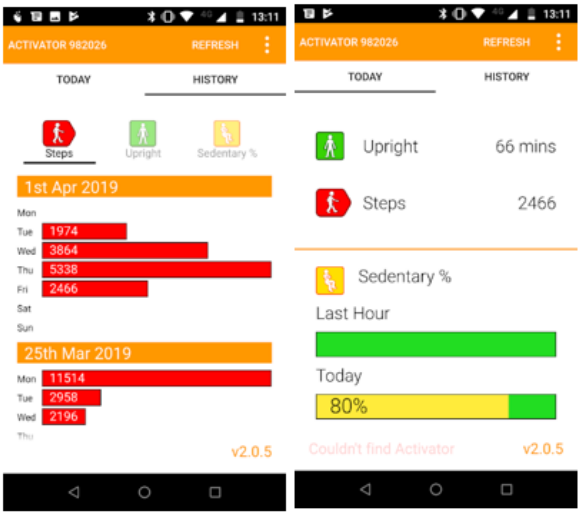


Self-Monitoring mHealth Intervention

The intervention consisted of general sedentary behavior information as well as visual and tactile feedback on participants' sedentary behavior. General sedentary behavior information was provided to participants by means of a 10-minute presentation. The presentation was given by an expert in the field during the second home visit. Visual and tactile feedback were provided using a novel self-monitoring device—the Activator (PAL Technologies). The Activator has recently been validated by Gill et al [20]. The Activator is worn on the front of the thigh, either in a pants pocket or attached with an elastic band to clothing covering the upper thigh (eg, trousers, jeans, shorts, leggings, tights, or dresses), and provides visual and tactile feedback [21]. Visual feedback is presented through a smartphone app via Bluetooth connection. Both real-time feedback and a 7-day historical overview are presented based on participants' sedentary time, upright time, and number of steps (see Figure 2). Visual feedback is constantly available and can be viewed whenever and as often as participants want. Tactile feedback is provided by means of a strong, but comfortable, vibration of the Activator device itself each time a participant is sitting for 30 uninterrupted minutes. If a participant remains sedentary, the vibration is repeated after another 30 minutes. Participants were able to turn the vibration function on and off.

XSL•FO
RenderX

Figure 2. Screenshots of visual feedback provided by the Activator.



Measurements

Structured Interview

Participants' sociodemographic characteristics were collected by a trained researcher during the first home visit. Sociodemographic characteristics included age, gender, family situation (ie, being single or a widow or widower, having a partner but living independently, living with a partner, or being married), number of children, number of grandchildren, residential area (ie, countryside, village, city suburb, or city), educational level (ie, no education, primary education, vocational secondary education, technical secondary education, general secondary education, college, or university), and employment status (ie, employed or not employed).

Activity Monitor

Total sedentary time, sit-to-stand transitions, standing time, and number of steps were objectively estimated by means of the activPAL accelerometer. The accelerometer was attached on the midline of the right anterior thigh. Participants were instructed to wear the accelerometer for 7 consecutive days (24 h/day), both at baseline and at postmeasurement. The activPAL accelerometer summarizes data in 15-second intervals and has shown to be a valid and reliable measure for estimating the time spent sitting, standing, and stepping [19]. The activPAL data were downloaded using activPAL3 software, version 7.2.38, and were then processed using ProcessingPAL, version 1.1 (University of Leicester, UK). This software uses a validated algorithm to separate valid waking wear data from sleep and nonwear data. A day was considered invalid if there was limited postural variation (ie, ≥95% of wear time in one activity), a limited number of steps (<500 steps/day), or fewer than 10 hours of valid waking wear time [22,23]. Summary data from the algorithm were quality checked using heat maps against participants' diaries, and corrections were made where needed [22,23]. Only participants with at least 5 days of valid activPAL data on both time points were included in the analyses [24].

Diary Log

Participants were asked to indicate sleep time (ie, time they went to bed and got up) and nonwearing time of the activPAL in a diary during the 7 days of baseline measurement and postmeasurement.

Semistructured Interview

Semistructured face-to-face interviews were conducted by trained researchers to explore (1) user engagement with the intervention and (2) the usability and acceptability of the intervention. User engagement was defined as the subjective experience of older adults with the intervention characterized by attention, interest, and affect. Acceptability was assessed by asking questions on how well the older adults perceived the intervention and by evaluating the extent to which the intervention met their needs. Usability included questions on the extent to which the intervention could be used by other older adults to reduce their sedentary behavior. The interview guide (see Multimedia Appendix 1) was developed by the first author (SC) based on an extensive literature search and on previous research by our research group examining user engagement, acceptability, and usability of eHealth and mHealth interventions [25]. Conceptual frameworks identified from the literature search, such as the conceptual framework of direct and indirect influences on engagement with digital behavior change interventions (DBCIs) by Perski et al [18] and the behavioral intervention technology (BIT) model of Mohr et al [26], guided the construction of the interview guide. The DBCI-related framework is an integrative conceptual framework involving potential direct and indirect influences on engagement and relationships between engagement and intervention effectiveness. The BIT model conceptually defines BITs, from the clinical aim to the technological delivery framework.

After thorough discussion with the last author (DVD), the interview guide was revised and pilot-tested with two older adults. Based on the pilot test, some minor changes were made, such as paraphrasing and simplifying some vocabulary. By doing so, the clarity of the questions was verified and the duration of the interview was estimated. Interviews were audio recorded (mean duration 11.11 minutes, SD 5.94) and transcribed verbatim, producing a document of 122 pages in length, using Calibri font, size 11.

XSL•FO
RenderX

System Usage Data

System usage data of the Activator (ie, the app) were stored on the cloud server of PAL Technologies and used to objectively estimate user engagement; data included (1) the number of days the Activator was worn and (2) the number of times the app was accessed.

Data Analysis

Descriptive statistics were used to describe the baseline characteristics of the participants and to assess the extent of usage (ie, engagement). Qualitative data were thematically analyzed, using the NVivo 12 software package (QSR International), using the six-phase approach by Braun and Clarke [27] to gain insight into participants' subjective experiences with the intervention (ie, engagement) and acceptability and usability of the intervention. More specifically, two researchers outside the project team—Charlotte Meersseman and Siel Mechelinck—read and reread the transcripts multiple times to become familiar with the data (phase 1). They independently coded the data line by line and defined an initial coding scheme using an inductive approach (phase 2). The coding schemes were then discussed with the first author (SC). By doing so, the triangulation technique was applied and the trustworthiness and validity of the findings were promoted. Based on the coding schemes, themes were searched (phase 3), reviewed (phase 4), and defined (phase 5). Subsequently, pen profiles (ie, diagrams of composite key emergent themes, frequency data, and verbatim quotes) were constructed based on the defined themes and results were written up (phase 6). This increasingly utilized technique is considered appropriate for presenting qualitative outcome data in a clear and useful manner [28]. Paired-samples *t* tests were performed to determine preliminary efficacy of the intervention on older adults' sedentary time. All quantitative analyses were conducted in SPSS Statistics for Windows, version 25.0 (IBM Corp).

Results

Participants

A total of 36 older adults expressed interest in participation. Out of these 36 participants, 2 (6%) of them could not be reached to make an appointment for the first home visit and 4 (11%) decided to withdraw from the study after receiving detailed study information. Reasons for withdrawal were health problems (2/36, 6%), lack of time (1/36, 3%), and death of a spouse (1/36, 3%). As such, 30 older adults completed the baseline measurements. Out of these 30 participants, 2 (7%) of them were excluded, as baseline data showed that they did not fulfill the inclusion criteria (ie, they were not able to walk 100 meters without severe difficulties). In addition, 2 (7%) participants dropped out during the intervention period due to health problems (1/30, 3%) and lack of motivation (1/30, 3%). Consequently, posttest data were collected from 26 out of 28 participants (93% retention).

Baseline characteristics of the participants are presented in Table 1. Just over half of the participants (15/28, 54%) were female, the average age was 65.0 years (SD 4.6), and the mean BMI was 25.4 kg/m^2 (SD 3.9). The majority of the participants were highly educated and were married or lived with a partner.

XSL•FO
RenderX

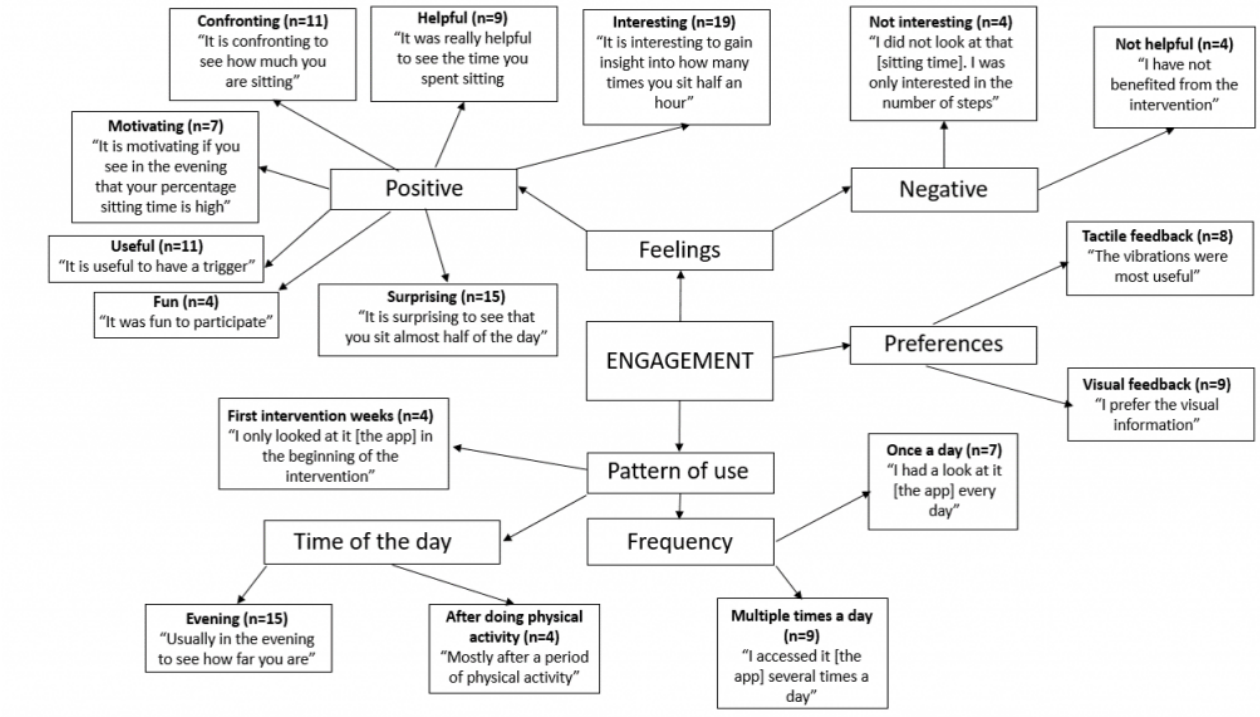
Table 1. Baseline characteristics of the participants.

Characteristic	Values (N=28)
Gender, n (%)	
Men	13 (46)
Women	15 (54)
Age in years	
All adults, mean (SD), range	64.3 (3.8), 60-76
Young older adults (<65 years), n (%)	15 (54)
Older adults (≥65 years), n (%)	13 (46)
Educational level, n (%)	
No education or primary education	1 (4)
Secondary education	11 (39)
College or university	16 (57)
Family situation, n (%)	
No partner (ie, single or widowed)	5 (18)
Have a partner but living separately	1 (4)
Married or living with a partner	22 (79)
BMI	
Mean (SD)	25.4 (3.9)
Healthy weight, n (%)	16 (57)
Overweight, n (%)	6 (21)
Obese, n (%)	5 (18)

User Engagement

Qualitative data on user engagement were thematically analyzed and are presented in Figure 3. The main themes that emerged were positive and negative feelings about the intervention, preferences for the kind of feedback, and the pattern of use.

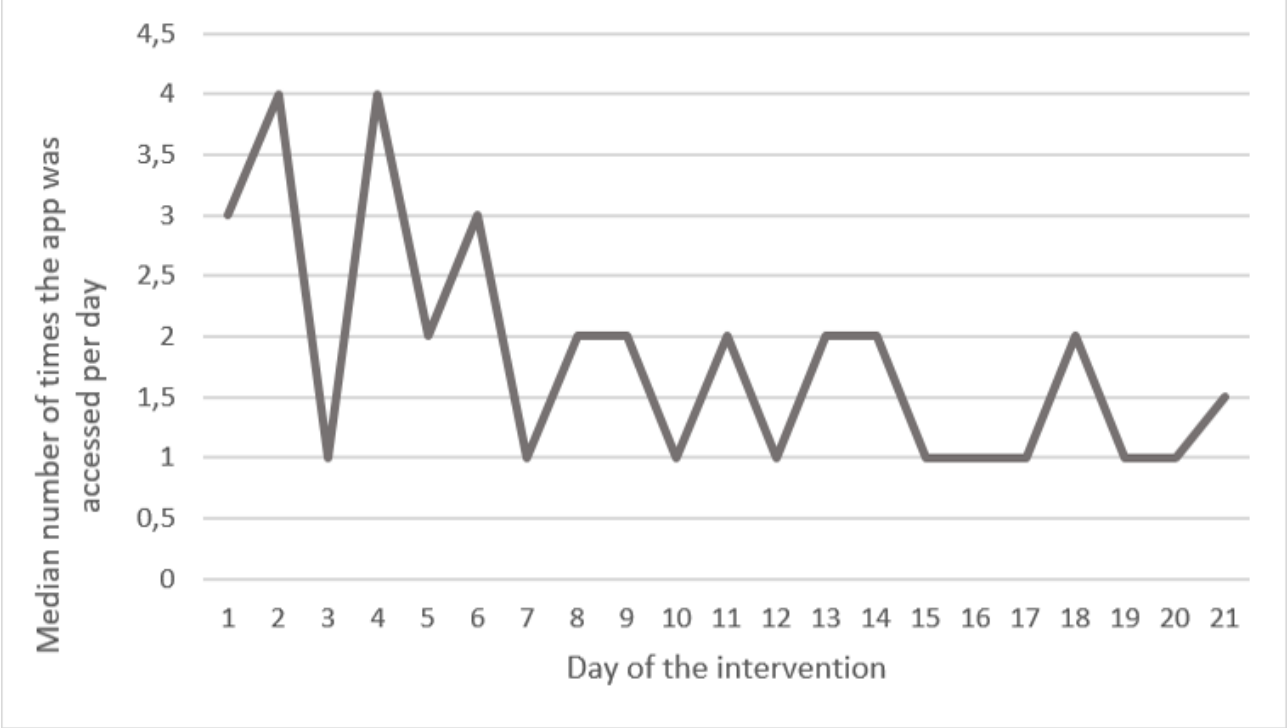
Figure 3. Pen profile of engagement with the intervention.



XSL•FO
RenderX

The participants mainly reported positive feelings, such as being motivated, surprised, and interested. Only a minority (3/28, 11%) indicated that they thought the intervention was not interesting and not helpful. There were mixed opinions on the preferred kind of feedback (ie, tactile vs visual). Some thought the vibrations were more useful, whereas others favored the visual information on the app. System usage data showed that the median number of days the self-monitoring device was worn by the participants was 20 out of 21 days (range 15-21). Half of the participants (16/28, 57%) reported that they accessed the app on a daily basis. This finding was confirmed by system usage data (see Multimedia Appendix 2), showing that 8 out of 28 participants (29%) consulted the app every day, while 5 participants (18%) consulted the app at least 80% of the days. Some participants reported that they consulted the visual feedback multiple times a day. Accordingly, system usage data showed that the median frequency of consulting the app ranged from 0 to 20 times a day (see Multimedia Appendix 2). Especially in the evening, and after doing physical activities, participants reported that they viewed the visual feedback. They indicated that the main reasons to access the app were out of curiosity, to go through their day, and to see the impact of certain physical activities. Participants also emphasized that they consulted the app more frequently in the beginning of the intervention period, compared to the end of the intervention period. This finding was also in line with the system usage data, which show that the median frequency of consulting the app ranged from 3 or 4 times a day in the beginning of the intervention to 1 or 2 times a day at the end of the intervention (see Figure 4).

Figure 4. Evolution of consulting visual feedback.

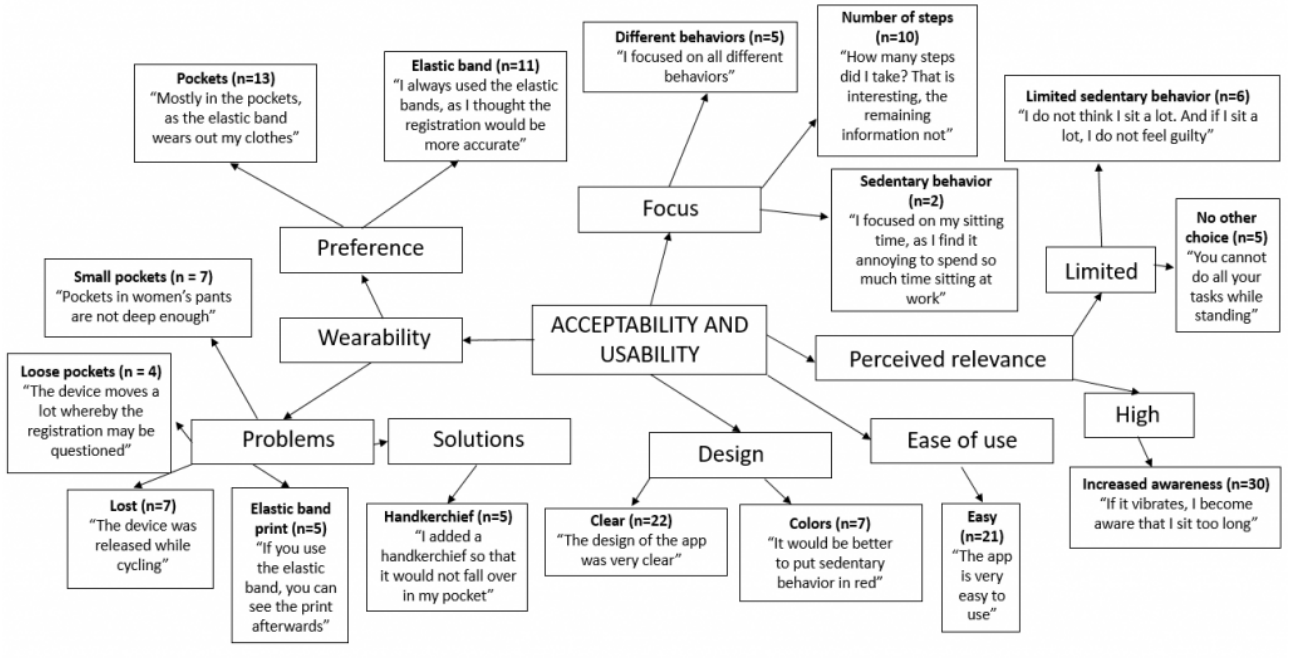


Acceptability and Usability

Results of the thematic analysis on acceptability and usability of the intervention are presented in Figure 5. The main themes that were identified were the design and the ease of use, wearing preferences, problems and solutions, the focus, and the perceived relevance. The intervention was considered easy to use, and most participants described the design as clear. The only remark on the design were the colors of the behaviors. Participants frequently cited that it would be more logical if sedentary behavior (ie, the behavior that should be limited) were displayed in red and the number of steps (ie, the behavior that should increase) in green. Participants expressed mixed preferences regarding the way to wear the device. Some participants (11/28, 39%) preferred to use the elastic band, whereas others (13/28, 46%) preferred to wear it in their pockets. Frequent problems that older adults, especially women, experienced included small or loose pockets, loss of the device, and the imprint of the elastic band on their clothes after wearing it. Out of the 28 older adults, 5 (18%) indicated that they used a handkerchief to ensure that the device was fixed in their pocket and could not flip over. Despite the fact that the aim of the intervention was to reduce sedentary behavior, only 2 participants (7%) mainly focused on the sedentary behavior information. Although the majority of older adults rated the intervention as highly relevant, some older adults were not convinced about the relevance. The most important reasons for limited perceived relevance were (1) the fact that they do not spend a lot of time sitting, or at least think they are not sitting a lot, and thus have no need to reduce their sedentary time and (2) the fact that they often see no other option but sitting to perform certain tasks.

XSL•FO
RenderX

Figure 5. Pen profile on acceptability and usability of the intervention.



Preliminary Efficacy

Participants commonly reported that the intervention changed their thinking (ie, they became more *aware* of their sedentary behavior) but not their actual sedentary behavior. The latter result was supported by quantitative data derived from the activity monitor (see Table 2). Sitting and standing time were very similar at pre- and postmeasurements. There was a small improvement in steps of around 400 per day. This improvement was not significant, probably because of the small sample size.

Table 2. Preliminary efficacy of the intervention.

Variable	Values (N=26)		
	Baseline	Postintervention	*P* value
Total sedentary time (h/day), mean (SD)	8.7 (1.9)	8.8 (2.0)	.91
Sit-to-stand transitions (times per day), mean (SD)	50.2 (7.2)	50.4 (10.0)	.92
Standing time (h/day), mean (SD)	4.7 (1.3)	4.7 (1.3)	.76
Stepping time (h/day), mean (SD)	2.0 (0.8)	2.2 (0.8)	.58
Number of steps	4786	5193	.50

Discussion

Principal Findings

This study provides novel and in-depth insights into the potential of a self-monitoring–based mHealth intervention in older adults to reduce sedentary behavior. Overall, our results indicated that the intervention was generally well perceived by older adults, but preliminary analyses showed no reduction in sedentary time after the 3-week intervention period.

Previous research has shown that building sustained user engagement over time is challenging in mHealth interventions [29]. Low user engagement results in limited exposure to the intervention and, in turn, small or no intervention effects [30]. Therefore, gaining insight into the user engagement of mHealth interventions is crucial. Both objective usage data and subjective experiences showed that older adults were highly engaged with this study's intervention. The participants generally expressed positive feelings, and the majority consulted the feedback frequently. They all agreed that the intervention made them more aware of their sedentary behavior, but the intervention did not result in a decrease in sedentary time.

The lack of a decrease in sedentary time is not entirely surprising given the following reasons. Firstly, the intervention period of 3 weeks was probably too short to actually change habitual behavior. Changing habits takes a long time [31] and, thus, it is likely that participants still need the cues to interrupt and/or reduce their sedentary behavior after the intervention has ended. Ending the cues might have resulted in relapse into their old and unhealthy habitual sedentary behavior [32]. Secondly, the intervention mainly targeted automatic processes underlying sedentary behavior by bringing the habitual behavior into conscious awareness. However, dual-process theories of motivation posit that both controlled and automatic processes regulate our sedentary behavior [33]. Thus, additional behavior change techniques (eg, goal setting, action planning, and coping planning) should be included in the intervention to affect the controlled processes and to actually achieve behavior change.

XSL•FO
RenderX

Given that participants often mentioned that they saw no other options to reduce their sedentary behavior, it might be worth including concrete examples on how to reduce sedentary behavior. Thirdly, physical activity information (ie, number of steps) was also provided in the app, notwithstanding that the only aim of this intervention was to reduce sedentary behavior. The physical activity information could not be removed from the Activator app before the start of the study. Existing literature has indicated that participants of interventions targeting both sedentary behavior and physical activity simultaneously are more likely to focus on increasing physical activity due to (1) the clearer guidelines for physical activity (ie, 150 minutes of moderate- to vigorous-intensity physical activity a week) compared to sedentary behavior (ie, sit less), (2) the better-known negative health consequences of too little physical activity compared to too much sedentary behavior, and (3) the fact that physical inactivity is often still considered a synonym for sedentary behavior [34]. The latter was confirmed by the results of the semistructured interviews: participants often mentioned that they wanted to increase the number of steps in order to reduce their sedentary time. Objective physical activity data showed that the average daily number of steps increased by approximately 400, or 10%, over the 3-week intervention period. Although this increase was not significant, this indicates that Activator feedback is more likely to affect the number of steps than the sedentary time. This finding is in line with the results of previous Activator studies [21,35] and suggests that more efforts should be made to clarify the difference between sedentary behavior and physical inactivity and to emphasize the importance of standing and light-intensity physical activity.

Despite the fact that common aging-related barriers (eg, visual impairment, reduced working memory, limited motivation, and reduced mobility) can influence the use of mHealth in older adults [36], general perceptions on the acceptability and usability of this study's intervention were positive. The app was easy to use and the design was clear. This is of great importance, as previous research has indicated that simplicity is one of the key principles for the design of mHealth interventions for older adults [37-39]. Although the Activator could be worn in different ways (ie, in the pants pockets or with an elastic band), the wearing of the device was often mentioned as challenging, especially among women. More research is therefore required to determine the ideal manner of attaching and wearing the Activator, especially when wearing pants without pockets or without deep pockets.

Strengths and Limitations

Strengths of this study include the innovativeness of the research. To our knowledge, this is the first study examining older adults' experiences with an electronic self-monitoring device specifically developed to reduce sedentary behavior. Moreover, by collecting both qualitative and quantitative data, a comprehensive view was obtained on self-monitoring–based mHealth interventions to reduce older adults' sedentary behavior. An important limitation of this study was the sampling method. Participants were not randomly selected and, therefore, selection bias may have occurred. The majority of the participants were highly educated, whereby generalization of the results to lower-educated groups might be limited. Moreover, no control group was included, as the main aim was to gain in-depth knowledge on participants' perceptions with the intervention. Although data saturation was achieved in the qualitative analysis, the small sample size was only sufficient to get a first indication on effect sizes and was not meant to provide sufficient statistical power for the quantitative analysis. Finally, the intervention lasted only 3 weeks and, thus, no conclusions can be drawn on the long-term adherence to the intervention. Based on these limitations, future studies should endeavor to recruit a larger, more generalizable sample and should use a randomized controlled trial design to draw firm conclusions on the effectiveness of a self-monitoring tool to reduce older adults' sedentary behavior. Furthermore, we believe that adding behavior change techniques to the mHealth intervention, ones that can affect the controlled processes underlying sedentary behavior, and extending the intervention duration might be recommended in future studies.

Conclusions

Results of this study suggest that the innovative self-monitoring–based mHealth intervention holds potential for the reduction of sedentary behavior in older adults. The intervention was considered interesting, helpful, and easy to use, and was able to increase awareness among older adults of their sedentary behavior. Despite the positive perceptions, no reductions in objective sedentary time were found in this study's sample. Hence, the intervention was probably of insufficient intensity to reduce the sedentary behaviors of participants. In order to effectively achieve behavior change, a number of modifications to the intervention are suggested, such as the addition of behavior change techniques that target controlled processes underlying sedentary behavior.

Acknowledgments

The authors would like to thank Charlotte Meersseman and Siel Mechelinck for their assistance with the qualitative data analysis, as well as the master students that supported the data collection. The first author (SC) is supported by funding from the Special Research Fund of Ghent University (BOF/PDO/2018/0006/01).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.

XSL•FO
RenderX

[DOCX File , 48 KB - mhealth_v8i10e18653_app1.docx]

Multimedia Appendix 2
Frequency of app consultation per participant.
[DOCX File , 54 KB - mhealth_v8i10e18653_app2.docx]

References

1. Decade of Healthy Ageing 2020–2030. Geneva, Switzerland: World Health Organization; 2019. URL: https://www.who.int/docs/default-source/decade-of-healthy-ageing/final-decade-proposal/decade-proposal-final-apr2020-en.pdf?sfvrsn=b4b75ebc_5 [accessed 2020-10-21]
2. de la Maisonneuve C, Oliveira Martins J. Public Spending on Health And Long-Term Care: A New Set of Projections. Paris, France: OECD; 2013 Jun. URL: oecd.org/economy/growth/Health%20FINAL.pdf
3. Maher JP, Sliwinski MJ, Conroy DE. Feasibility and preliminary efficacy of an intervention to reduce older adults' sedentary behavior. Transl Behav Med 2017 Mar;7(1):52-61 [FREE Full text] [doi: 10.1007/s13142-016-0394-8] [Medline: 27734281]
4. de Rezende LFM, Rey-López JP, Matsudo VKR, do Carmo Luiz O. Sedentary behavior and health outcomes among older adults: A systematic review. BMC Public Health 2014 Apr 09;14:333 [FREE Full text] [doi: 10.1186/1471-2458-14-333] [Medline: 24712381]
5. Copeland JL, Ashe MC, Biddle SJ, Brown WJ, Buman MP, Chastin S, et al. Sedentary time in older adults: A critical review of measurement, associations with health, and interventions. Br J Sports Med 2017 Nov;51(21):1539. [doi: 10.1136/bjsports-2016-097210] [Medline: 28724714]
6. Rosenberg DE, Bellettiere J, Gardiner PA, Villarreal VN, Crist K, Kerr J. Independent associations between sedentary behaviors and mental, cognitive, physical, and functional health among older adults in retirement communities. J Gerontol A Biol Sci Med Sci 2016 Jan;71(1):78-83 [FREE Full text] [doi: 10.1093/gerona/glv103] [Medline: 26273024]
7. Ekelund U, Steene-Johannessen J, Brown WJ, Fagerland MW, Owen N, Powell KE, et al. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. Lancet 2016 Sep;388(10051):1302-1310. [doi: 10.1016/S0140-6736(16)30370-1]
8. Harvey JA, Chastin SFM, Skelton DA. Prevalence of sedentary behavior in older adults: A systematic review. Int J Environ Res Public Health 2013 Dec 02;10(12):6645-6661 [FREE Full text] [doi: 10.3390/ijerph10126645] [Medline: 24317382]
9. Maher JP, Conroy DE. A dual-process model of older adults' sedentary behavior. Health Psychol 2016 Mar;35(3):262-272. [doi: 10.1037/hea0000300] [Medline: 26690644]
10. Gardner B, Smith L, Lorencatto F, Hamer M, Biddle SJH. How to reduce sitting time? A review of behaviour change strategies used in sedentary behaviour reduction interventions among adults. Health Psychol Rev 2016;10(1):89-112 [FREE Full text] [doi: 10.1080/17437199.2015.1082146] [Medline: 26315814]
11. Welk GJ. The youth physical activity promotion model: A conceptual bridge between theory and practice. Quest 1999 Feb;51(1):5-23. [doi: 10.1080/00336297.1999.10484297]
12. Verplanken B, Wood W. Interventions to break and create consumer habits. J Public Policy Mark 2018 Oct 02;25(1):90-103. [doi: 10.1509/jppm.25.1.90]
13. Aarts H, Dijksterhuis A. The silence of the library: Environment, situational norm, and social behavior. J Pers Soc Psychol 2003;84(1):18-28. [doi: 10.1037/0022-3514.84.1.18]
14. Hermsen S, Frost J, Renes RJ, Kerkhof P. Using feedback through digital technology to disrupt and change habitual behavior: A critical review of current literature. Comput Human Behav 2016 Apr;57:61-74. [doi: 10.1016/j.chb.2015.12.023]
15. Quinn JM, Pascoe A, Wood W, Neal DT. Can't control yourself? Monitor those bad habits. Pers Soc Psychol Bull 2010 Apr;36(4):499-511. [doi: 10.1177/0146167209360665] [Medline: 20363904]
16. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: The CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
17. Compernolle S, DeSmet A, Poppe L, Crombez G, De Bourdeaudhuij I, Cardon G, et al. Effectiveness of interventions using self-monitoring to reduce sedentary behavior in adults: A systematic review and meta-analysis. Int J Behav Nutr Phys Act 2019 Aug 13;16(1):63 [FREE Full text] [doi: 10.1186/s12966-019-0824-3] [Medline: 31409357]
18. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
19. Ayala GX, Elder JP. Qualitative methods to ensure acceptability of behavioral and social interventions to the target population. J Public Health Dent 2011;71 Suppl 1:S69-S79 [FREE Full text] [doi: 10.1111/j.1752-7325.2011.00241.x] [Medline: 21656958]
20. Gill JMR, Hawari NSA, Maxwell DJ, Louden D, Mourselas N, Bunn C, et al. Validation of a novel device to measure and provide feedback on sedentary behavior. Med Sci Sports Exerc 2018 Mar;50(3):525-532 [FREE Full text] [doi: 10.1249/MSS.0000000000001458] [Medline: 29040225]

XSL•FO
RenderX

21. Renaud LR, Jelsma JG, Huysmans MA, van Nassau F, Lakerveld J, Speklé EM, et al. Effectiveness of the multi-component dynamic work intervention to reduce sitting time in office workers: Results from a pragmatic cluster randomised controlled trial. Appl Ergon 2020 Apr;84:103027. [doi: 10.1016/j.apergo.2019.103027] [Medline: 31987512]
22. Morris AS, Murphy RC, Shepherd SO, Healy GN, Edwardson CL, Graves LEF. A multi-component intervention to sit less and move more in a contact centre setting: A feasibility study. BMC Public Health 2019 Mar 12;19(1):292 [FREE Full text] [doi: 10.1186/s12889-019-6615-6] [Medline: 30866872]
23. Winkler EAH, Bodicoat DH, Healy GN, Bakrania K, Yates T, Owen N, et al. Identifying adults' valid waking wear time by automated estimation in activPAL data collected with a 24 h wear protocol. Physiol Meas 2016 Oct;37(10):1653-1668. [doi: 10.1088/0967-3334/37/10/1653] [Medline: 27652827]
24. Aguilar-Farias N, Martino-Fuentealba P, Salom-Diaz N, Brown WJ. How many days are enough for measuring weekly activity behaviours with the ActivPAL in adults? J Sci Med Sport 2019 Jun;22(6):684-688. [doi: 10.1016/j.jsams.2018.12.004] [Medline: 30579814]
25. Poppe L, Van der Mispel C, Crombez G, De Bourdeaudhuij I, Schroé H, Verloigne M. How users experience and use an eHealth intervention based on self-regulation: Mixed methods study. J Med Internet Res 2018 Oct 01;20(10):e10412 [FREE Full text] [doi: 10.2196/10412] [Medline: 30274961]
26. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: An integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014 Jun 05;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
27. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
28. Boddy LM, Knowles ZR, Davies IG, Warburton GL, Mackintosh KA, Houghton L, et al. Using formative research to develop the healthy eating component of the CHANGE! school-based curriculum intervention. BMC Public Health 2012 Aug 29;12:710 [FREE Full text] [doi: 10.1186/1471-2458-12-710] [Medline: 22931457]
29. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
30. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: A systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
31. Neal DT, Wood W, Quinn JM. Habits—A repeat performance. Curr Dir Psychol Sci 2016 Jun 24;15(4):198-202. [doi: 10.1111/j.1467-8721.2006.00435.x]
32. Wood W, Neal DT. Behav Sci Policy 2016;2(1):71-83. [doi: 10.1353/bsp.2016.0008]
33. Maher JP, Conroy DE. A dual-process model of older adults' sedentary behavior. Health Psychol 2016 Mar;35(3):262-272. [doi: 10.1037/hea0000300] [Medline: 26690644]
34. van der Ploeg HP, Hillsdon M. Is sedentary behaviour just physical inactivity by another name? Int J Behav Nutr Phys Act 2017 Oct 23;14(1):142 [FREE Full text] [doi: 10.1186/s12966-017-0601-0] [Medline: 29058587]
35. Wyke S, Bunn C, Andersen E, Silva MN, van Nassau F, McSkimming P, et al. The effect of a programme to improve men's sedentary time and physical activity: The European Fans in Training (EuroFIT) randomised controlled trial. PLoS Med 2019 Feb;16(2):e1002736 [FREE Full text] [doi: 10.1371/journal.pmed.1002736] [Medline: 30721231]
36. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
37. Compernolle S, De Cocker K, Cardon G, De Bourdeaudhuij I, Van Dyck D. Older adults' perceptions of sedentary behavior: A systematic review and thematic synthesis of qualitative studies. Gerontologist 2019 Oct 31:1. [doi: 10.1093/geront/gnz127] [Medline: 31670766]
38. Gilbert BJ, Goodman E, Chadda A, Hatfield D, Forman DE, Panch T. The role of mobile health in elderly populations. Curr Geriatr Rep 2015 Aug 21;4(4):347-352. [doi: 10.1007/s13670-015-0145-6]
39. Farage MA, Miller KW, Ajayi F, Hutchins D. Design principles to accommodate older adults. Glob J Health Sci 2012 Feb 29;4(2):2-25 [FREE Full text] [doi: 10.5539/gjhs.v4n2p2] [Medline: 22980147]

Abbreviations

BIT: behavioral intervention technology
DBCI: digital behavior change intervention
mHealth: mobile health

XSL•FO
RenderX

Edited by G Eysenbach, L Buis; submitted 11.03.20; peer-reviewed by S Mukherjee, C Jacob; comments to author 12.06.20; revised version received 07.07.20; accepted 15.10.20; published 29.10.20.

Please cite as:
Compernolle S, Cardon G, van der Ploeg HP, Van Nassau F, De Bourdeaudhuij I, Jelsma JJ, Brondeel R, Van Dyck D
Engagement, Acceptability, Usability, and Preliminary Efficacy of a Self-Monitoring Mobile Health Intervention to Reduce Sedentary Behavior in Belgian Older Adults: Mixed Methods Study
JMIR Mhealth Uhealth 2020;8(10):e18653
URL: http://mhealth.jmir.org/2020/10/e18653/
doi:10.2196/18653
PMID:33118951

©Sofie Compernolle, Greet Cardon, Hidde P van der Ploeg, Femke Van Nassau, Ilse De Bourdeaudhuij, Judith J Jelsma, Ruben Brondeel, Delfien Van Dyck. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 29.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Costing and Cost-Effectiveness of a Mobile Health Intervention (ImTeCHO) in Improving Infant Mortality in Tribal Areas of Gujarat, India: Cluster Randomized Controlled Trial

Dhiren Modi[1], MD; Somen Saha[2], PhD; Prakash Vaghela[3], MD; Kapilkumar Dave[1], BHMS, PGDHHM; Ankit Anand, PhD; Shrey Desai[1], MD, MPH; Pankaj Shah[1], MD

[1]Society for Education, Welfare and Action (SEWA) Rural, Bharuch, India

[2]Indian Institute of Public Health Gandhinagar, Gandhinagar, India

[3]Health & Family Welfare Department, Gujarat, Gandhinagar, India

Corresponding Author:
Somen Saha, PhD
Indian Institute of Public Health Gandhinagar
NH-147, Palaj Village, opp. New Air Force Station HQ
Gandhinagar, 382042
India
Phone: 91 9687207549
Email: ssaha@iiphg.org

Abstract

Background: During 2013, a mobile health (mHealth) program, Innovative Mobile Technology for Community Health Operation (ImTeCHO), was launched in predominantly tribal and rural communities of Gujarat, India. ImTeCHO was developed as a job aid for Accredited Social Health Activists (ASHAs) and staff of primary health centers to increase coverage of maternal, neonatal, and child health care.

Objective: In this study, we assessed the incremental cost per life-years saved as a result of the ImTeCHO intervention as compared to routine maternal, neonatal, and child health care programs.

Methods: A two-arm, parallel, stratified cluster randomized trial with 11 clusters (primary health centers) randomly allocated to the intervention (280 ASHAs, n=2,34,134) and control (281 ASHAs, n=2,42,809) arms was initiated in 2015 in a predominantly tribal and rural community of Gujarat. A system of surveillance assessed all live births and infant deaths in the intervention and control areas. All costs, including those required during the start-up and implementation phases, were estimated from a program perspective. Incremental cost-effectiveness ratios were estimated by dividing the incremental cost of the intervention with the number of deaths averted to estimate the cost per infant death averted. This was further analyzed to estimate the cost per life-years saved for the purpose of comparability. Sensitivity analysis was undertaken to account for parameter uncertainties.

Results: Out of a total of 5754 live births (3014 in the intervention arm, 2740 in the control arm) reported in the study area, per protocol analysis showed that the implementation of ImTeCHO resulted in saving 11 infant deaths per 1000 live births in the study area at an annual incremental cost of US $163,841, which is equivalent to US $54,360 per 1000 live births. Overall, ImTeCHO is a cost-effective intervention from a program perspective at an incremental cost of US $74 per life-years saved or US $5057 per death averted. In a realistic environment with district scale-up, the program is expected to become even more cost-effective.

Conclusions: Overall, the findings of our study strongly suggest that the mHealth intervention as part of the ImTeCHO program is cost-effective and should be considered for replication elsewhere in India.

Trial Registration: Clinical Trials Registry of India CTRI/2015/06/005847; http://www.ctri.nic.in/Clinicaltrials/pdf_generate.php?trialid=11820&EncHid=&modid=&compid=%27,%2711820det%27

(JMIR Mhealth Uhealth 2020;8(10):e17066) doi:10.2196/17066

KEYWORDS

mHealth, cost-effectiveness; life-years saved; India, ASHA

XSL•FO
RenderX

Introduction

The National Health Mission of 2005 introduced a cadre of village-based frontline health worker communities named Accredited Social Health Activists (ASHAs) to facilitate the delivery of proven community-based maternal, neonatal, and child health (MNCH) services in rural areas of India, with one ASHA for every 1000 individuals. Despite this initiative, the coverage of selected MNCH services aimed at reducing mortality and undernutrition has remained low [1,2]. Coverage of key MNCH services and outcomes are inequitably distributed among the states of India [3]. The tribal communities have worse health indicators compared to those of nontribal communities [4]. The reasons for this low coverage are inadequate institutional capacity for optimal supervision and support to ASHAs in addition to insufficient skills, poor quality of training, and complexity of tasks to be performed [5,6].

Mobile telecommunication technology for health or mobile health (mHealth) has emerged as an important tool for global health. Its potential for improving the coverage and outcome of MNCH services through enhancing performance of frontline health workers has been well documented [7,8]. Despite its promise and potential, the Global Observatory for eHealth survey found that most initiatives on mHealth have not expanded beyond small-scale pilot projects [9]. In addition, a World Bank report on mobile apps for the health sector found that evidence of mHealth was limited, particularly for moving beyond intermediate outcomes to better health, especially in rural settings [10]. There is a gap in terms of rigor, intervention type, measurement of effectiveness, and, more importantly, cost-effectiveness of mHealth programs as well as detailed discussion on the scale-up of mHealth interventions.

An mHealth-based intervention named Innovative Mobile Technology for Community Health Operation (ImTeCHO) was designed for predominantly tribal and rural communities of Gujarat, India. ImTeCHO was developed as a job aid for ASHAs and staff of primary health centers to increase the coverage of MNCH care. A cluster randomized controlled trial was carried out to assess the effectiveness of the ImTeCHO intervention between 2015 to 2018, details of which have been published elsewhere [11]. In the present study, we assessed the incremental cost per life-year saved as a result of the ImTeCHO intervention as compared to routine MNCH programs.

Methods

Study Design

In 2015, a voluntary organization named Society for Education, Welfare and Action-Rural (SEWA-Rural) in active partnership with the government of Gujarat and the software partner Argusoft India Ltd began evaluating the effectiveness of the ImTeCHO intervention over a 3-year period through a cluster randomized controlled trial to improve the delivery of proven MHCH services through community-based ASHAs by enhancing their motivation and strengthening supervision in tribal areas of Gujarat, India. This trial was conducted within the existing public health system involving 22 primary health centers serving a total population of 4,76,943 individuals. Eleven primary health centers (280 ASHAs, n=2,34,134) were randomized to the intervention arm and 11 were randomized to the control arm (281 ASHAs, n=2,42,809). The control arm continued to receive usual health services from the government and other providers, while the intervention arm received the ImTeCHO intervention. There were four components of the ImTeCHO intervention:

1. Scheduling and task management: ASHAs received reminders on their mobile phone apps regarding the services to be provided every day.
2. Health promotion using multimedia: nine mobile-based short videos assisted ASHAs to provide counseling about key healthy behaviors during their home visits to households.
3. Decision support system screening, risk stratification, and treatment: the ImTeCHO app showed a diagnosis and customized treatment plan based on entry provided by the ASHAs on a mobile phone.
4. Support and supervision: the ImTeCHO web interface provided tools and real-time information to medical officers for offering timely support and supervision. ImTeCHO integrated a checklist (to ensure standardization of services) with other features offered through mobile technology, such as the ability to transfer data instantly and apply an algorithm automatically to data entered, along with features to ensure check-and-balance for truthfulness and accuracy of the collected information.

ASHAs from the control and intervention areas received a one-time refresher training on Home-based Newborn Care and Integrated Management of Neonatal and Childhood Illnesses. This was necessitated for ethical reasons as well as to ensure that ASHAs from both the intervention and control arms were updated about newer guidelines on the management of childhood illness to attribute effectiveness to the ImTeCHO intervention itself instead of only training.

Along with the intervention effectiveness, efforts were made to measure the incremental cost-effectiveness ratio of the ImTeCHO intervention. Estimates of intervention cost-effectiveness were calculated for different scenarios.

Framework for Assessing the Incremental Cost of Delivering the ImTeCHO Intervention

Cost data were collected from a program perspective. As the intervention was delivered at the doorstep and part of routine home visits of ASHAs, there were insignificant, if any, costs incurred by the households. Hence, costing from a program perspective was deemed appropriate. An incremental costing approach was adopted for the present study. The incremental costing approach takes into consideration the difference in cost and the additional benefits incurred on implementing the intervention compared with the routine case scenario (control arm). An incremental costing approach involves collecting additional financial costs, representing actual monetary flow on the goods and services purchased for delivering the program from a provider perspective. This type of costing enabled an analysis of cost by different program phases and program activities, and the unit cost of the program. Following the World Health Organization guide to cost-effectiveness analysis, all

XSL•FO
RenderX

research costs such as field testing and data collection were not included in the calculation of costs [12].

Two cost heads, start-up cost and annual implementation cost, were included in the 2016-2017 price. Cost data collection included a study of financial records of the ImTeCHO project. Start-up costs included those associated with activities conducted during the preimplementation period prior to February 2016, such as expenses for training, and a one-time capital cost that included software development. Capital costs, including start-up costs such as software development, were estimated by completing a checklist of all equipment (such as mobile phones) and furniture used in the program, and the useful life of the equipment. The orientation training cost was assumed to last for 3 years. Refresher training was assumed to be a recurrent activity. All capital costs, including start-up costs, were annualized assuming a useful life year. Initial software development costs (excluding annual costs on maintenance and upgrades) were assumed to last for 10 years (the cost incurred on development is expected to last until strategies for the program remain in place); four-wheelers were assumed to last for 10 years; two-wheelers were assumed to last for 7 years; and laptops, projectors, and printers were assumed to last for 5 years. Mobile handsets were assumed to last for 3 years on average. The cost of capital items was annualized across the project life, with discounting at an annual rate of 3%. The data sources included program financial reports, interviews with key officials, and surveillance data.

In addition to capital costs, other fixed costs included the costs of development of guidelines for implementation and their adherence, training costs for both staff and supervisors, communication material costs, supplies, and an additional incentive to ASHAs. The staff involved in the project were asked to estimate the amount of time spent on various activities at different times of the intervention.

The intervention implementation period was from February 2016 to January 2017. Software maintenance, internet data plan, additional incentives for ASHAs, and salaries for supervisory staff were important annual expenditures during the implementation period.

Start-up costs and annual implementation costs were summed to determine the total cost. All costs are presented in US dollars. Costs were converted to constant values and are reported as an annualized cost at the 2016-2017 price.

Surveillance to Collect Program Effectiveness Outcome Indicators: Infant Mortality

Infant deaths was the main outcome of interest to reflect program effectiveness for this study. To obtain complete information of all infant deaths, all pregnancy registrations, their outcomes, and survival status of all live births up to 1 year of age in all study clusters throughout the study period were counted as part of ongoing, prospective, pregnancy, and mortality surveillance. This was carried out by a data collection team who conducted a field survey of the entire study area every 3 months. One data collector covered approximately 25,000 individuals in the population. During their field surveys, the data collectors visited all localities of the study villages and met with ASHAs to register pregnant women who were native of the study village. All live births to women native to the study village were included, irrespective of place of the birth.

Once pregnancy was registered, the data collectors prospectively tracked pregnancy outcome and survival status up to 1 year after delivery by visiting the household at regular intervals. The data collectors recorded all maternal and infant deaths as part of the surveillance. Data entry was done in a customized, Android-based mobile phone app. The verbal autopsy method was used to review the cause of maternal and infant deaths [13]. When a data collector recorded a maternal or infant death, a supervisor visited the household of the deceased infant or mother, validated the death, and performed a verbal autopsy to determine the cause of death. A qualified, experienced doctor assigned the cause of death for each infant death after reviewing the verbal autopsy report. Supervisors also validated the surveillance data regarding the accuracy of pregnant registrations and pregnancy outcomes by calling random respondents over the phone and making random field visits. Deaths that were averted were calculated as the differences in deaths reported in the intervention and control arms from the surveillance activity.

Statistical Analysis to Estimate the Incremental Cost-Effectiveness Ratio of the Program

Cost-effectiveness ratios were estimated by dividing the incremental cost of the intervention with the number of infant deaths averted to estimate the cost per infant death averted. As per World Bank Data, a life expectancy of 68.35 years (2016 value) [14] was assumed to estimate the number of life years saved by averting an infant death, and this value as a denominator provided the cost per life years saved. According to the most commonly cited cost-effectiveness thresholds, an intervention is considered to be cost-effective if the incremental cost-effectiveness ratio (cost per life years saved) is less than the per capita gross domestic product (GDP) [15].

Both intention to treat (ITT) and per-protocol (PP) analyses were used in this study. ITT involved all live births, including women who were not native to the village but gave birth there. In other words, ITT included all live births that occurred in the study villages. However, there is a local custom in which some of the pregnant women leave their in-laws' homes during the last trimester and stay at their maternal home up to 2-3 months after delivery. These women and infants were partially exposed to the treatment arm. Therefore, PP analysis was performed by excluding deliveries of such women and infants to assess the effectiveness of the intervention only among those who were fully exposed.

A sensitivity analysis was conducted using the upper and lower estimates for various variables in the model to determine the impact of changes on cost per life years saved. A decision tree model was constructed to combine information from a wide variety of sources, extrapolate costs and health effects beyond the time period of the ImTeCHO study, and evaluate multiple potential interventions packaged into the strategies. Cost estimates that are relevant to the government staff in one representative district were estimated to assess the cost of scale up of such a program by the government. A sensitivity analysis for different program costs and effect estimates was conducted.

XSL•FO
RenderX

The study report adhered to the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) requirement for reporting the economic evaluation of health interventions [16]. The study protocol was approved by the SEWA-Rural Institutional Ethics Committee on January 29, 2016.

Results

Costs

The costs incurred in both the intervention and control arm were annualized based on the life span of the equipment. More than three-quarters (76%) of the cost was directed toward annual recurrent implementation costs, comprising personnel, training, software development, annual maintenance, supportive supervision, and monitoring costs. The item-specific annualized costs and the differences in costs for the intervention and control arms are detailed in Table 1.

Table 1. Annual start-up and implementation costs (US dollars) of the ImTeCHO[a] program in the implementation and control arms from a program perspective.

Cost Category	Annualized cost in intervention arm	Age (%)	Annualized cost in control arm	Cost difference
Annual start-up costs				
Total	45,647	26	17,789	27,858
Software development cost	7951	4.4	—[b]	7951
Vehicles	1135	0.6	—	1135
Mobile handset	11,873	6.5	—	11,873
Other IT[c] equipment	397	0.2	—	397
Training cost	24,291	13.4	17,789	6,502
Annual implementation costs				
Total	126,405	74	—	126,405
Personnel	24,919	14.5	—	24,919
Training cost	256	0.1	—	256
Software annual development and maintenance	49,599	28.8	—	49,599
Travel	2123	1.2	—	2123
ASHA[d] incentive	35,166	20.4	—	35,166
IT expense	12,935	7.5	—	12,935
Office expenses	1406	0.8	—	1406
Total costs	172,052	100	17,789	154,263

[a]IMTeCHO: Innovative Mobile Technology for Community Health Operation.

[b]—:not relevant; no costs incurred in the control arm.

[c]IT: information technology.

[d]ASHA: Accredited Social Health Activists.

Valuation of Study Outcome

Implementation of the ImTeCHO intervention with 561 ASHAs across 22 primary health centers of Gujarat resulted in 11 infant deaths per 1000 live births averted in the PP analysis (infant mortality rate of 56.4 per 1000 live births in the intervention area as compared to 67.2 per 1000 live births in the control area). This implies a reduction of 16% infant deaths per-protocol in the study area (Table 2). This resulted in an increase in 735 life years, with a life expectancy of 68.35 years. Although the protective effect of the ImTeCHO intervention was observed across other indicators such as early neonatal mortality rate, neonatal mortality rate, and stillbirths, to avoid double counting, these were not accounted for in the present analysis.

XSL•FO
RenderX

Table 2. Surveillance data on study outcome variables from February 2016 to January 2017.

Variable	Intention to Treat		Per Protocol	
	Control	Intervention	Control	Intervention
Number of live births	4059	4171	2740	3014
Hospital deliveries (%)	83.5	80.3	80.8	77.1
Number of early neonatal deaths	106	113	81	83
Number of neonatal deaths	138	142	102	104
Number of stillbirths	107	90	79	72
Number of infant deaths	236	233	184	170

Cost-Effectiveness

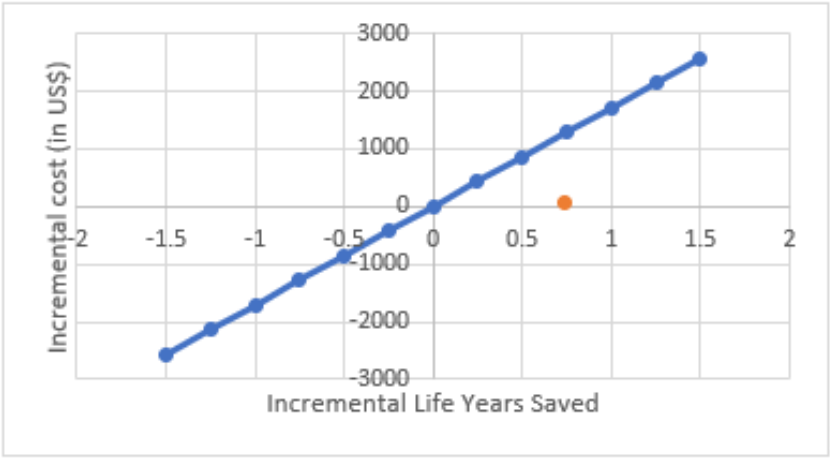
The implementation of ImTeCHO resulted in saving 11 infant deaths per 1000 live births in the study area at an annual incremental cost of US $54,360 per 1000 live births. Overall, ImTeCHO is a cost-effective intervention from a program perspective at an incremental cost of US $74 per life years saved or US $5057 per death averted (Table 3). At a per capita GDP of US $1709 in 2016, the ImTeCHO intervention is considered to be cost-effective (Figure 1).

Table 3. Cost-effectiveness of the ImTeCHO[a] program.

Point estimate of infant mortality rate	Value
Total births in the study area (n)	3014
Cost per live birth (US $)	54
Cost per 1000 live births (US $)	54,360
Infant deaths averted per 1000 live births (n)	11
Life years saved (life expectancy: 68.35 years)	735
Cost per infant deaths averted (US $)	5057
Cost per life years saved due to infant deaths averted (US $)	74

[a]ImTeCHO: Innovative Mobile Technology for Community Health Operation.

Figure 1. Cost-effectiveness plane with an incremental cost-effectiveness ratio, 2016-2017.



Sensitivity Analysis

To test the cost-effectiveness of the ImTeCHO program under different scenarios, we hypothesized three cases: (1) overall infant mortality rate reported in the study area, (2) estimated cost for district scale up with the actual ImTeCHO effect, and (3) estimated cost for district scale up with 50% of the reported ImTeCHO effect. The ImTeCHO program remained cost-effective across all scenarios (Table 4).

XSL•FO
RenderX

Table 4. Sensitivity analysis under different scenarios.

Variable	Value
Reference Case (IMR[a] per protocol in the study area)	
Infant deaths averted (n)	11
Cost per ASHA[b] (US $)	579
Cost per infant death averted (US $)	5057
Cost per LYS[c] (US $)	74
Scenario 1: IMR as intention-to-treat in the study area	
Infant deaths averted (n)	2
Cost per ASHA (US $)	578.95
Cost per infant death averted (US $)	17,225
Cost per LYS (US $)	252
Scenario 2: Wider district scale-up with 100% observed effectiveness	
Infant deaths averted (n)	11
Cost per ASHA (US $)	75
Cost per infant death averted (US $)	824
Cost per LYS (US $)	12
Scenario 3: Wider district scale-up with 50% observed effectiveness	
Infant deaths averted (n)	5
Cost per ASHA (US $)	75
Cost per infant death averted (US $)	1649
Cost per LYS (US $)	24

[a]IMR: infant mortality rate.

[b]ASHA: Accredited Social Health Activists.

[c]LYS: life-years saved.

Discussion

This economic evaluation study compared the costs and consequences of implementing an mHealth program (ImTeCHO) in the existing routine health services of a tribal block of Gujarat, India, compared to routine MNCH provided by frontline health workers (ASHAs). The findings are presented from a program perspective using an incremental cost approach. The per capita GDP of India at the 2016 price was US $1709 [17]. Overall, ImTeCHO is a cost-effective intervention from a program perspective at an incremental cost of US $74 per life years saved or US $5057 per death averted. Analysis under different scenarios showed that the program is expected to be even more cost-effective in a realistic environment with district scale up. The program is expected to be cost-effective even with a 50% reduction in effectiveness reported under the trial phase.

Our study used actual mortality data to report cost-effectiveness of an mHealth intervention. ReMiND, a nonrandomized study, implemented an mHealth app in 2012 through 259 ASHAs from two blocks of the Kaushambi district of the state of Uttar Pradesh in India, which resulted in a reduction of 0.2% maternal and 5.3% neonatal deaths. The incremental cost of the ReMiND program was US $205 per disability-adjusted life year averted or US $5865 per death averted [18]. However, the ReMiND study relied on modeling to estimate the number of deaths averted.

Implementation of an intervention as part of a routine government system has a distinct cost advantage for two reasons: first, the program could leverage the existing public infrastructure, and second, there is greater potential for the program to be replicated across the state of Gujarat, particularly in high-focused districts with a high infant mortality rate. The initial investment in software development, induction training, and supportive supervision were the key drivers for the success of the ImTeCHO program.

ImTeCHO, as a mobile phone app in the hands of health workers, has potential to bridge an important gap in the delivery of existing public health programs through ensuring data entry at the point of service delivery by the health provider through a handheld device, thereby ensuring better quality of data. This remains an area of future research once ImTeCHO is scaled up.

A limitation of our analysis is that we did not assess the health care input cost or time spent by health workers in training, supportive supervision by medical officers, and other supervisors from the health system. As a standard cost-effectiveness analysis, it was implicitly assumed that opportunity costs were equal and that it does not matter from which health care input

XSL•FO
RenderX

the resources were drawn [19]. The ImTeCHO trial was conducted in partnership with the local health government, and thus, the health care input cost was borne by the state to aid in potential scale up of the intervention. As the program operated in a realistic environment, it was assumed that the health care input cost will be absorbed as part of the health budget for replication of the program. The total cost of implementing the intervention was assumed to be the same in PP and ITT analyses; however, the total cost incurred to deliver the intervention among less than 5 live births included in the PP analysis would have been lower. There is potential for the cost-effectiveness to decrease over time as other interventions designed to directly or indirectly enhance maternal health and lower infant mortality take effect. However, this remains an area of future research.

Overall, the findings of our study strongly suggest that the mHealth intervention as part of the ImTeCHO program is cost-effective and should be considered for replication elsewhere in India and beyond.

Acknowledgments

This study was funded by the Indian Council of Medical Research, John D and Catherine T MacArthur Foundation, and the World Health Organization. We acknowledge support from Argusoft India Ltd in designing the ImTeCHO software and mobile app. We thank the project team from SEWA-Rural for their efforts toward implementation and evaluation. We thank the local community of the study area, ASHAs, primary health care staff, and district and state health officers for their support with the intervention.

Authors' Contributions

DM, SS, SD, and PS contributed to conceptualization, investigation, analysis, and writing of the manuscript. KD and AA contributed to analysis, review, and editing of the manuscript. PV facilitated implementation of the ImTeCHO intervention.

Conflicts of Interest

None declared.

References

1. Rai RK, Kumar C, Singh PK. District level coverage gap in Maternal, Newborn and Child Health care services in India. J Epidemiol Glob Health 2012 Dec;2(4):221-224 [FREE Full text] [doi: 10.1016/j.jegh.2012.12.004] [Medline: 23856504]
2. Balarajan Y, Selvaraj S, Subramanian S. Health care and equity in India. Lancet 2011 Feb 05;377(9764):505-515 [FREE Full text] [doi: 10.1016/S0140-6736(10)61894-6] [Medline: 21227492]
3. Kapur A. National Rural Health Mission (NRHM). SSRN Journal 2013 Mar 13;5(2):2231855 [FREE Full text] [doi: 10.2139/ssrn.2231855]
4. GoI. Report of the High Level Committee on Socio-Economic, Health and Educational Status of Tribal Communities in India. New Delhi: Ministry of Tribal Affairs, Government of India. 2014 May. URL: https://ruralindiaonline.org/library/resource/report-of-the-high-level-committee-on-socio-economic-health-and-educational-status-of-the-tribals-of-india/ [accessed 2020-08-27]
5. Bajpai N, Dholokia R. Improving the performance of accredited social health activists in India. In: Columbia Global Centers - South Asia Working Papers Series 1. Mumbai: Columbia Global Centers - Mumbai, Columbia University; 2011.
6. NHSRC. ASHA: which way forward...? Evaluation of ASHA programme. New Delhi: National Rural Health Mission; 2011. URL: https://nhm.gov.in/images/pdf/communitisation/asha/Studies/Evaluation_of_ASHA_Program_2010-11_Report.pdf [accessed 2020-08-27]
7. Lee S, Nurmatov U, Nwaru B, Mukherjee M, Grant L, Pagliari C. Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: Systematic review and meta-analysis. J Glob Health 2016 Jun;6(1):010401. [doi: 10.7189/jogh.06.010401] [Medline: 26649177]
8. Philbrick WC. mHealth and MNCH: State of the Evidence: Trends, Gaps, Stakeholder Needs, and Opportunities for Future Research on the Use of Mobile Technology to Improve Maternal, Newborn, and Child Health.: mHealth Alliance, UN Foundation; 2013 Jan. URL: https://www.gfmer.ch/mhealth/coursefiles2013/mhealthmnch-evidence-final.pdf [accessed 2020-08-26]
9. Kay M, Santos J, Takane M. mHealth: New horizons for health through mobile technologies. Geneva: World Health Organization; Jun 2011:66-71.
10. Qiang C, Yamamichi M, Hausman V, Altman D, Unit I. Mobile applications for the health sector. ICT Sector Unit World Bank. 2012 Apr. URL: http://documents1.worldbank.org/curated/en/751411468157784302/pdf/726040WP0Box370th0report00Apr020120.pdf [accessed 2020-08-26]
11. Modi D, Dholakia N, Gopalan R, Venkatraman S, Dave K, Shah S, et al. mHealth intervention "ImTeCHO" to improve delivery of maternal, neonatal, and child care services-A cluster-randomized trial in tribal areas of Gujarat, India. PLoS Med 2019 Oct 24;16(10):e1002939 [FREE Full text] [doi: 10.1371/journal.pmed.1002939] [Medline: 31647821]
12. Baltussen R, Adam T, Tan-Torres ET, Hutubessy R, Acharya A, Evans D, et al. In: Tan-Torres Edejer T, editor. Making choices in health: WHO guide to cost-effectiveness analysis. Geneva: World Health Organization; 2003.

XSL•FO
RenderX

13. Kanungo S, Tsuzuki A, Deen JL, Lopez AL, Rajendran K, Manna B, et al. Use of verbal autopsy to determine mortality patterns in an urban slum in Kolkata, India. Bull World Health Organ 2010 May 21;88(9):667-674. [doi: 10.2471/blt.09.073742]
14. Life expetancy at birth. The World Bank data. URL: https://data.worldbank.org/indicator/sp.dyn.le00.in [accessed 2020-08-26]
15. Sachs JD. Macroeconomics and health: investing in health for economic development. Rev Panam Salud Publica 2002 Aug;12(2):143-144. [doi: 10.1590/s1020-49892002000800017]
16. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, ISPOR Health Economic Evaluation Publication Guidelines-CHEERS Good Reporting Practices Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health 2013 Mar;16(2):231-250 [FREE Full text] [doi: 10.1016/j.jval.2013.02.002] [Medline: 23538175]
17. GDP Per Capita (current US$)- India. The World Bank. URL: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?end=2016&locations=IN&start=1960 [accessed 2020-08-26]
18. Prinja S, Bahuguna P, Gupta A, Nimesh R, Gupta M, Thakur JS. Cost effectiveness of mHealth intervention by community health workers for reducing maternal and newborn mortality in rural Uttar Pradesh, India. Cost Eff Resour Alloc 2018 Jun 25;16(1):25 [FREE Full text] [doi: 10.1186/s12962-018-0110-2] [Medline: 29983645]
19. van Baal P, Morton A, Severens JL. Health care input constraints and cost effectiveness analysis decision rules. Soc Sci Med 2018 Mar;200:59-64 [FREE Full text] [doi: 10.1016/j.socscimed.2018.01.026] [Medline: 29421472]

Abbreviations

ASHA: Accredited Social Health Activists
GDP: gross domestic product
ImTeCHO: Innovative Mobile Technology for Community Health Operation
ITT: intention to treat
mHealth: mobile health
MNCH: maternal, neonatal, and child health care
PP: per protocol
SEWA-Rural: Society for Education, Welfare and Action-Rural

Edited by G Eysenbach; submitted 15.11.19; peer-reviewed by M Dave, N Hussain-Shamsy; comments to author 31.12.19; revised version received 28.01.20; accepted 28.01.20; published 14.10.20.

Please cite as:
Modi D, Saha S, Vaghela P, Dave K, Anand A, Desai S, Shah P
Costing and Cost-Effectiveness of a Mobile Health Intervention (ImTeCHO) in Improving Infant Mortality in Tribal Areas of Gujarat, India: Cluster Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(10):e17066
URL: https://mhealth.jmir.org/2020/10/e17066
doi:10.2196/17066
PMID:33052122

©Dhiren Modi, Somen Saha, Prakash Vaghela, Kapilkumar Dave, Ankit Anand, Shrey Desai, Pankaj Shah. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 14.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Admissions to a Low-Resource Neonatal Unit in Malawi Using a Mobile App: Digital Perinatal Outcome Audit

Caroline Crehan[1], MBChB; Erin Kesler[1], MSN, CRNP; Indira Angela Chikomoni[2], BScNM; Kristi Sun[3], BSc, MBBS; Queen Dube[4], PhD; Monica Lakhanpaul[1,3], DM; Michelle Heys[1,5], MD

[1]UCL-Great Ormond Street Hospital Institute of Child Health, University College London, London, United Kingdom

[2]Zomba Central Hospital, Zomba, Malawi

[3]Whittington Hospital, London, United Kingdom

[4]Paediatric Department, College of Medicine, University of Malawi, Blantyre, Malawi

[5]East London NHS Foundation Trust, London, United Kingdom

Corresponding Author:
Michelle Heys, MD
UCL-Great Ormond Street Hospital Institute of Child Health
University College London
30 Guilford Street
Holborn
London, WC1N 1EH
United Kingdom
Phone: 44 (0)2079052212
Email: m.heys@ucl.ac.uk

Abstract

Background: Mobile health (mHealth) is showing increasing potential to address health outcomes in underresourced settings as smartphone coverage increases. The NeoTree is an mHealth app codeveloped in Malawi to improve the quality of newborn care at the point of admission to neonatal units. When collecting vital demographic and clinical data, this interactive platform provides clinical decision support and training for the end users (health care professionals [HCPs]), according to evidence-based national and international guidelines.

Objective: This study aims to examine 1 month's data collected using NeoTree in an outcome audit of babies admitted to a district-level neonatal nursery in Malawi and to demonstrate proof of concept of digital outcome audit data in this setting.

Methods: Using a phased approach over 1 month (November 21-December 19, 2016), frontline HCPs were trained and supported to use NeoTree to admit newborns. Discharge data were collected by the research team using a discharge form within NeoTree, called *NeoDischarge*. We conducted a descriptive analysis of the exported pseudoanonymized data and presented it to the newborn care department as a digital outcome audit.

Results: Of 191 total admissions, 134 (70.2%) admissions were completed using NeoTree, and 129 (67.5%) were exported and analyzed. Of 121 patients for whom outcome data were available, 102 (84.3%) were discharged alive. The overall case fatality rate was 93 per 1000 admitted babies. Prematurity with respiratory distress syndrome, birth asphyxia, and neonatal sepsis contributed to 25% (3/12), 58% (7/12), and 8% (1/12) of deaths, respectively. Data were more than 90% complete for all fields. Deaths may have been underreported because of phased implementation and some families of babies with imminent deaths self-discharging home. Detailed characterization of the data enabled departmental discussion of modifiable factors for quality improvement, for example, improved thermoregulation of infants.

Conclusions: This digital outcome audit demonstrates that data can be captured digitally at the bedside by HCPs in underresourced newborn facilities, and these data can contribute to a meaningful review of the quality of care, outcomes, and potential modifiable factors. Coverage may be improved during future implementation by streamlining the admission process to be solely via digital format. Our results present a new methodology for newborn audits in low-resource settings and are a proof of concept for a novel newborn data system in these settings.

(JMIR Mhealth Uhealth 2020;8(10):e16485) doi:10.2196/16485

XSL•FO
RenderX

KEYWORDS

infant, newborn; mHealth; data collection; clinical audit; digital health; low income population; mobile phone

Introduction

Background

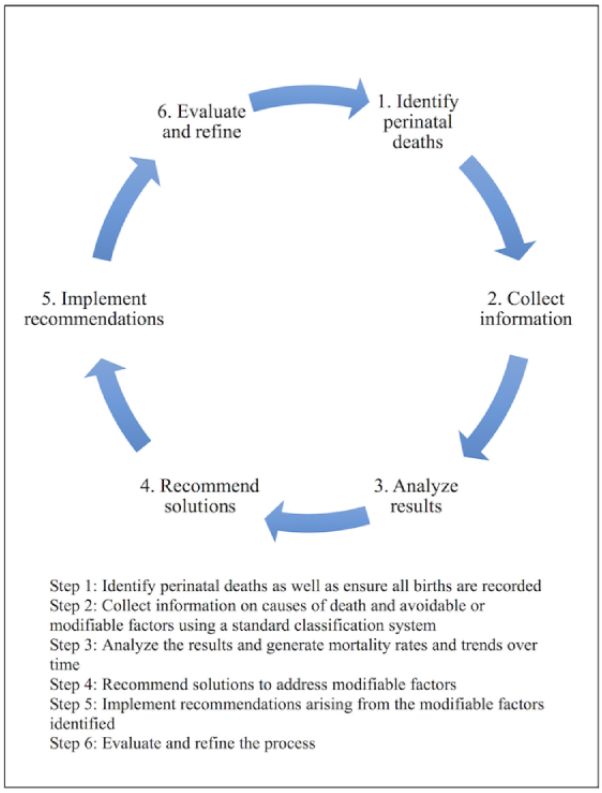
Each year, 2.5 million newborns die with no registration of their death or any documentation of how or why they died [1]. Half of the world's newborn babies do not receive a birth certificate [2,3], and stillbirths are statistically invisible [3]. Despite this, it is widely acknowledged that to prevent newborn deaths, more information on the number of births and deaths, causes of deaths, and avoidable factors linked to deaths is needed [4]. The process of a perinatal death audit aims to establish the profile of facility-based causes of death and has been shown to reduce perinatal mortality by 30% in low-resource countries [5]. Improving the quality of care for newborns is a key priority in tackling persistently high neonatal mortality rates (NMRs), particularly for hospitalized sick newborns in low- and middle-income countries [6]. At the facility level, the process of audit and feedback is considered the cornerstone of quality improvement, particularly when the process includes a clear action plan and targets [4]. Meanwhile, smartphone technology is becoming increasingly commonplace in low-resource settings. Mobile health (mHealth; "the use of mobile and wireless technologies to support the achievement of health objectives" [7]) has been harnessed for accurate and efficient data collection, particularly in community settings and in the context of clinical trials [8,9]. To the best of our knowledge, mHealth has not been utilized in the context of a hospital perinatal death audit. In response to recognized challenges in the scale-up of audits in these settings and a call for electronic health systems [5], we present the results of a novel digital outcome audit collected by health care professionals (HCPs) on an mHealth app: the NeoTree [10].

Different types of clinical audits exist for different purposes, for example, a structural audit examines the availability of resources in a system and a process audit assesses the process of case management [5]. An outcome audit assesses the end results of care, either deaths or near misses, depending on the NMR. In high-income countries where the NMR is low, the focus is more on near misses, and the management of specific cases are typically discussed in a monthly multidisciplinary *morbidity and mortalitymeeting*. In low-resource settings where the NMR is high, the emphasis is on deaths, and these are often discussed in a *death audit* meeting. Perinatal death audits have been defined as "the process of capturing information on the number and causes of all still births and neonatal deaths, or near misses where applicable, with an aim toward identifying specific cases for systematic, critical analysis of the quality of perinatal care received in order to improve the care provided to all mothers and babies" [4]. It aims to follow a 6-step cycle summarized in Figure 1 (adapted from Pattinson et al [5]).

Despite being one of the least developed countries in the world, Malawi has seen great success in achieving the Millennium Development Goal 4 [11], particularly for underfive mortality; however, for newborns, mortality is persistently high. A previous pediatric death audit in Kamuzu Central Hospital (KCH), a large referral center in Malawi, was recently published and cited reliable records as a significant limitation, with 6% of charts missing and documentation deficiencies in 58% of the charts reviewed [12]. To our knowledge, a perinatal death audit has not yet been published from Malawi or other sub-Saharan countries.

XSL•FO
RenderX

Figure 1. Six-step cycle for perinatal mortality audits.



Objectives

Our objective was to conduct and report a digital outcome audit of admission and discharge information collected by HCPs in a district-level facility using a novel mHealth app, NeoTree, on electronic mobile devices.

Methods

Setting

Zomba Central Hospital (ZCH) is a district-level hospital in Malawi with a large neonatal unit. During our study period and in discussions with the local clinical management team, it was noted by the researcher-in-residence (CC) that the 40-bed nursery takes around 7 admissions per day with 9 staff covering day and night shifts on 3 neonatal wards (high dependency, transit, and kangaroo wards). At maximum capacity, this equates to up to 3 nurses looking after approximately 50 babies per shift. Admissions are referred from a range of areas, defined as their *place of origin*, from within ZCH (eg, Theater), or from outside ZCH (eg, home, health center, or other hospital). A health center in this context is a facility that provides outpatient care services for common diseases in the local population, whereas a hospital is a larger facility providing more specialized care to a district population. Admissions are usually clerked on a Malawian Ministry of Health (MOH) paper proforma. Each patient's paper-based medical record consists of this admission form and any other loose-leaf sheets held together with a string. The records of NNDs are examined by pediatric and neonatal HCPs at monthly death audit meetings, often without an obstetric input. A rudimentary root cause analysis for each death was postulated, and subsequent recommendations were made. It was observed informally by the researcher-in-residence (CC) that the quality of the paper records is very poor [12], with faded, illegible writing, and the death audit process is time-consuming and often takes a whole day. She also observed that clinicians find it particularly difficult to attend death audits as their clinical duties continue. Oxygen concentrators and heaters in the unit rely on an electricity supply, which is affected by power outages on a daily basis. When this occurs, a backup supply of electricity is provided by a generator within a few minutes.

Identifying Deaths and Collecting Information Using a Digital App: Stages 1 and 2 of the Audit Cycle

The NeoTree App

A bedside app, NeoTree, was used (in addition to the MOH paper admission form) by frontline neonatal staff (mainly nurses) to record 1 month of admissions (November 21-December 19, 2016) to the neonatal unit on 3 low-cost Android electronic tablet devices, which were provided and installed at the nursing station. This audit took place in the context of an intervention development study [10], which presents the co-development process and acceptability, feasibility, and usability data.

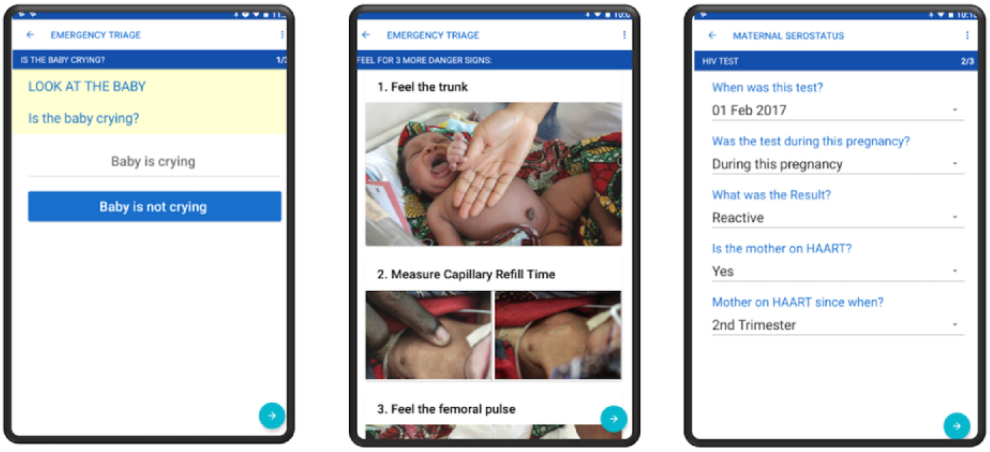
The NeoTree app was structured as an electronic admission form (Figure 2), which included all the fields of the standardized MOH paper neonatal admission form. Owing to HCP feedback, some fields were removed (eg, physical gestational score), but no new fields were added. By the end of the study period, the digital form followed the structure as below:

- Emergency triage and vital signs.
- Patient identification and demographics.
- Reason for admission (presenting history).

XSL•FO
RenderX

- Examination.
- Place of origin.
- Maternal history.
- Provisional HCP admission diagnoses.

Figure 2. Example app screens.



Digital Admission Form

Most fields on the electronic admission form were compulsory in that the HCP users could not continue until a plausible value had been recorded. However, in attempts to make the form more user-friendly and practical, some fields that relied on the availability of specific equipment (such as blood sugar and oxygen saturations) were made optional. Data collected at admission and discharge were stored locally on the tablets, and a printed copy (printed via a wireless printer on the ward) was kept in each patient's paper-based medical record, which included patient identifiers. When a network connection became available, data were exported pseudoanonymized with a unique ID to the researcher-in-residence's (CC) laptop as a file using Excel (Microsoft) program. Fields containing identifiable information were preconfigured as confidential by the researcher-in-residence so that they were not exported.

During each digital admission, a *reason for admission* was recorded. This was the presenting complaint entered by the nurse according to what the baby was referred to from the labor ward or other facility. If the referred patient arrived without accompanying information, the HCPs had to use their own clinical judgment. The reason for admission may have differed considerably from the actual diagnosis, where, for example, labor ward referrals were labeled *meconium aspiration* simply because there was meconium at delivery. The reason for admission was mutually exclusive, as there was only 1 presenting complaint recorded for each baby. Later, at the end of the digital admission form, *provisional HCP admission diagnoses* were decided by the HCPs based on their assessment of the baby. They could choose more than one diagnosis when necessary; hence, they were not mutually exclusive.

During the admission process, gestation was estimated from fundal height (recorded antenatally in the maternal record) or length of pregnancy (reported by the mother), both of which are unreliable methods [13,14]; hence, a physical maturity score was included in early iterations of NeoTree (Multimedia Appendix 1). Occasionally, wireless printing of forms was momentarily delayed by power outages, but these did not affect the completion of forms on the tablets because they had an 8-hour battery life.

Digital Discharge Form

The discharge form, the *NeoDischarge*, was completed by a researcher (EK) upon review of the patient's record just after discharge or death. Discharge information collected included identifiable information, outcome (discharged alive, absconded, or dead), HCP discharge diagnoses 1, 2, and 3 for discharges, and cause of death for NNDs. Often, there was more than one HCP discharge diagnosis documented; hence, the need for 3 fields, and these data were not mutually exclusive. It was noted that some patients in the sample had not had a documented review by a clinician during admission; hence, the researcher-in-residence (CC, a UK pediatrician in training with 4 years of neonatal experience, including 18 months in low income country newborn care), decided the most salient singular discharge diagnosis for all babies. This was labeled *researcher-in-residence discharge diagnosis*.

Recruitment and Initial Training

All 9 permanent staff on the neonatal rota were invited to attend a scenario-based one-on-one session in which they were trained to record new admissions using the app. Written consent to participate in the study was obtained according to ethical approval from The Malawi College of Medicine Research and Ethics Committee (reference 17PP12). Each HCP was then supervised clerking (recording) 1 new admission on the ward, completing both the MOH paper form and the form on the NeoTree app. When patients arrived with partially completed paper forms, HCPs were encouraged to cross-check information already documented, with mothers and the mother's medical record book, before entering it to the app. Permanent staff who could not attend the initial training session gave written consent when they entered the study (n=1). The audit and implementation were explained to other ad hoc locum staff or nursing students by the head nurse of each shift, who obtained their verbal consent.

XSL•FO
RenderX

Phased Implementation and Iterative Changes

NeoTree was then implemented on the ward in a phased approach over 1 month, with increasing coverage and decreasing levels of supervision. In the first week of the study, day-time admissions were supervised by the researcher-in-residence (CC) who was present in the ward. She was also available throughout the 1-month study to support any technical issues that occurred. Electronically clerked admissions were cross-checked by the researcher-in-residence with the admissions book and the ward manager at the end of each day to see if any patients had not been captured on NeoTree. Incomplete sessions were deleted. During the study, via a web-based editor platform, the researcher-in-residence was able to act on verbal feedback from the nurses as they used the app, and reconfigure NeoTree to best fit an efficient admission process in consultation with the wider NeoTree team. For example, the order of sections was adjusted to start with triage and examination and end with mothers' history (the opposite of the order of the MOH paper form), and pictures were added to explain exactly how to check danger signs.

Statistical Analysis: Stage 3 of the Audit Cycle

The primary outcome of this study was to report and describe deaths over a 1-month period using an electronic app, NeoTree. To this end, the admission and outcome datasets were exported from the app and merged by matching unique identification numbers. The SPSS program was then used to analyze the data and produce simple graphs [15]. Descriptive statistics included totals (n), percentages (%), mean and SD for normally distributed data, and median and IQR for skewed data and ranges. Charts included simple bar and pie charts for categorical data and histograms for continuous data (eg, a histogram for temperature). To measure the digital process, the number of missing pieces of data for each data point was reported as a total (n) and as a percentage (%). The average time taken to complete the app was also calculated. Case fatality rates (CFRs) for each diagnosis were calculated as the number of deaths per 1000 babies admitted with that diagnosis, for example, CFR for respiratory distress syndrome (RDS) was 148 deaths per 1000 cases of RDS. The overall CFR was calculated as the number of deaths per 1000 babies admitted.

Presentation of Results at an Audit Meeting: Stage 4 of the Audit Cycle

Simple graphs and statistics produced were presented by the researcher-in-residence (CC) to the department 2 weeks later in conjunction with that month's death audit meeting. The researcher facilitated the discussion regarding each graph sequentially; however, the discussion was led by the participants. Emerging patterns in admission and outcome data, CFRs, and modifiable factors contributing to morbidity and mortality were identified, and possible solutions were discussed. Individual cases were not discussed.

Results

Participants

A total of 31 HCPs from 4 different cadres clerked newborns using the NeoTree app, including all 9 permanent staff on the rota, 8 of whom attended the one-on-one training workshops at the beginning of the study. All the other 22 participants were nonpermanent locum staff or nursing students temporarily working on the unit (Table 1).

Participants attending the audit meeting included 3 female nurse midwife technicians, 1 male clinical officer (who was also the head of the department), and 1 male nursing officer.

Table 1. Participant characteristics.

Characteristics	Values (N=31)
Gender (female), n (%)	22 (73)
Age (years), mean (SD); range	30 (12.3); 19-65
Cadre, n (%)	
Nurse midwife technician	7 (23)
Nursing officer	6 (19)
Nurse	1 (3)
Student-nurse	17 (55)
Years of neonatal experience, mean (SD); range	1 (4.4); 0-8
Previously used a tablet device, n (%)	15 (50)
Regularly used a tablet device, n (%)	4 (13)
Received COIN[a] training, n (%)	5 (16)
Received HBB[b] training, n (%)	22 (71)

[a]COIN: care of the infant newborn [16].

[b]HBB: helping babies breathe.

XSL•FO
RenderX

Digital Outcome Audit Findings: Stage 3 of the Audit Cycle

During the study period, there were 191 admissions to the neonatal ward. A total of 70.2% (134/191) admissions were completed using NeoTree, and data from 130 babies were exported for analysis. Of these, 129 were analyzed because of 1 repeat entry. Hence, 67.5% (129/191) of all admissions were analyzed.

Table 2 describes the patient demographics and clinical data. The key highlights are described here. The mean birth weight was 2616 g, with just over one-third 37.6% (44/117) of babies born at low birth weight (LBW) and 45.0% (58/129) born prematurely at <37 weeks' gestation. In total, 25 maturity scores were carried out by HCPs, most of which estimated higher gestation than the fundal height or length of pregnancy methods (data not shown).

Figure 3 depicts the clinical reason for admission, recorded at the beginning of the app before the clinical assessment, the most common of which was fever 30.2% (39/129), followed by birth asphyxia 17.1% (22/129) and prematurity 14.0% (18/129). The provisional admission diagnoses made by HCPs (at the end of the app, after clinical assessment) are reported in Figure 4.

The most common discharge diagnoses recorded by the HCP and the researcher-in-residence were very similar (Table 2): sepsis followed by birth asphyxia and LBW. In total, 84.3% (102/121) babies were discharged alive, 5.8% (7/121) left the hospital against medical advice, and 9.9% (12/121) died.

Table 3 describes the clinical findings at the emergency triage and HIV status of the babies. Two-third of the babies did not cry at triage and hence underwent an airway, breathing, circulation, and disability examination with only a minority considered *unstable* following this assessment. Grunting was the most common danger sign in unstable babies. Over a third (49/129, 38.0%) of neonates admitted were hypothermic. A fifth of the babies were known to be exposed to HIV before or during birth.

The cause of death for the 12 newborn deaths, as recorded by the second researcher (EK) on the discharge form, are shown in Table 4. In all 12 patients, the cause of death was the same as the researcher-in-residence's discharge diagnosis. Prematurity with RDS and birth asphyxia were the leading causes of death.

XSL•FO
RenderX

Table 2. Patient demographics and main findings of neonatal admissions.

Characteristics	Findings, n (%)	Missing (n=129), n (%)	Complete (n=129), n (%)
Demographics (n=129)			
Gender, n (%)		0 (0)	129 (100)
Male	79 (61.2)		
Female	50 (38.7)		
Age (hours), n (%)		0 (0)	129 (100)
≤48	85 (65.9)		
>48	44 (34.1)		
Type of birth, n (%)		0 (0)	129 (100)
Singletons	124 (96.1)		
Twins	5 (3.8)		
Admitted from, n (%)		0 (0)	129 (100)
Within Zomba Central Hospital	84 (65.1)		
Outside health facility	45 (34.8)		
Weight and gestation			
Birth weight (g; n=117)		12 (9.3)	117 (90.7)
Mean (SD)	2616 (750)		
Median (IQR)	2800 (2000-3200)		
Range	600-4000		
<2500 g (LBW[a]), n (%),	44 (37.6)		
>2500 g (normal birth weight), n (%)	73 (62.4)		
Admission weight (g; n=129)		0 (0)	129 (100)
Mean (SD)	2638 (841)		
Median (IQR)	2750 (2030-2750)		
Range	630-5320		
<2500 g, n (%)	49 (38.0)		
>2500 g, n (%)	80 (62.0)		
Gestation (weeks) by fundal height and LMP[b] (n=129)		0 (0)	129 (100)
Mean (SD)	36 (4)		
Median (IQR)	37 (34-38)		
Range	24-42		
<37 weeks, n (%)	58 (45.0)		
≤30 weeks, n (%)	9 (7.0)		
Discharge diagnoses and outcome			
Researcher-in-residence diagnosis (mutually exclusive[c]; n=129), n (%)		0 (0)	129 (100)
Neonatal sepsis	44 (34.1)		
Birth asphyxia (mild or moderate or severe)	30 (23.3)		
LBW	13 (10.1)		
Prematurity with RDS[d]	6 (4.7)		
Pneumonia or bronchiolitis	6 (4.7)		
Prematurity only	13 (10.1)		
Congenital abnormality	6 (4.7)		

XSL•FO
RenderX

Characteristics	Findings, n (%)	Missing (n=129), n (%)	Complete (n=129), n (%)
Other	6 (4.7)		
Well baby	4 (3.1)		
HCP[e] discharge diagnoses (not mutually exclusive[f]; n=129), n (%)		0 (0)	129 (100)
Neonatal sepsis	58 (45.0)		
Birth asphyxia	28 (21.7)		
Prematurity with RDS	20 (15.5)		
Congenital anomaly	5 (3.9)		
Other	42 (32.6)		
Outcome (n=121), n (%)		8 (6.2)	121 (93.8)
Absconded	7 (5.8)		
Discharged alive	102 (84.3)		
Neonatal death	12 (9.9)		

[a]LBW: low birth weight.

[b]LMP: last menstrual period.

[c]Mutually exclusive: assumes one single diagnosis only.

[d]RDS: respiratory distress syndrome.

[e]HCP: health care professionals

[f]Not mutually exclusive: assumes on discharge some babies had more than one diagnosis, which may have contributed to their presentation.

Figure 3. Reasons for admission to neonatal ward.

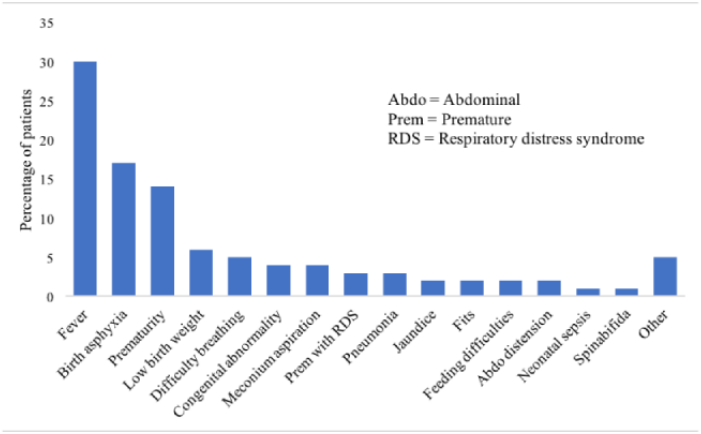
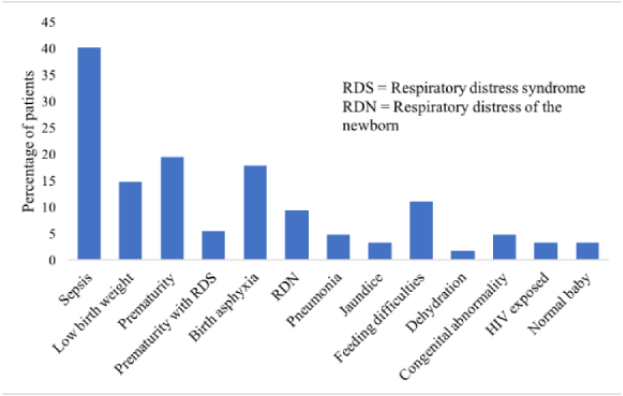


Figure 4. Provisional health care professional admission diagnoses (not mutually exclusive).



XSL•FO
RenderX

Table 3. Findings at emergency triage and HIV status of babies.

Characteristics	Findings, n (%)	Missing, n (%)	Complete, n (%)
Emergency triage, danger signs and vital signs			
Baby crying triage (n=129), n (%)		0 (0)	129 (100)
Crying	42 (32.6)		
Not crying	87 (67.4)		
ABCD[a] triage (n=87), n (%)		0 (0)	87 (100)
Stable	76 (87.4)		
Not stable	11 (12.6)		
Danger signs (n=129), n (%)		0 (0)	129 (100)
Grunting	18 (14.0)		
Cold trunk	7 (5.4)		
Prolonged capillary refill time	3 (2.3)		
Cyanosis	1 (0.8)		
Convulsions	1 (0.8)		
Weak femoral pulses	0 (0)		
None	101 (78.3)		
Vital signs, mean (SD)			
Admission HR[b] (bpm[c]; n=122)	141 (31)	7 (5.4)	122 (94.6)[d]
Admission oxygen saturations in air (%; n=122)	90 (12)	7 (5.4)	122 (94.6)[d]
Manual HR (bpm; n=7)	102 (20)	0 (0)	7 (100)
Admission RR[e] (bpm; n=129)	65 (18)	0 (0)	129 (100)
Admission oxygen saturations in oxygen (%; n=24)	86 (20)	2 (7.7)	24 (92.3)
Admission temperature (°C; n=129)	37.0 (1.6)	0 (0)	129 (100)
Admission blood sugar (mmol/l; n=124)	4.9 (4.1)	5 (3.9)	124 (96.1)[d]
Abnormal vital signs (n=129), n (%)		0 (0)	129 (100)
Tachycardic (HR>160 bpm)	21 (16.3)		
Hypoxic (oxygen saturations <90% in air)	26 (20.2)		
Tachypneic (RR>60 bpm)	66 (51.2)		
Hypothermic (temperature <36.5°C)	49 (38.0)		
Hyperthermic (temperature >37.5°C)	52 (40.3)		
Hypoglycemic (blood sugar <2.6 mmol/l)	9 (7.0)		
HIV status			
HIV status (n=129), n (%)		0 (0)	129 (100)
Exposed	25 (19.4)		
Unexposed	101 (78.3)		
Unknown	1 (0.8)		

[a]ABCD: airway, breathing, circulation, and disability.

[b]HR: heart rate.

[c]bpm: beats/breaths per minute for HR/RR, respectively.

[d]n=129.

[e]RR: respiratory rate.

XSL•FO
RenderX

Table 4. Cause of death and case fatality rates.

Cause of death[a]	Percentage deaths from total cases[b], n (%)	Case fatality rate per 1000 cases	Percentage deaths from total deaths (n=12), n (%)
Prematurity with RDS[c] (n=7)	3 (43)	428	3 (25)
Birth asphyxia (n=30)	7 (23)	233	7 (58)
Neonatal sepsis (n=44)	1 (2)	23	1 (8)
Congenital anomaly (n=6)	1 (17)	167	1 (8)
Total (N=129)	12 (9)	93	12 (100)

[a]Causes of death are mutually exclusive, that is, only 1 cause of death per neonate.

[b]Diagnosis refers to a researcher-in-residence's discharge diagnosis, which was the same as the cause of death for all neonatal deaths.

[c]RDS: respiratory distress syndrome.

Examination of the Newborn

A summary of findings on examination of the newborn is detailed in Multimedia Appendix 2. A total of 53 admissions (53/129, 41.1%) had signs of difficulty breathing or respiratory distress where chest in-drawings was the most commonly reported sign (36/129, 27.9%). On the basis of clinical judgement of the use of accessory muscles and prominence of chest retractions to aid breathing, 12 (12/129, 9.3%) were deemed to have severe work of breathing. Of HCPs, 56 (56/129, 43.4%) reported confidence in using a stethoscope; however, there was a lack of chest findings reported, and no heart murmurs were reported at all. Examination of the fontanelle was unavailable in the data because the confidential button had been pressed in error by the researcher-in-residence during the configuration of this field.

Place of Origin

In terms of place of birth, 93 (93/129, 72.1%) of the admissions were born in a hospital, 3 (3/129, 2.3%) were born at home, 27 (27/129, 20.9%) were born in a health center, and 4 (4/129, 3.1%) were born before arrival. Only 2 (2/129, 1.6%) were born with a traditional birth attendant despite a *ban* of traditional birth attendants by the government. Other facilities referred 40 patients (Multimedia Appendices 3 and 4).

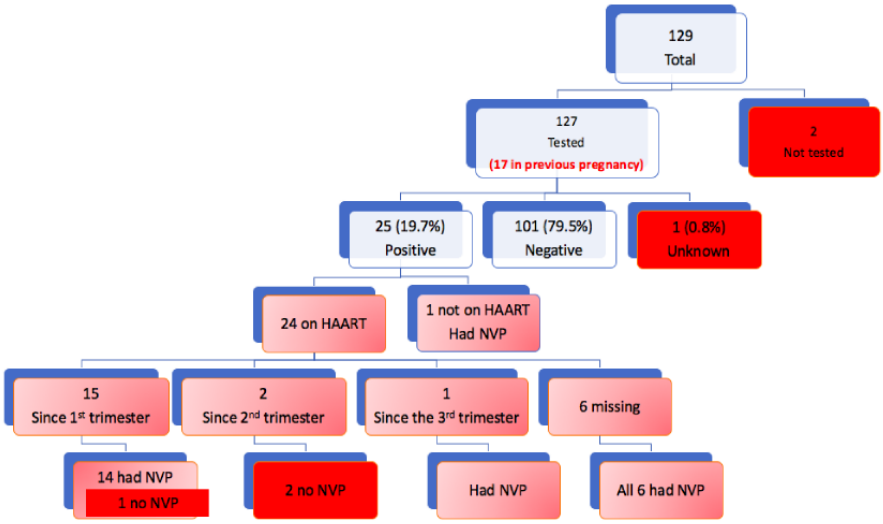
Maternal and Antenatal History

The mothers' ages were poorly recorded as 62 (62/129, 48.1%) mothers did not know their exact date of birth, and it was not exported for analysis because it is potentially identifiable information. A summary of the maternal history captured by the app is shown in Multimedia Appendix 5. Attendance at antenatal care was generally poor (Multimedia Appendix 6), with 40 (40/129, 31.0%) attending 2 or fewer antenatal appointments. A minimum of 3 antenatal appointments is needed to receive all doses of the tetanus vaccine. Most mothers (127/129, 98.4%) had been tested for HIV, and 25 (25/127, 19.7%) of those tested had a positive result (Figure 5). Of the 3 HIV-exposed babies who did not receive nevirapine prophylaxis, 2 had mothers who delayed highly active antiretroviral therapy until the second trimester; hence, they were the most vulnerable to vertical transmission of HIV.

Syphilis status was much more poorly recorded in comparison with HIV status, with *unknown* status in 21 mothers (21/129, 16.3%). A total of 35 mothers (35/129, 27.1%) had definitely not had a syphilis test. Of the 73 (73/129, 56.6%) mothers who had been tested, 3 were positive, and all of their babies were treated with penicillin. Medical conditions in pregnancy included malaria (16/129, 12.4%), hypertension (2/129, 1.6%), other sexually transmitted diseases (2/129, 1.6%), and anemia (2/129, 1.6%). No maternal heart disease, diabetes, or thyroid disease was reported in NeoTree. Antenatal steroids were administered to 8 mothers (8/129, 6.2%).

XSL•FO
RenderX

Figure 5. HIV status of mothers of babies admitted to neonatal unit using the NeoTree app. HAART: highly active antiretroviral therapy; NVP: nevirapine.



Labor History

A summary of the data captured for labor history is shown in Multimedia Appendices 7 and 8. Maternal conditions reported in the labor field of the NeoTree app included significant vaginal bleeding seen in 4 mothers, but no other problems were reported.

The Digital Process

Completeness of Data

Most fields in the app admission form were compulsory; hence, data for these fields were 100% complete with no missing data. For example, respiratory rate and temperature were recorded for all 129 babies in the sample (these were compulsory fields; Table 2). Only 5 admission fields were not compulsory and had less than 100% data completion rates: admission heart rate (HR; 122/129, 94.6%), admission oxygen saturation (122/129, 94.6%), saturations in oxygen (24/26, 92.3%), birth weight (117/129, 90.7%), and blood sugar (124/129, 96.1%). For the 7 missing admission HR and saturation readings (likely because of the lack of a pulse oximeter), the app directed the HCP to take and record a manual HR; hence, some form of HR (electronic or manual) was recorded for every baby. There were 8 missing outcomes at discharge because of a lack of clear outcome documentation in the patient record.

Time Taken

The mean time taken to complete an admission using the NeoTree app was 37 min (range 18-59 min), excluding 1 outlier (n=1). Anecdotal reports from HCPs suggest that longer sessions may have been interrupted by urgent tasks. Approximately one-fifth (22%) of NeoTree admissions were supervised.

Emerging Patterns and Corresponding Modifiable Factors Discussed: Stage 4 of the Audit Cycle

The audit meeting took just over 1 hour, after which HCPs could return to their daily duties; 4 factors were discussed. First, the high rate of hypothermia on admission (Table 3) and thermoregulation of babies by drying and wrapping was identified as a modifiable factor for improvement and future reauditing. Hypothermia is also an example of a factor that could be highlighted in the anticipated next phase of NeoTree development, that is, feedback data dashboards linked to the NeoTree data. Second, inadequate reporting of antenatal syphilis testing was discussed, and it was suggested that this could be fed back to health centers via the District Health Officer. Furthermore, the timing and completeness of penicillin treatment were requested to be added to the NeoTree form. Third, difficulties in knowing mothers' age were also highlighted as important because younger mothers typically experience more premature deliveries and more complications of pregnancy. A request was made for the app to calculate this in the future. Finally, because of the lack of reported problems in labor other than bleeding, it was deliberated that this question may have been answered poorly and that midwives and newborn HCPs should be shown where to find this information.

Discussion

Principal Findings

This paper presents a novel approach to capturing documentation in an inpatient setting that could signal a start to inpatient computerization. To our knowledge, this is the first digital outcome audit of neonatal admissions to a low-resource newborn facility. Our digital outcome audit achieved 70% coverage of admissions during a phased implementation approach where HCPs collected the data themselves on a novel app, NeoTree [10]. The overall CFR for newborns admitted on NeoTree was 92 per 1000. The most common diagnoses were sepsis, prematurity, and birth asphyxia for which the CFRs were 34, 250, and 250 per 1000, respectively. The completeness of data was high or 100% for much of the data set, exemplifying how the digital method significantly improves the quality of data in terms of completeness. In comparison, other studies have commented on how >50% of the charts had missing documentation [12]. Our 1-month audit has completed steps 1 to 4 of the audit cycle (Figure 1) [5] and has the potential for reaudit, evaluation, and refinement of recommendations and hence the completion of the whole audit cycle.

XSL•FO
RenderX

Discussion of Findings in Context

In a previous study at the ZCH nursery, demographic data were collected over a 2-month period using the MOH paper admission form. The mortality rate was 160 per 1000, significantly higher than that of our digital outcome audit [17]. Therefore, our digital outcome audit could have underestimated CFRs because of systematically missing data on babies who died. As HCPs were required to complete a paper form in addition to NeoTree, they may not have filled out a NeoTree admission form for babies that died soon after birth. If the NeoTree app completely replaced the paper option and HCPs were trained to clerk all babies, including those arriving moribund, this could cease to be a problem.

Another previous paper-based death audit in KCH hospital, Malawi, audited pediatric patients, with ages ranging from 1 day to 16.5 years, rather than newborns, and showed mortality rates ranging from 22 to 44 per 1000 [12]. The lower mortality rates most likely reflect the older age range, but may also be because of the retrospective nature of their study and missing data. The authors reported that >50% of the charts had missing documentation [12]. The prospective nature of our study, the presence of a researcher onsite overseeing data collection, and the use of a digital method may have aided in improving the completeness of data in our study.

Field validation and compulsory fields within the app may have also contributed to the completeness of data. Saturations in oxygen, for example, were not a compulsory field and were recorded in 24 of 26 babies. This may reflect the power of compulsory fields but also a lack of time to wait for a second saturation reading once oxygen had been applied. Indeed, the adult pulse oximeters available took time to pick up a reading, particularly in smaller premature infants. Local protocols (care of the infant newborn) [16] specify how to measure oxygen saturation and that it should be taken in air and oxygen as part of the assessment for starting continuous positive airway pressure, but these were newly implemented at the time of the study. For optional fields, lack of available information (eg, birth weight of older infants born in other facilities) and equipment (eg, test strips for the blood sugar monitor) may have contributed to incomplete data, and these fields were intentionally configured as optional for these reasons. For other fields (eg, syphilis status), an *unknown* option was added to allow progress through the app, where data were unavailable. A researcher collected the outcome data in our study by reading the documentation of HCPs (which was only 94% complete); hence, in the future, the recording of outcomes by the HCPs themselves in real time might improve the completeness of outcome data. Although the completeness of data was generally high, there is always the possibility of false data being entered. Since we did not include any quality assurance in our study, we can only assume that HCPs were entering correct data.

Discussion of Key Fields Within the NeoTree App

The percentage of LBW at admission (<2.5 kg) may be a useful indicator for the procurement of feeding cups and nasogastric tubes and the provision of kangaroo mother care beds. These data could also potentially influence the maternal and obstetric department and potentially government policies to tackle the nutrition of Malawian mothers and their babies.

It is important to note that LBW or small for gestational age is not the same as prematurity; hence, a maturity score was included within NeoTree. The difference in maturity scores and estimated gestation exposed the inaccuracy of fundal height and length of pregnancy, suggesting a significant underestimation of gestation using these methods. Feedback that the maturity score was time-consuming and required additional training prompted its removal from NeoTree halfway through the study.

Our results from the subjective assessment of the severity of work of breathing (WOB) suggest that nasal flare and chest in-drawings were not considered *severe* WOB. Head nodding, grunting, and tracheal tug made up for 9.3% of severe WOB. This could be further analyzed in the next phase to improve the understanding of the training needs of HCPs in assessing respiratory distress and potentially develop a scoring system in the future.

Regarding the examination of the newborn, the lack of chest findings reported and the complete lack of heart murmurs auscultated suggest that related fields may not be appropriate for nursing cadres, but their relevance for doctors could be examined in the future. The flexible nature of the NeoTree app means that these fields can be optional. For head circumference and birth weight, the app could ideally plot these automatically on a growth chart. However, it may represent a training challenge for HCPs to interpret these. Nevertheless, this is certainly a consideration for future iterations.

Limitations

The coverage of admissions was only 70.2% during a phased implementation; hence, this may not be a representative sample. The paper charts for the remaining 29.8% could have been checked to improve coverage of key indicators such as age and sex; however, this was not considered a reliable alternative because of missing charts and missing documentation [12]. As this was part of a proof-of-concept study, the digital form had to be completed in addition to the paper form, adding time and workload to already pressured staff. The researcher-in-residence was present throughout the study, which may have enhanced uptake. There were difficulties completing antenatal fields, particularly when a guardian accompanied the infant to the nursery while the mother was still recovering in the labor ward. As the study progressed, midwives started to bring the mother's labor ward notes, in addition to her hand-held record, with new admissions from the labor ward, but the problem persisted for out-born babies. Hence, the option *unknown* was added to many of the drop-down menus to preserve practical feasibility.

A major problem with our digital outcome audit is that the proposed system only collected data at the point of admission and discharge. What occurred during the crucial period between admission and death was not recorded, and therefore modifiable factors contributing to deaths and reciprocal solutions could not be identified. In turn, because of time constraints, steps 5 and 6 of the audit cycle could not be executed, leaving the audit loop unclosed. However, we have identified a considerable number of modifiable factors from patterns in aggregate data;

XSL•FO
RenderX

hence, with the right resources and staff available, we could potentially close the audit loop.

Future Steps

To allow the scrutiny of individual causes of death, a free text field will be added to the *NeoDischarge* form for the reviewing clinician to record (in a nonblame anonymous fashion) any possible modifiable factors that might have prevented that death. Copies of these *death summaries* could potentially be printed and collated for review in monthly death audits, which would significantly increase the efficiency of these meetings and provide valuable contemporary insights into how and why an individual newborn died. Other next steps include using NeoTree where it completely replaces the paper form, or where no paper form exists in the first place, without the presence of the researcher on site. A study where clinicians or doctors use the app in another low-resource country would also be recommended.

Conclusions

Using an mHealth app, NeoTree, a digital outcome audit was successfully carried out by health care workers at a neonatal unit of a district hospital in Malawi with high completeness of data. These results were discussed at a local audit meeting and demonstrated that data collected digitally could stimulate quality improvement initiatives, such as improving the thermoregulation of babies. Limitations are noted in this study, with only 70% coverage of all admissions. Overall, this study illustrates how a digital audit using an app can improve documentation and richness of clinical data to help support the delivery and configuration of local services. This study demonstrates huge potential for the use of a daily electronic record in low-resource settings, and these findings can inform the next stage of development for the NeoTree app, in particular, for guiding the development of linked data dashboards.

Acknowledgments

This work was supported by the National Institute for Health Research (NIHR) Great Ormond Street Hospital Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the National Health Service (NHS), the NIHR, or the Department of Health. The authors would like to thank the babies, mothers, and guardians, who allowed their data to contribute to this digital outcome audit. The authors are extremely grateful to the nursing team at ZCH who diligently contributed to this digital outcome audit with open minds and hearts. They extend special thanks to Dr Joshua, the director of ZCH, and Indira Angela Chikomoni, the nurse in charge during the time of the study, whose support and coordination of the project was invaluable. The authors thank Dr Rizine Mzikamanda, who kindly facilitated ethical approval from the Malawi University College of Medicine, and thank all the members of the technical team—Matteo Giacconne (main developer), Jane Minto (UX design), and Charles Normand (information technology architect), whose hard work created the app needed to deliver this audit. The authors would also like to thank Mr Vincent Harding at the Computer Sciences Department of the University College London, Professor Jens Kaasball at the University of Oslo, and the computer science team at Chancellor College, Zomba, University of Malawi, for their valuable technical advice. Finally, the authors thank the National Institute for Health Research Collaboration for the Leadership in Applied Health Research and Care North Thames at Barts Health NHS Trust for supporting Professor Monica Lakhanpaul.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Maturity scores compared to fundus height and length of pregnancy measured.
[PDF File (Adobe PDF File), 216 KB - mhealth_v8i10e16485_app1.pdf]

Multimedia Appendix 2
Findings on newborn examination.
[PDF File (Adobe PDF File), 48 KB - mhealth_v8i10e16485_app2.pdf]

Multimedia Appendix 3
Place of origin.
[PDF File (Adobe PDF File), 23 KB - mhealth_v8i10e16485_app3.pdf]

Multimedia Appendix 4
Map of referral centres showing number of referrals from outlying referral centres.
[PNG File , 292 KB - mhealth_v8i10e16485_app4.png]

Multimedia Appendix 5
Maternal history.
[PDF File (Adobe PDF File), 24 KB - mhealth_v8i10e16485_app5.pdf]

XSL•FO
RenderX

Multimedia Appendix 6
Summary of antenatal care visits and medications/vaccinations received.
[PNG File , 74 KB - mhealth_v8i10e16485_app6.png]

Multimedia Appendix 7
Summary of characteristics of labour as captured by the NeoTree app.
[PDF File (Adobe PDF File), 30 KB - mhealth_v8i10e16485_app7.pdf]

Multimedia Appendix 8
Reported reasoning for either elective or emergency C-section.
[PDF File (Adobe PDF File), 204 KB - mhealth_v8i10e16485_app8.pdf]

References

1. Neonatal Mortality. UNICEF DATA - Child Statistics. 2018. URL: https://data.unicef.org/topic/child-survival/neonatal-mortality/ [accessed 2019-04-22] [WebCite Cache ID 77ppaFTWl]
2. Lawn JE, Blencowe H, Oza S, You D, Lee AC, Waiswa P, Lancet Every Newborn Study Group. Every newborn: progress, priorities, and potential beyond survival. Lancet 2014 Jul 12;384(9938):189-205. [doi: 10.1016/S0140-6736(14)60496-7] [Medline: 24853593]
3. Lawn JE, Cousens S, Zupan J, Lancet Neonatal Survival Steering Team. 4 million neonatal deaths: when? Where? Why? Lancet 2005;365(9462):891-900. [doi: 10.1016/S0140-6736(05)71048-5] [Medline: 15752534]
4. Kerber KJ, Mathai M, Lewis G, Flenady V, Erwich JJ, Segun T, et al. Counting every stillbirth and neonatal death through mortality audit to improve quality of care for every pregnant woman and her baby. BMC Pregnancy Childbirth 2015;15(Suppl 2):S9 [FREE Full text] [doi: 10.1186/1471-2393-15-S2-S9] [Medline: 26391558]
5. Pattinson R, Kerber K, Waiswa P, Day LT, Mussell F, Asiruddin SK, et al. Perinatal mortality audit: counting, accountability, and overcoming challenges in scaling up in low- and middle-income countries. Int J Gynaecol Obstet 2009 Oct;107(Suppl 1):S113-21, S121. [doi: 10.1016/j.ijgo.2009.07.011] [Medline: 19815206]
6. Zaka N, Alexander EC, Manikam L, Norman IC, Akhbari M, Moxon S, et al. Quality improvement initiatives for hospitalised small and sick newborns in low- and middle-income countries: a systematic review. Implement Sci 2018 Jan 25;13(1):20 [FREE Full text] [doi: 10.1186/s13012-018-0712-2] [Medline: 29370845]
7. Mhealth New Horizons for Health Through Mobile Technologies. World Health Organization. 2011. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2019-04-22] [WebCite Cache ID 77pIuSWkm]
8. Dillon DG, Pirie F, Rice S, Pomilla C, Sandhu MS, Motala AA, African Partnership for Chronic Disease Research (APCDR). Open-source electronic data capture system offered increased accuracy and cost-effectiveness compared with paper methods in Africa. J Clin Epidemiol 2014 Dec;67(12):1358-1363 [FREE Full text] [doi: 10.1016/j.jclinepi.2014.06.012] [Medline: 25135245]
9. le Jeannic A, Quelen C, Alberti C, Durand-Zaleski I, CompaRec Investigators. Comparison of two data collection processes in clinical studies: electronic and paper case report forms. BMC Med Res Methodol 2014 Jan 17;14:7 [FREE Full text] [doi: 10.1186/1471-2288-14-7] [Medline: 24438227]
10. Crehan C, Kesler E, Nambiar B, Dube Q, Lufesi N, Giaccone M, et al. The NeoTree application: developing an integrated mHealth solution to improve quality of newborn care and survival in a district hospital in Malawi. BMJ Glob Health 2019;4(1):e000860 [FREE Full text] [doi: 10.1136/bmjgh-2018-000860] [Medline: 30713745]
11. Kanyuka M, Ndawala J, Mleme T, Chisesa L, Makwemba M, Amouzou A, et al. Malawi and millennium development goal 4: a countdown to 2015 country case study. Lancet Glob Health 2016 Mar;4(3):e201-e214 [FREE Full text] [doi: 10.1016/S2214-109X(15)00294-6] [Medline: 26805586]
12. Fitzgerald E, Mlotha-Mitole R, Ciccone EJ, Tilly AE, Montijo JM, Lang H, et al. A pediatric death audit in a large referral hospital in Malawi. BMC Pediatr 2018 Feb 21;18(1):75 [FREE Full text] [doi: 10.1186/s12887-018-1051-9] [Medline: 29466967]
13. Jehan I, Zaidi S, Rizvi S, Mobeen N, McClure EM, Munoz B, et al. Dating gestational age by last menstrual period, symphysis-fundal height, and ultrasound in urban Pakistan. Int J Gynaecol Obstet 2010 Oct;110(3):231-234 [FREE Full text] [doi: 10.1016/j.ijgo.2010.03.030] [Medline: 20537328]
14. White LJ, Lee SJ, Stepniewska K, Simpson JA, Dwell SL, Arunjerdja R, et al. Estimation of gestational age from fundal height: a solution for resource-poor settings. J R Soc Interface 2012 Mar 7;9(68):503-510 [FREE Full text] [doi: 10.1098/rsif.2011.0376] [Medline: 21849388]
15. IBM SPSS Statistics for Windows, Version 20. IBM Corp. 2011. URL: https://www.ibm.com/support/pages/how-cite-ibm-spss-statistics-or-earlier-versions-spss [accessed 2020-09-09]
16. Care of the Infant and Newborn in Malawi: the Coin Course - Participants Manual. St Andrews Research Repository. URL: https://research-repository.st-andrews.ac.uk/bitstream/handle/10023/12839/COIN.2017_participants_manual.pdf [accessed 2019-04-22]

XSL•FO
RenderX

17. Crehan C, Colbourn T, Heys M, Molyneux E. Evaluation of 'TRY': an algorithm for neonatal continuous positive airways pressure in low-income settings. Arch Dis Child 2018 Aug;103(8):732-738. [doi: 10.1136/archdischild-2017-313867] [Medline: 29514813]

Abbreviations

CFR: case fatality rate
HCP: health care professional
HR: heart rate
KCH: Kamuzu Central Hospital
LBW: low birth weight
mHealth: mobile health
MOH: Malawian Ministry of Health
NHS: National Health Service
NIHR: National Institute for Health Research
NMR: neonatal mortality rate
NND: neonatal death
RDS: respiratory distress syndrome
WOB: work of breathing
ZCH: Zomba Central Hospital

Edited by G Eysenbach; submitted 03.10.19; peer-reviewed by C King, N Muinga, R Dewey; comments to author 12.02.20; revised version received 15.06.20; accepted 21.06.20; published 21.10.20.

Please cite as:
Crehan C, Kesler E, Chikomoni IA, Sun K, Dube Q, Lakhanpaul M, Heys M
Admissions to a Low-Resource Neonatal Unit in Malawi Using a Mobile App: Digital Perinatal Outcome Audit
JMIR Mhealth Uhealth 2020;8(10):e16485
URL: https://mhealth.jmir.org/2020/10/e16485
doi:10.2196/16485
PMID:33084581

©Caroline Crehan, Erin Kesler, Indira Angela Chikomoni, Kristi Sun, Queen Dube, Monica Lakhanpaul, Michelle Heys. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 21.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effect of a Mobile App for the Pharmacotherapeutic Follow-Up of Patients With Cancer on Their Health Outcomes: Quasi-Experimental Study

Roberto Collado-Borrell[1], PharmD, PhD; Vicente Escudero-Vilaplana[1], PharmD, PhD; Almudena Ribed[1], PharmD, PhD; Cristina Gonzalez-Anleo[2], PharmD, PhD; Maite Martin-Conde[2], PharmD, PhD; Rosa Romero-Jimenez[1], PharmD, PhD; Irene Iglesias-Peinado[3], PharmD, PhD; Ana Herranz-Alonso[1], PharmD, PhD; Maria Sanjurjo-Saez[1], PharmD, PhD

[1]Hospital General Universitario Gregorio Marañon, Madrid, Spain

[2]Hospital Clinic Barcelona, Barcelona, Spain

[3]Faculty of Pharmacy, Universidad Complutense de Madrid, Madrid, Spain

Corresponding Author:
Vicente Escudero-Vilaplana, PharmD, PhD
Hospital General Universitario Gregorio Marañon
Dr. Esquerdo 46
Madrid
Spain
Phone: 34 915867714
Email: vicente.escudero@salud.madrid.org

Abstract

Background: Oral antineoplastic agents (OAAs) have revolutionized cancer management. However, they have been reported with adverse side effects and drug-drug interactions. Moreover, patient adherence to OAA treatment is critical. Mobile apps can enable remote and real-time pharmacotherapeutic monitoring of patients, while also promoting patient autonomy in their health care.

Objective: The primary objective was to analyze the effect of using a mobile app for the follow-up of patients with oncohematological malignancies undergoing treatment with OAAs on their health outcomes. The secondary objectives were to analyze the role of the app in communication with health care professionals and patient satisfaction with the app.

Methods: We performed a comparative, quasi-experimental study based on a prepost intervention with 101 patients (control group, n=51, traditional pharmacotherapeutic follow-up vs intervention group, n=50, follow-up through e-OncoSalud, a custom-designed app that promotes follow-up at home and the safety of patients receiving OAAs). The effect of this app on drug safety, adherence to treatment, and quality of life was evaluated.

Results: With regard to drug safety, 73% (37/51) of the patients in the control group and 70% (35/50) of the patients in the intervention group (P=.01) presented with drug-related problems. The probability of detecting an insufficiently treated health problem in the intervention group was significantly higher than that in the control group (P=.04). The proportion of patients who presented with side effects in the intervention group was significantly lower than that in the control group (P>.99). In the control group, 49% (25/51) of the patients consumed some health resources during the first 6 months of treatment compared with 36% (18/50) of the patients in the intervention group (P=.76). Adherence to treatment was 97.6% (SD 7.9) in the intervention group, which was significantly higher than that in the control group (92.9% [SD 10.0]; P=.02). The EuroQol-5D in the intervention group yielded a mean (SD) index of 0.875 (0.156), which was significantly higher than that in the control group (0.741 [0.177]; P<.001). Approximately 60% (29/50) of the patients used the messaging module to communicate with pharmacists. The most frequent types of messages were acknowledgments (77/283, 27.2%), doubts about contraindications and interactions with OAAs (70/283, 24.7%), and consultations for adverse reactions to treatment (39/283, 13.8%). The satisfaction with the app survey conducted in the intervention group yielded an overall mean (SD) score of 9.1 (0.4) out of 10.

Conclusions: Use of e-OncoSalud for the real-time follow-up of patients receiving OAAs facilitated the optimization of some health outcomes. The intervention group had significantly higher health-related quality of life and adherence to treatment than

XSL•FO
RenderX

the control group. Further, the probability of the intervention group presenting with side effects was significantly lower than that of the control group.

(JMIR Mhealth Uhealth 2020;8(10):e20480) doi:10.2196/20480

KEYWORDS

e-OncoSalud; app; smartphone; oral antineoplastic agent; oncology

Introduction

Cancer research has grown exponentially in recent years. It is currently estimated that 40% of the chemotherapeutic drugs are oral antineoplastic agents (OAAs), which have changed the model for administration of chemotherapy. The treatment for cancer has changed from controlled administration in the day hospital to administration at home without the supervision of a health care professional, thus requiring greater autonomy on the part of the patient [1,2]. However, these treatments are subject to drug-related problems (DRPs), such as adherence, interactions with the usual medication, and side effects [3,4]. The Institute for Safe Medication Practices classifies OAAs as high-risk medications [5]. According to Walsh et al [6], up to 20% of the patients treated with OAAs experience severe side effects.

In this sense, information and communication technologies, specifically the area of mobile health, can provide patients with greater autonomy and facilitate communication with health care professionals. Similarly, mobile health can enable health care professionals to improve monitoring and patient care [7]. According to a study in more than 7800 patients, 83% considered the use of technology to be an essential or very important component of their health care. The majority valued the use of virtual care very positively in various scenarios such as receiving reminders for health promotion, taking medication, changing an appointment, and even follow-up on discharge and contacting their health care professional [8].

Since mobile apps are accessible by a vast majority of the population and since these apps facilitate the possibility of remote monitoring, they offer patients with cancer the opportunity to participate in the management of their disease and endow greater responsibility for the control of their health, thus favoring empowerment and improving the safety and quality of care [7].

According to the report “Global Oncology Trends 2018,” it is estimated that there are more than 2500 mobile health cancer-related apps and that their use in clinical practice is increasing, especially in the case of health care apps [9]. Nevertheless, evidence for the benefits that these apps can bring to patients is limited. In fact, 2015 was the first year that this advancement was addressed in the literature, with 5 studies analyzing the effectiveness of apps in oncology. All 5 studies reported positive results, although 2 showed efficacy similar to that in the standard of care. In a review of 17 studies on the use of apps and websites for patients with cancer, it was observed that most were focused exclusively on the design, feasibility, and acceptance of the app. There is hardly any evidence on the effect of apps intended for the patient on the improvement of their health outcomes [10].

Therefore, the objective of this study was to analyze the effect of an app for the pharmacotherapeutic follow-up of patients with oncohematological malignancies undergoing treatment with OAAs on their health outcomes. The secondary objective was to analyze the role of the app in communication with health care professionals and patient satisfaction with the app.

Methods

Design and Scope of the Study

We performed a comparative, quasi-experimental study with a prepost intervention design. The health outcomes of a group of patients in which pharmacotherapeutic follow-up was according to the usual clinical practice of the Pharmacy Service (control group) was compared with those of a group of patients in which the said follow-up was managed through an app (intervention group).

There were 2 periods of recruitment: the first began in January 2015 and ended in December 2015 (control group) and the second began in May 2017 and ended in May 2018 (intervention group). The follow-up period in both the groups was 6 months. This study was carried out in the outpatient unit of the Pharmacy Services of 2 University Hospitals. Both services have a pharmaceutical care program for patients undergoing treatment with OAAs to train them in the appropriate management of their medication (optimal adherence, management of side effects, and interactions with usual medication). This study was approved by the Hospital Clinical Research Ethics Committee (code PI13/02056). All patients signed an informed consent document. This study adhered to the basic ethics principles and the tenets of the Declaration of Helsinki.

Study Population

The control group consisted of adult patients who started treatment with OAAs between January 1, 2015 and December 31, 2015. These patients were followed up at the outpatient unit at the beginning of the treatment and 6 months later according to the established pharmaceutical care program. The intervention group consisted of patients over 18 years of age who started treatment with OAAs between May 31, 2017 and May 31, 2018. Patients had to have a smartphone and were followed up daily through an app (e-OncoSalud) for 6 months from the start of the treatment.

Pharmacotherapeutic Follow-Up Via the App

Between January 2016 and May 2017, a team of hospital pharmacists, oncologists, hematologists, and computer scientists designed the e-OncoSalud app [2]. e-OncoSalud is a

XSL•FO
RenderX

custom-designed app that promotes pharmacotherapeutic follow-up at home and the safety of patients treated with OAAs. It comprises the following 5 modules that integrate all relevant treatment information.

1. An agenda module, wherein the patient can register various events (appointment with the doctor, laboratory analyses, imaging test appointments, and medication collection) with customizable alerts.

2. A treatment module, in which both the patient and the pharmacist can register the drugs consumed by the patient and the dosage. Patients can see their package insert and schedule alerts for administration, thus improving adherence.

3. A module providing tips on the disease and management of symptoms, as well as links to websites of interest and instructions for using the app.

4. A messaging module so that both the patient and the health care professional can contact each other at any time.

5. A module in which the patient can record general progress, blood pressure, weight, and side effects. The patient can register these with a defined periodicity, except for the side effects, which are recorded when they occur. The management of the side effects is based on a decision algorithm that uses a series of questions to classify severity according to the severity of the Common Terminology Criteria for Adverse Events (CTCAE, grades 1-4) and issues appropriate recommendations. The side effect information is focused on the management of fatigue, diarrhea, nausea, vomiting, hand-foot syndrome, and fever. In diarrhea and vomiting, the decision algorithm works based on a count of the number of events registered in the last 24 hours, and when it reaches a defined number of events, the different recommendations appear. All the information that the patient records in the app is sent through a web interface to ensure real-time pharmacotherapeutic follow-up by pharmacists.

Variables

The effect of the app was evaluated through the following health outcomes: drug safety, adherence, and quality of life. Patient communication and satisfaction with the app were also analyzed. The variables analyzed were as follows.

1. Demographic/clinical characteristics: age, sex, diagnosis, type of OAA, and concomitant medications.

2. Safety: DRPs according to the 3rd Consensus of Granada, the number of interactions and their severity, side effects according to the CTCAE scale v.4.03 [11], and resource consumption (unscheduled consultations, emergency visits, and hospital admissions).

3. Quality of life: assessed using the EuroQol-5D (EQ-5D) questionnaire [12].

4. Adherence to treatment: measured through the medication possession ratio, calculated through the dispensation record of the outpatient unit's computer program.

5. Communication: analysis of the messages that patients sent through the app.

6. Satisfaction: evaluated through a specific satisfaction survey for the app. The survey consisted of 8 closed questions on an additive assessment scale that included aspects of the ease of communication with the pharmacist, usefulness for managing the treatment, and grade of recommendation. The survey was delivered to patients 6 months after having installed the app.

Sample Size

The objective of the app was to improve the pharmaceutical follow-up and safety of patients treated with OAAs by increasing the detection of DRPs by at least 20% (from 15% to 35%). Therefore, accepting an alpha risk of .05 and a beta risk of .10 in a bilateral contrast, 45 subjects are needed in the control group and 45 in the intervention group. A loss to follow-up rate of 10% has been estimated.

Statistical Analysis

Data were analyzed using SPSS Statistics for Windows, version 21.0 (IBM Corp). The results were expressed as mean and standard deviation (SD). Categorical variables were expressed as frequencies and percentages. The homogeneity of the 2 groups was analyzed using a univariate analysis by applying the chi-square test for qualitative variables and the one-sided *t* test or Mann-Whitney *U* test to compare quantitative variables. The Bonferroni correction was applied for multiple comparisons. *P* values less than .05 were considered statistically significant.

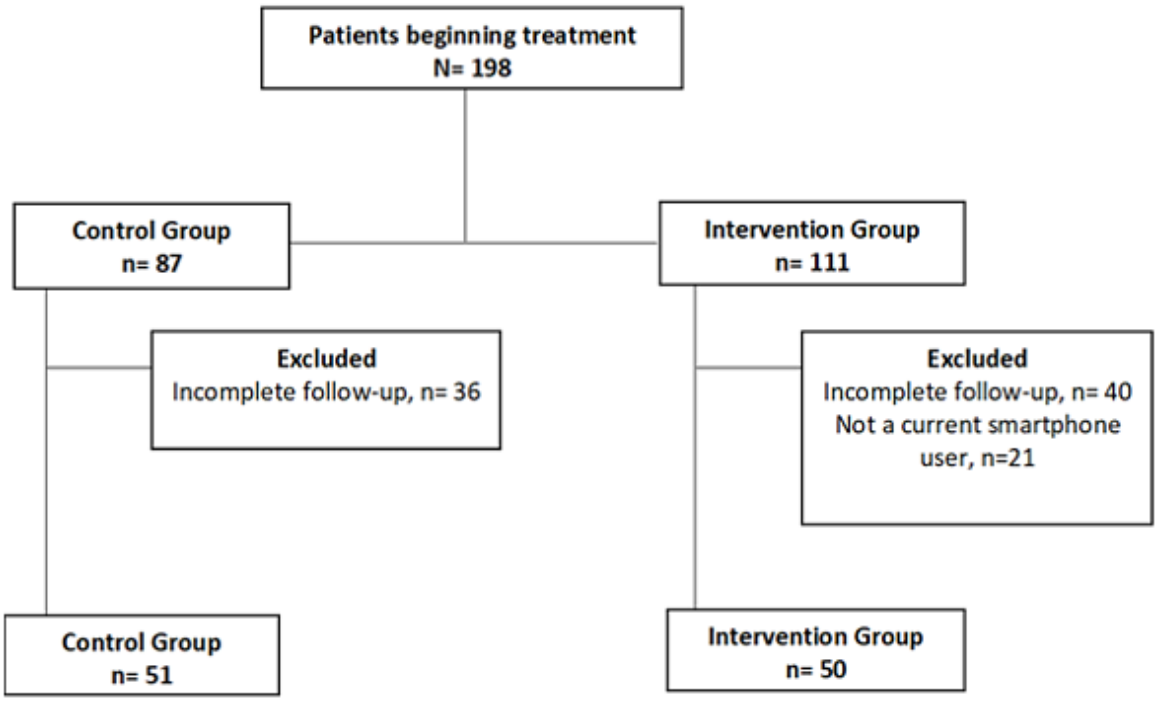
Results

Demographic Data of the Patients

During the study period, 198 patients started treatment with OAAs (Figure 1). A total of 87 patients were assigned to the control group; of these, 36 were excluded because they did not complete the 6 months of treatment. In the intervention group, 111 patients started treatment but 61 patients were excluded (40 for not completing the follow-up period and 21 for not having a smartphone). Finally, 101 patients were analyzed in a 1:1 ratio, with 51 patients in the control group (without the app) and 50 patients in the intervention group (with the app). The mean (SD) age of the patients was 62.7 (13.6) years: 68.7 (10.7) years in the control group and 56.6 (13.6) years in the intervention group ($P<.001$). Table 1 describes the demographic and clinical characteristics of the study patients.

XSL•FO
RenderX

Figure 1. Description of the recruitment of the patients.



XSL•FO
RenderX

Table 1. Demographic and clinical characteristics of the control group and the intervention group.

Demographic/clinical characteristics	Control group, n=51, n (%)	Intervention group, n=50, n (%)	Total, N=101, n (%)	*P* value
Gender				.20
Female	19 (37)	25 (50)	44 (43.6)	
Male	32 (63)	25 (50)	57 (56.4)	
Tumor				.004[a]
Multiple myeloma	13 (26)	5 (10)[a]	18 (17.8)	
Non–small cell lung cancer	12 (24)	5 (10)	17 (16.8)	
Kidney cancer	11 (22)	4 (8)	15 (14.9)	
Prostate cancer	5 (10)	6 (12)	11 (10.9)	
Chronic myeloid leukemia	2 (4)	4 (8)	6 (5.9)	
Breast cancer	3 (6)	2 (4)	5 (5.0)	
Gastrointestinal stromal tumors	1 (2)	3 (6)	4 (4.0)	
Soft tissue sarcoma	0 (0)	4 (8)	4 (4.0)	
Hepatocellular carcinoma	0 (0)	3 (6)	3 (3.0)	
Central nervous system	0 (0)	3 (6)	3 (3.0)	
Thyroid cancer	0 (0)	3 (6)	3 (3.0)	
Chronic lymphatic leukemia	0 (0)	3 (6)	3 (3.0)	
Colon cancer	2 (4)	0 (0)	2 (2.0)	
Myelodysplastic syndrome	2 (4)	0 (0)	2 (2.0)	
Ovarian cancer	0 (0)	2 (4)	2 (2.0)	
Acute lymphoid leukemia	0 (0)	1 (2)	1 (1.0)	
Melanoma	0 (0)	1 (2)	1 (1.0)	
Other	0 (0)	1 (2)	1 (1.0)	
Eastern Cooperative Oncology Group performance status				.04
Grade 0	20 (41)	31 (66)	51 (53.1)	
Grade 1	25 (51)	15 (32)	40 (41.7)	
Grade 2	4 (8)	1 (2)	5 (5.2)	
Smoker				.15
No	51 (100)	48 (96)	99 (98.0)	
Yes	0 (0)	2 (4)	2 (2.0)	
Living alone				.13
No	39 (77)	44 (88)	83 (82.2)	
Yes	12 (24)	6 (12)	18 (17.8)	
Treatment				.001[a]
Lenalidomide	15 (29)	4 (8)[a]	19 (18.8)	
Pazopanib	8 (16)	3 (6)	11 (10.9)	
Gefitinib	7 (14)	1 (2)[a]	8 (7.9)	
Imatinib	4 (8)	4 (8)	8 (7.9)	
Capecitabine	6 (12)	1 (2)	7 (6.9)	
Sorafenib	0 (0)	6 (12)	6 (5.9)	
Erlotinib	5 (10)	0 (0)	5 (5.0)	

XSL•FO
RenderX

Demographic/clinical characteristics	Control group, n=51, n (%)	Intervention group, n=50, n (%)	Total, N=101, n (%)	*P* value
Abiraterone	3 (6)	2 (4)	5 (5.0)	
Enzalutamide	1 (2)	4 (8)	5 (5.0)	
Sunitinib	2 (4)	2 (4)	4 (4.0)	
Dasatinib	0 (0)	3 (6)	3 (3.0)	
Afatinib	0 (0)	3 (6)	3 (3.0)	
Ibrutinib	0 (0)	3 (6)	3 (3.0)	
Olaparib	0 (0)	2 (4)	2 (2.0)	
Regorafenib	0 (0)	2 (4)	2 (2.0)	
Procarbazine	0 (0)	2 (4)	2 (2.0)	
Axitinib	0 (0)	1 (2)	1 (1.0)	
Crizotinib	0 (0)	1 (2)	1 (1.0)	
Everolimus	0 (0)	1 (2)	1 (1.0)	
Ruxolitinib	0 (0)	1 (2)	1 (1.0)	
Dabrafenib/trametinib	0 (0)	1 (2)	1 (1.0)	
Temozolomide	0 (0)	1 (2)	1 (1.0)	
Lenvatinib	0 (0)	1 (2)	1 (1.0)	
Treatment line[b]				.61
1	26 (53)	27 (54)	53 (53.5)	
2	14 (29)	14 (28)	28 (28.3)	
3	8 (16)	5 (10)	13 (13.1)	
4	1 (2)	3 (6)	4 (4.0)	
5	0 (0)	1 (2)	1 (1.0)	
Previous oral chemotherapy				.15
No	36 (72)	42 (84)	78 (78.0)	
Yes	14 (28)	8 (16)	22 (22.0)	

[a]Significant difference between the control and the intervention group at $P<.05$.

[b]Number of previous treatments the patient has received.

Drug Safety

DRPs

DRPs were recorded in 73% (37/51) of the patients in the control group and in 70% (35/50) of the patients in the intervention group (*P*=.01) (Table 2). The most frequent DRPs in both groups were interactions with the usual medication. The probability of the intervention group presenting with side effects was significantly lower than that of the control group presenting with side effects (*P*=.01). The probability of detecting an insufficiently treated health problem in the intervention group was significantly higher than that of detecting an insufficiently treated health problem in the control group (*P*=.04).

XSL•FO
RenderX

Table 2. Drug-related problems in the control group and intervention group.

Drug-related problem	Control group, n=51, n (%)	Intervention group, n=50, n (%)	Total, N=101, n (%)
Erroneous administration of the drug	2 (3)	3 (5)	5 (3.5)
Personal characteristics	0 (0)	2 (3)	2 (1.4)
Contraindication	0 (0)	2 (3)	2 (1.4)
Dose, schedule, or inadequate duration	3 (4)	2 (3)	5 (3.5)
Prescription errors	0 (0)	1 (2)	1 (0.7)
Nonadherence	9 (11)	6 (9)	15 (10.4)
Drug-drug interactions	26 (33)	21 (32)	47 (32.6)
Other health problems affecting treatment	1 (1)	0 (0)	1 (0.7)
Likelihood of side effects	19 (24)[a]	2 (3)	21 (14.6)
Insufficiently treated health problem	5 (6)	11 (17)[a]	16 (11.1)

[a]Significant difference between the control and the intervention groups at $P<.05$.

Side Effects

The proportion of patients who presented with side effects in the group that used the app was significantly lower than that of patients who presented with side effects in the control group (45/50, 90% vs 51/51, 100%; $P>.99$). Table 3 shows the distribution of the patients according to side effects in the control group and intervention group. No statistically significant differences were found between the 2 groups based on side effects. The mean (SD) time to onset of the first adverse effect was 8.2 (10.6) days in the intervention group.

Table 3. Distribution of patients according to side effects in the control group and intervention group.

Side effects	Control group, n=51, n (%)	Intervention group, n=50, n (%)	Total, N=101, n (%)
Nausea	12 (24)	26 (28)	38 (37.6)
Vomiting	3 (6)	13 (20)	16 (15.8)
Diarrhea	21 (41)	46 (50)	67 (66.3)
Hypertension	7 (14)	17 (20)	24 (23.8)
Hematologic toxicity	9 (18)	16 (14)	25 (24.8)
Fatigue	27 (53)	56 (58)	83 (82.2)
Hand-foot syndrome	7 (14)	11 (8)	18 (17.8)
Cutaneous adverse effects	16 (31)	34 (36)	50 (49.5)
Others	28 (55)	59 (62)	87 (86.1)

Resource Consumption

Almost half (25/51, 49%) of the patients in the control group consumed health care resources during the first 6 months of treatment compared with 36% (18/50) of the patients in the intervention group ($P=.78$) (Table 4). Four emergency visits were avoided in the intervention group, since the events were managed remotely. No significant differences were found between the control group and the intervention group with regard to resource consumption.

Table 4. Resource consumption in the control and intervention groups.

Resource consumption	Control group, n=51, n (%)	Intervention group, n=50, n (%)	Total, N=101, n (%)
Unscheduled visits to the oncology department	4 (11)	3 (14)	7 (12.3)
Emergency visits	22 (63)	17 (77)	39 (68.4)
Admission to hospital	9 (26)	2 (9)	11 (19.3)

Health-Related Quality of Life

When a linear transformation was applied to standardize the score, the EQ-5D in the intervention group yielded a mean (SD) index of 0.875 (0.156), which was significantly higher than that in the control group (0.741 [0.177]; $P<.001$).

Adherence

The mean (SD) rate of adherence to treatment in the intervention group was 97.6% (7.9), which was significantly higher than that in the control group (92.9% [10.0]; $P=.02$)

XSL•FO
RenderX

Patient Communication Through e-OncoSalud

Approximately 60% (29/50) of the patients in the intervention group used the messaging module to communicate with pharmacists. They sent 283 messages, that is, an average of 8 messages per patient. The most frequent messages concerned doubts about contraindications and interactions with OAAs (70/283, 24.7%), consultations for adverse reactions to treatment (39/283, 13.8%), and acknowledgment of the care received (77/283, 27.2%) (Table 5).

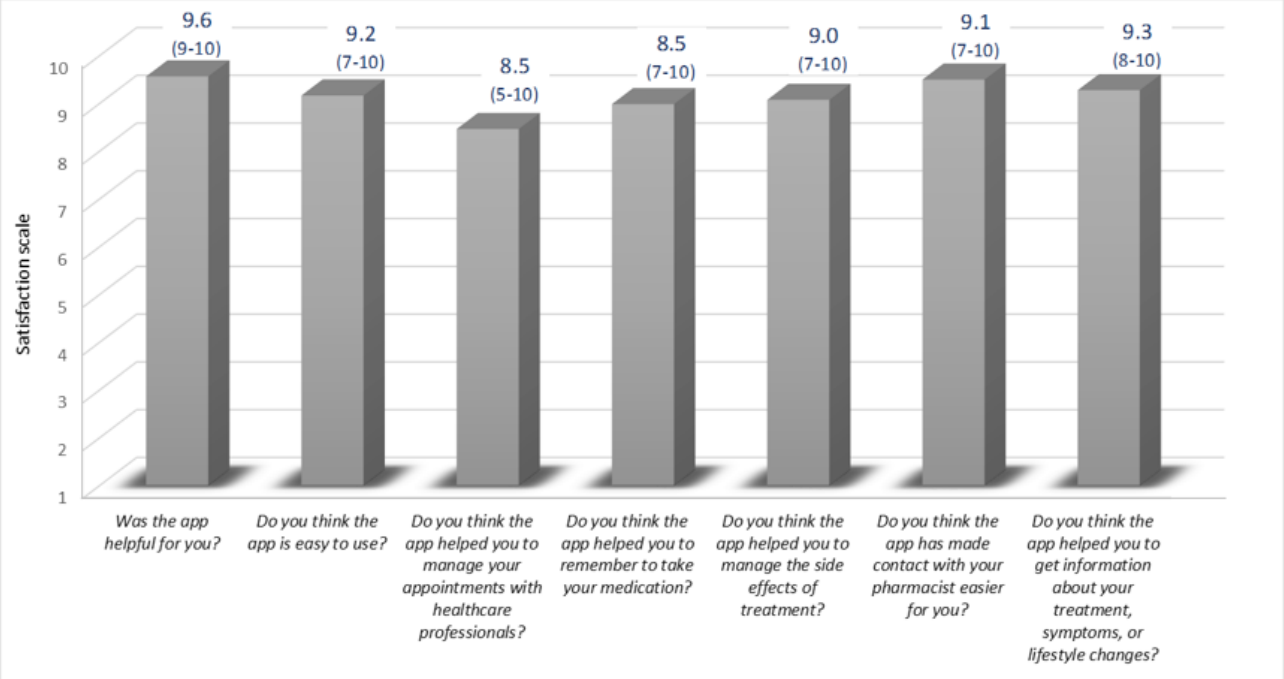
Table 5. Classification of the messages sent by patients through e-OncoSalud (n=283).

Message classification	Values, n (%)
Acknowledgment of the pharmaceutical care received	77 (27.2)
Contraindications and drug interactions	70 (24.7)
Side effects	39 (13.8)
Medication collection alert	36 (12.7)
Dosage and administration	24 (8.5)
Hospital logistics	15 (5.3)
Use of therapy and efficacy of other treatments	7 (2.5)
Availability of other treatments	6 (2.1)
App-related problems	3 (1.1)
Suggestions for improving the app	3 (1.1)
Use and efficacy of complementary medicinal products	2 (0.7)
Other	1 (0.4)

Satisfaction With e-OncoSalud

The satisfaction survey conducted in the intervention group yielded an overall mean (SD) score of 9.1 (0.4) out of 10. The results are shown in Figure 2.

Figure 2. Satisfaction of the patients in the intervention group with the e-OncoSalud app.



Discussion

Principal Findings and Comparison With Previous Studies

To our knowledge, this is the first study to show that an app for the pharmacotherapeutic follow-up of patients with cancer treated with OAAs has a positive impact on their safety and adherence to treatment as well as on health-related quality of life. e-OncoSalud is a new app that includes a unique decision-making algorithm. Depending on the type and severity

XSL•FO
RenderX

of the side effects registered by the patient, the app provides personalized recommendations instantly.

Although OAAs are not subject to the specific problems of intravenous chemotherapy, they are not without side effects [3,13]. In our study, the most common side effects reported by patients in both groups were characteristic of OAAs (asthenia, diarrhea, skin disorders, nausea, and vomiting) [3]. However, by having the decision-making algorithm, the patients in the intervention group were able to manage these symptoms from home. In a 3-module app [14] aimed at patients with nasopharyngeal carcinoma receiving chemoradiotherapy, one of the recommendations decreased the frequency of the complications associated with treatment. The incidence of mucositis, xerostomia, difficulty opening the mouth, and nasal congestion was lower in the patients who used the app (67 patients) than in those who did not (65 patients) [14]. In a clinical trial conducted in 76 patients with breast cancer [15], the authors analyzed the impact of the ILOVEBREAST app in reducing side effects, improving psychological status, and improving adherence to treatment. The patients were randomized in a 1:1 ratio to follow-up through the app (intervention group) or traditional follow-up (control group). Consistent with our findings, the results showed that the use of the app was associated with better quality of life, greater adherence, and fewer side effects such as nausea, fatigue, hand-foot syndrome, and hair loss. Although the differences were not significant in most cases, the incidence of grade 3 fatigue was significantly lower in patients who used the app (1 vs 13; *P*=.02) [15].

According to a study conducted at the Dana-Farber Cancer Institute, patients undergoing treatment with OAAs are more likely to be admitted to hospital with side effects. However, patients who were closely followed by their health care professional were significantly less likely to be admitted in the hospital [16]. Thus, information and communication technologies are demonstrating an improvement in the management of various conditions, with results that show a reduction in complications and hospital admissions. For example, a study that analyzed the use of text messages in patients with diabetes found a significant decrease in HbA_{1c} levels, improved medication adherence, and decreased emergency department visits [17]. Another study evaluated the effect of home telemonitoring in patients with heart failure and demonstrated a significant improvement in health outcomes [18]. However, we found no studies showing that apps for patients with cancer reduce their emergency visits or hospital admissions. Through e-OncoSalud, patients' consumption of resources could be reduced from 49% to 36% (unscheduled consultations, emergency visits, or hospital admissions), although there was no statistically significant difference between the control and the intervention groups.

Side effects can also affect the quality of life of patients. According to Rincon et al [19], real-time monitoring and symptom management significantly improved emotional status, insomnia, and urinary tract symptoms. In another study that analyzed an app for the follow-up of patients with lung cancer, registration and follow-up of symptoms and subsequent management by the oncologist improved the quality of life [20]. According to the authors, this improvement was due to the early detection of side effects, complications, and signs or symptoms of progression [20]. In our study, follow-up based on the use of the e-OncoSalud app improved the health-related quality of life more than the traditional follow-up.

While patients with cancer face many challenges, adherence to OAAs is crucial if they are to maximize treatment outcomes and avoid complications [21]. In the follow-up based on e-OncoSalud, adherence at 6 months was 97.6% (SD 7.9), which is significantly higher than that of patients who did not receive home follow-up through the app (92.9% [SD 10.0]). These results agree with those reported in the literature, where the app is positioned as another strategy to improve adherence to treatment [22].

Regarding patient-pharmacist communication, 60% of the patients used the messaging module to communicate with the pharmacist. This percentage is similar to that reported with the WebChoice app [23], in which 61% of the patients used the messaging module to contact their nurse. However, we were unable to find studies that analyze the type of messages sent through an app between patients and health care professionals. This is not surprising, given that few electronic systems for patients with cancer incorporate messaging modules. A recently published review showed that only 6 (15%) of 40 electronic systems incorporated this modality [10]. Although several studies have shown that patients highly value this function, its complexity and maintenance limit its usability [10,24-26].

Finally, our results revealed a high degree of satisfaction with most of the aspects assessed. Those aspects that were the best rated were the ability to communicate with the pharmacist from anywhere, having the treatment registered with the notification system, the immediacy of the response by the pharmacist, and the ease of use. In addition, 100% of the patients agreed that they would recommend the app. These data are similar to those reported in studies that analyzed the app satisfaction of patients with cancer [15,25]. In order to assess the satisfaction of the ILOVEBREAST app, Kim et al [15] conducted a survey consisting of 8 questions that covered the ease of use and help with taking medication. Patients were also asked if they would recommend the app to others. The functions most highly valued by the patients were the ease of obtaining information and managing side effects. However, unlike our app, ease of use was one of the least valued aspects, possibly because it is an app based on an avatar type game. In their evaluation of the iCancerHealth app, Berry et al [25] observed that the aspects best valued by patients were ease of communication with their health care professional and the recording and monitoring of side effects.

Limitations

This study has the following limitations. The main limitation is the absence of randomization. However, the inclusion and exclusion criteria ensured that the selection of patients was representative of a usual clinical practice, and the sample size was sufficiently large to achieve the objectives. Likewise, the control and intervention groups were recruited 3 years apart. This could have an impact on not only which antineoplastic drugs are available but also other factors in terms of people's

XSL•FO
RenderX

willingness to communicate via an app. In addition, the use of 2 separate study periods with respect to a type of drugs undergoing development makes it difficult to draw reasonable comparisons between the groups.

Conclusions

Real-time pharmacotherapeutic follow-up of patients receiving OAAs by using the e-OncoSalud enabled the optimization of some health outcomes. Although there was no significant difference between the control and intervention groups in terms of unscheduled visits to the oncology department, emergency visits, or admissions to the hospital, the intervention group had significantly higher health-related quality of life and adherence to treatment. Further, the probability of the intervention group presenting with side effects was significantly lower than that of the control group. However, more randomized studies should be conducted to confirm the observed findings.

Acknowledgments

This project was supported by iPharma (Pharmacy Innovation Center at the Hospital General Universitario Gregorio Marañón) through funding received from FEDER Project PI13/02056. Support was cofinanced by the European Regional Development Fund (FEDER)–A way of making Europe.

Conflicts of Interest

None declared.

References

1. Bender JL, Yue RYK, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. J Med Internet Res 2013 Dec 23;15(12):e287 [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
2. Collado-Borrell R, Escudero-Vilaplana V, Ribed A, Anglada-Martínez H, Martín-Conde M, Herranz-Alonso A, et al. Design of the functional requirements of a smartphone app for patients receiving oral antineoplastic agents: The e-OncoSalud® app. J Oncol Pharm Pract 2020 Jan;26(1):105-115. [doi: 10.1177/1078155219840419] [Medline: 30955468]
3. Escudero-Vilaplana V, Ribed A, Romero-Jimenez RM, Herranz-Alonso A, Sanjurjo-Saez M. Pharmacotherapy follow-up of key points in the safety of oral antineoplastic agents. Eur J Cancer Care (Engl) 2017 May;26(3). [doi: 10.1111/ecc.12463] [Medline: 26872286]
4. Collado-Borrell R, Escudero-Vilaplana V, Romero-Jiménez R, Iglesias-Peinado I, Herranz-Alonso A, Sanjurjo-Sáez M. Oral antineoplastic agent interactions with medicinal plants and food: an issue to take into account. J Cancer Res Clin Oncol 2016 Nov;142(11):2319-2330. [doi: 10.1007/s00432-016-2190-8] [Medline: 27316629]
5. Institute For Safe Medication Practices. ISMP List of High-Alert Medication. Institute for Safe Medications Practices. URL: https://www.ismp.org/sites/default/files/attachments/2018-08/highAlert2018-Acute-Final.pdf [accessed 2019-12-10]
6. Walsh KE, Dodd KS, Seetharaman K, Roblin DW, Herrinton LJ, Von Worley A, et al. Medication errors among adults and children with cancer in the outpatient setting. J Clin Oncol 2009 Feb 20;27(6):891-896. [doi: 10.1200/JCO.2008.18.6072] [Medline: 19114695]
7. Collado-Borrell R, Escudero-Vilaplana V, Ribed-Sánchez A, Ibáñez-García S, Herranz-Alonso A, Sanjurjo-Sáez M. Smartphone applications for cancer patients; what we know about them? Farm Hosp 2016 Jan 01;40(1):25-35 [FREE Full text] [doi: 10.7399/fh.2016.40.1.8993] [Medline: 26882831]
8. Encuesta DUSSD. Accenture Cloud First. Accenture. URL: https://www.accenture.com/_acnmedia/pdf-83/accenture-health-meet-todays-health [accessed 2019-12-20]
9. Aitken M, Kleinrock M, Simorellis A, Nass D. Global oncology trends 2018, innovation, expansion and disruption. IQVIA Institute for Human Data Science. 2018. URL: https://www.iqvia.com/insights/the-iqvia-institute/reports/global-oncology-trends-2018 [accessed 2019-12-20]
10. Warrington L, Absolom K, Conner M, Kellar I, Clayton B, Ayres M, et al. Electronic Systems for Patients to Report and Manage Side Effects of Cancer Treatment: Systematic Review. J Med Internet Res 2019 Jan 24;21(1):e10875 [FREE Full text] [doi: 10.2196/10875] [Medline: 30679145]
11. U.S. DOHAHS. Common terminology criteria for adverse events (CTCAE). Version 4. URL: https://www.eortc.be/services/doc/ctc/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf [accessed 2019-12-10]
12. 2016. EQ-5D, EuroQol-5D questionnaire. URL: https://euroqol.org/eq-5d-instruments/ [accessed 2019-12-10]
13. Ribed A, Romero-Jiménez RM, Escudero-Vilaplana V, Iglesias-Peinado I, Herranz-Alonso A, Codina C, et al. Pharmaceutical care program for onco-hematologic outpatients: safety, efficiency and patient satisfaction. Int J Clin Pharm 2016 Apr;38(2):280-288. [doi: 10.1007/s11096-015-0235-8] [Medline: 26715547]
14. Di R, Li G. Use of a Smartphone Medical App Improves Complications and Quality of Life in Patients with Nasopharyngeal Carcinoma Who Underwent Radiotherapy and Chemotherapy. Med Sci Monit 2018 Sep 04;24:6151-6156 [FREE Full text] [doi: 10.12659/MSM.908146] [Medline: 30178780]

XSL•FO
RenderX

15. Kim HJ, Kim SM, Shin H, Jang J, Kim YI, Han DH. A Mobile Game for Patients With Breast Cancer for Chemotherapy Self-Management and Quality-of-Life Improvement: Randomized Controlled Trial. J Med Internet Res 2018 Oct 29;20(10):e273 [FREE Full text] [doi: 10.2196/jmir.9559] [Medline: 30578205]
16. Siefert ML, Bonquist TM, Berry DL, Hong F. Symptom-related emergency department visits and hospital admissions during ambulatory cancer treatment. J Community Support Oncol 2015 May;13(5):188-194 [FREE Full text] [doi: 10.12788/jcso.0134] [Medline: 27119127]
17. Valles Ortiz PM, Miranda Félix P, García Sosa ES. Mensajes de texto para el control glucémico en adultos con diabetes tipo 2: revisión sistemática. eglobal 2015 Jan 01;14(1). [doi: 10.6018/eglobal.14.1.188461]
18. Margolis KL, Asche SE, Dehmer SP, Bergdall AR, Green BB, Sperl-Hillen JM, et al. Long-term Outcomes of the Effects of Home Blood Pressure Telemonitoring and Pharmacist Management on Blood Pressure Among Adults With Uncontrolled Hypertension: Follow-up of a Cluster Randomized Clinical Trial. JAMA Netw Open 2018 Sep 07;1(5):e181617 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.1617] [Medline: 30646139]
19. Rincon E, Monteiro-Guerra F, Rivera-Romero O, Dorronzoro-Zubiete E, Sanchez-Bocanegra CL, Gabarron E. Mobile Phone Apps for Quality of Life and Well-Being Assessment in Breast and Prostate Cancer Patients: Systematic Review. JMIR Mhealth Uhealth 2017 Dec 04;5(12):e187 [FREE Full text] [doi: 10.2196/mhealth.8741] [Medline: 29203459]
20. Denis F, Lethrosne C, Pourel N, Molinier O, Pointreau Y, Domont J, et al. Randomized Trial Comparing a Web-Mediated Follow-up With Routine Surveillance in Lung Cancer Patients. J Natl Cancer Inst 2017 Sep 01;109(9). [doi: 10.1093/jnci/djx029] [Medline: 28423407]
21. Marin D, Bazeos A, Mahon F, Eliasson L, Milojkovic D, Bua M, et al. Adherence is the critical factor for achieving molecular responses in patients with chronic myeloid leukemia who achieve complete cytogenetic responses on imatinib. J Clin Oncol 2010 May 10;28(14):2381-2388 [FREE Full text] [doi: 10.1200/JCO.2009.26.3087] [Medline: 20385986]
22. Dayer L, Heldenbrand S, Anderson P, Gubbins PO, Martin BC. Smartphone medication adherence apps: potential benefits to patients and providers. J Am Pharm Assoc (2003) 2013;53(2):172-181 [FREE Full text] [doi: 10.1331/JAPhA.2013.12202] [Medline: 23571625]
23. Ruland CM, Jeneson A, Andersen T, Andersen R, Slaughter L, Bente-Schjødt-Osmo, et al. Designing tailored Internet support to assist cancer patients in illness management. AMIA Annu Symp Proc 2007 Oct 11:635-639 [FREE Full text] [Medline: 18693913]
24. Collado-Borrell R, Escudero-Vilaplana V, Calles A, Garcia-Martin E, Marzal-Alfaro B, Gonzalez-Haba E, et al. Oncology Patient Interest in the Use of New Technologies to Manage Their Disease: Cross-Sectional Survey. J Med Internet Res 2018 Oct 23;20(10):e11006 [FREE Full text] [doi: 10.2196/11006] [Medline: 30355554]
25. Berry DL, Blonquist TM, Nayak MM, Grenon N, Momani TG, McCleary NJ. Self-Care Support for Patients with Gastrointestinal Cancer: iCancerHealth. Appl Clin Inform 2018 Oct;9(4):833-840 [FREE Full text] [doi: 10.1055/s-0038-1675810] [Medline: 30463094]
26. Andersen T, Ruland CM. Cancer patients' questions and concerns expressed in an online nurse-delivered mail service: preliminary results. Stud Health Technol Inform 2009;146:149-153. [Medline: 19592825]

Abbreviations

CTCAE: Common Terminology Criteria for Adverse Events
DRP: drug-related problem
EQ-5D: EuroQol-5D
OAA: oral antineoplastic agent

Edited by G Eysenbach; submitted 20.05.20; peer-reviewed by S Davis, A Lambert; comments to author 29.06.20; revised version received 23.08.20; accepted 13.09.20; published 16.10.20.

Please cite as:
Collado-Borrell R, Escudero-Vilaplana V, Ribed A, Gonzalez-Anleo C, Martin-Conde M, Romero-Jimenez R, Iglesias-Peinado I, Herranz-Alonso A, Sanjurjo-Saez M
Effect of a Mobile App for the Pharmacotherapeutic Follow-Up of Patients With Cancer on Their Health Outcomes: Quasi-Experimental Study
JMIR Mhealth Uhealth 2020;8(10):e20480
URL: http://mhealth.jmir.org/2020/10/e20480/
doi:10.2196/20480
PMID:33064100

XSL•FO
RenderX

©Roberto Collado-Borrell, Vicente Escudero-Vilaplana, Almudena Ribed, Cristina Gonzalez-Anleo, Maite Martin-Conde, Rosa Romero-Jimenez, Irene Iglesias-Peinado, Ana Herranz-Alonso, Maria Sanjurjo-Saez. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 16.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Hospital-Community-Family–Based Telemedicine (HCFT-AF) Program for Integrative Management of Patients With Atrial Fibrillation: Pilot Feasibility Study

Jiang Jiang[1*], MA; Xiang Gu[2,3*], MD, PhD; Chen-Di Cheng[4*], MA; Hong-Xiao Li[2,3*], MD; Xiao-Lin Sun[2,3], MD; Ruo-Yu Duan[1], MA; Ye Zhu[2,3], MD; Lei Sun[2,3], MA; Fu-Kun Chen[2,3], MD; Zheng-Yu Bao[2,3], MD; Yi Zhang[2,3], MA; Jian-Hua Shen[2,3], MA

[1]Dalian Medical University, Dalian, China

[2]Medical College of Yangzhou University, Yangzhou, Jiangsu, China

[3]Department of Cardiology, Northern Jiangsu People's Hospital, Yangzhou, China

[4]Second Affiliated Hospital, Xiang-Ya Medical College of Central South University, Changsha, China

[*]these authors contributed equally

Corresponding Author:
Xiang Gu, MD, PhD
Medical College of Yangzhou University
No. 98, Nantong West Road
Yangzhou, Jiangsu, 225001
China
Phone: 86 051487373366
Email: guxiang@yzu.edu.cn

Abstract

Background: The potential effectiveness of integrated management in further improving the prognosis of patients with atrial fibrillation has been demonstrated; however, the best strategy for implementation remains to be discovered.

Objective: The aim of this study was to ascertain the feasibility of implementing integrated atrial fibrillation care via the Hospital-Community-Family–Based Telemedicine (HCFT-AF) program.

Methods: In this single-arm, pre-post design pilot study, a multidisciplinary teamwork, supported by efficient infrastructures, provided patients with integrated atrial fibrillation care following the Atrial fibrillation Better Care (ABC) pathway. Eligible patients were continuously recruited and followed up for at least 4 months. The patients' drug adherence, and atrial fibrillation–relevant lifestyles and behaviors were assessed at baseline and at 4 months. The acceptability, feasibility, and usability of the HCFT-AF technology devices and engagement with the HCFT-AF program were assessed at 4 months.

Results: A total of 73 patients (mean age, 68.42 years; 52% male) were enrolled in November 2019 with a median follow up of 132 days (IQR 125–138 days). The patients' drug adherence significantly improved after the 4-month intervention (P<.001). The vast majority (94%, 64/68) of indicated patients received anticoagulant therapy at 4 months, and none of them received antiplatelet therapy unless there was an additional indication. The atrial fibrillation–relevant lifestyles and behaviors ameliorated to varying degrees at the end of the study. In general, the majority of patients provided good feedback on the HCFT-AF intervention. More than three-quarters (76%, 54/71) of patients used the software or website more than once a week and accomplished clinic visits as scheduled.

Conclusions: The atrial fibrillation–integrated care model described in this study is associated with improved drug adherence, standardized therapy rate, and lifestyles of patients, which highlights the possibility to better deliver integrated atrial fibrillation management.

Trial Registration: Clinicaltrials.gov NCT04127799; https://clinicaltrials.gov/ct2/show/NCT04127799

(JMIR Mhealth Uhealth 2020;8(10):e22137) doi:10.2196/22137

KEYWORDS

atrial fibrillation; integrative management; telemedicine; self-management; feasibility study

XSL•FO
RenderX

Introduction

Atrial fibrillation is the most prevalent cardiac arrhythmia and has cumulatively been acknowledged as a major health care burden globally [1]. Despite the fact that atrial fibrillation is infrequent among young people (<1% in individuals aged <40 years), its incidence dramatically increases with age, reaching up to 10%-17% in individuals above the age of 80 years [1,2]. The lifetime risk of atrial fibrillation is 21%-23% in women and 17%-26% in men [3,4]. Approximately 33.5 million individuals were suffering from atrial fibrillation in 2010 globally, which is predicted to double by 2050 owing to widespread population aging [5]. Additionally, atrial fibrillation is associated with a 5-fold increase in ischemic stroke risk and accounts for 15%-20% of all strokes [6,7]. Strokes arising from atrial fibrillation are more catastrophic and disabling than those of other etiologies, which has been described as an "atrial fibrillation–correlative stroke tsunami" [8].

Oral anticoagulants therapy can markedly reduce the risk of stroke by 64% and the risk of death by 26% in patients with atrial fibrillation [9]. However, the underuse or improper use of oral anticoagulants is fairly common in real-world clinical practice, even in the new era of nonvitamin K antagonist oral anticoagulants (NOACs), especially in many Asian countries [10-12]. In comparison with Europe (90.1%) and North America (78.3%), the reported usage rate of oral anticoagulants was notably lower in Asia (55.2%) and well below the global average (79.9%) [11]. Furthermore, 17.7% of high-risk patients were not anticoagulated (Europe 8.8%; Asia 42.4%), whereas 76.5% of low-risk patients were inappropriately anticoagulated [11]. In Asia, oral anticoagulant use varies from 21.0% in China (5.8% for NOACs) to 89.7% in Japan [11].

The integrative management of atrial fibrillation patients is deemed to have the potential to improve the oral anticoagulant rate and patient prognosis, which comprises the following core elements: (1) patient-centeredness, (2) multidisciplinary teamwork, (3) utilization of intelligent technology, and (4) application of comprehensive strategies with access to all therapy options [13-15]. Additionally, the European Society of Cardiology developed an app to support the management of patients with atrial fibrillation [16]. To date, several preliminary studies have demonstrated the benefits of integrated atrial fibrillation care in reducing readmission rates [17], cardiovascular and all-cause mortality [17-20], and atrial fibrillation–associated health care expenditures [21]. However, a majority of prior studies have been dominated by nurses [17-19] with latent gaps in the better implementation of guidelines [22]. Additionally, these were mostly single-center studies covering a finite region or supported by a single management strategy such as a structured telephone follow up [23], web platform [16], or app [24].

To our knowledge, no subsistent program has integrated a mobile app, web platform, and intelligent health monitoring devices with cooperation among specialists, general practitioners, patients, and their caregivers to care and empower patients in disease self-management. There are scarce data on the implementation of such an integrated program to manage atrial fibrillation patients, particularly regarding its feasibility and safety. On the basis of our heart failure management program [25,26], we conducted the Hospital-Community-Family–Based Telemedicine (HCFT-AF) program incorporating the core elements mentioned above. The aim of this study was to ascertain the feasibility of the HCFT-AF program for implementation of integrated atrial fibrillation care.

Methods

Study Design and Participants

This single-arm, pre-post design study was part of the HCFT-AF program (ClinicalTrials.gov NCT04127799), which received ethics approval (2019093) from the Institutional Review Board of Northern Jiangsu People's Hospital [27]. Eligible patients (Textbox 1) were enrolled consecutively in November 2019 and were followed up for at least 4 months (Figure 1). All patients provided written informed consent before enrolling in the study.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

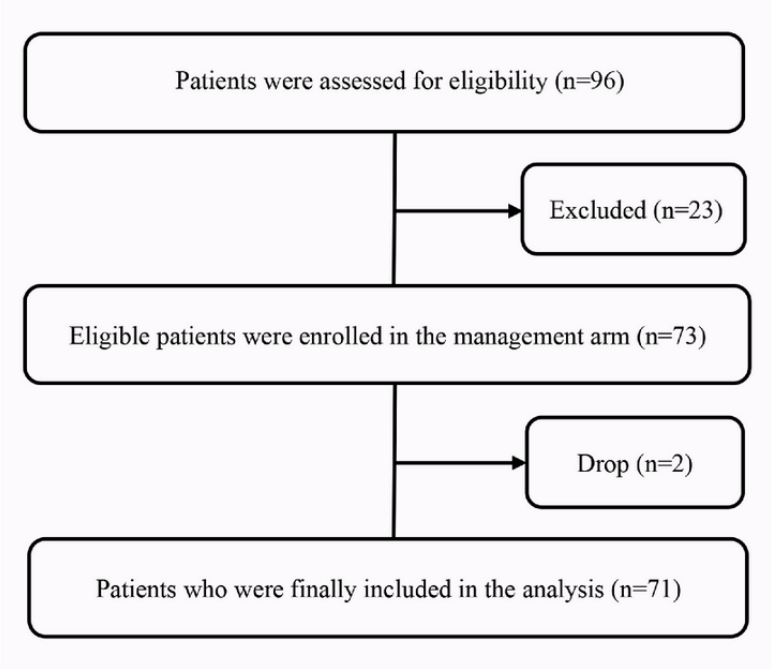
Inclusion criteria

- Aged≥18 years
- Meeting the diagnostic criteria for atrial fibrillation
- Understand the nature of the study, and agree to sign informed consent and continue follow up

Exclusion criteria

- Atrial fibrillation due to a reversible cause (eg, untreated hyperthyroidism, acute myocardial infarction, or acute myocarditis within 1 month)
- No recurrence of atrial fibrillation after surgical treatment
- Combined with other diseases with a life expectancy less than 1 year
- Severe liver and kidney disfunction: serum creatinine>5.0 mg/dL; alanine transaminase exceeds the reference value by more than 3 times (>100 U/L)
- Systolic/diastolic blood pressure≥180/110 mmHg, but can be enrolled after achieving blood pressure control
- Diagnosed or suspected blood system disease (except for mild to moderate anemia), leading to coagulopathy or combined with bleeding tendency
- Pregnant and lactating women
- Unable to use remote monitoring equipment (eg, depression, dementia)
- Participating in other treatment research or remote patient management programs
- Investigators considered that it is inappropriate to participate in the study

Figure 1. Flow diagram of the pilot study.



Supportive System in the HCFT-AF Program

System Overview

A user-friendly supportive system was developed to better deliver integrated atrial fibrillation management. In general, this system incorporates the following three major components: (1) a multifunctional service platform (Physio-Gate PG 2000, GETEMED AG), (2) a personal health data app (King OPTO-Electronic AG), and (3) a few health monitoring devices.

Multifunctional Service Platform

The multifunctional service platform [28] covers patients' data collection, storage, education, and audio-video portions. In our initial version, patients' data could only be collected manually, which was time-consuming, labor-intensive, and had a high error rate. In version 2.0, the platform can automatically create a structured medical record via a document exported from the hospital information system. Additionally, data uploaded by participants through the app can be deposited on the platform. Abnormal values are coded with different colors: green denotes reduction, orange denotes moderate risk, red denotes high risk,

XSL•FO
RenderX

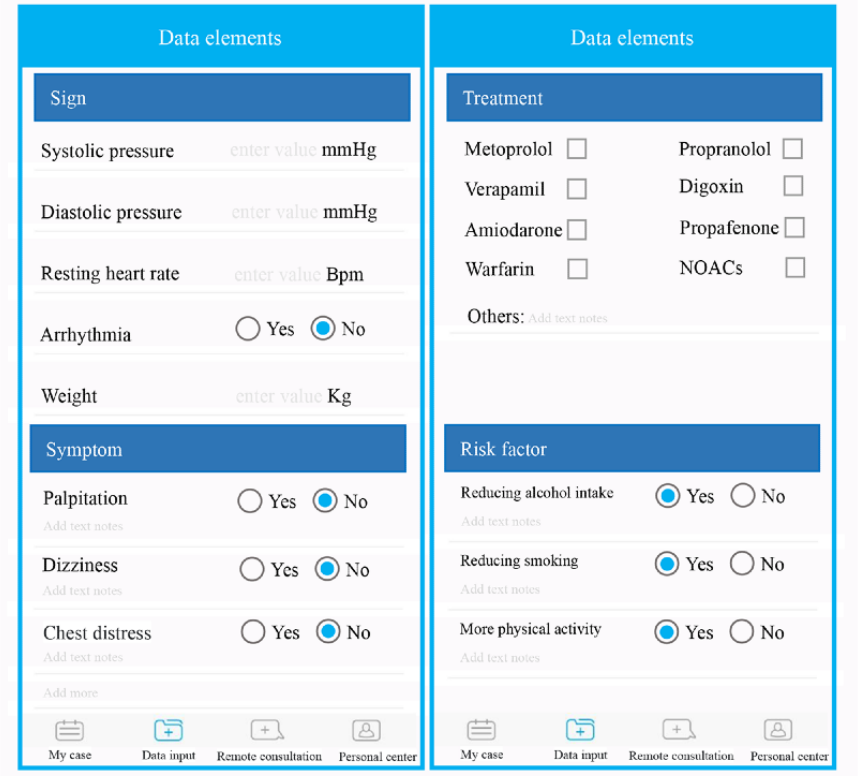
and no color denotes normal or approximating the normal range. In this way, physicians can master patients' condition more comprehensively and rapidly. The online library on the platform provides multifaceted knowledge on atrial fibrillation (eg, guideline-based therapy, guidance for self-management). The platform is also equipped with a highly confidential audio-video system to protect patient privacy. In brief, this platform serves as a bridge for information interaction and resource sharing between all stakeholders in this program.

Mobile App

A user-friendly mobile app, available on both Android and iOS, was respectively developed for patients and physicians. The data obtained with access to an internet connection can be accessed offline, after the network is disconnected, while subsequent updates require connection to the network. The version for physicians assists them in better managing patients and allows them to observe and track the health status of patients more conveniently. Physicians can view the medical records of patients in their charge. When patients upload data or seek remote consultation, their attending physicians will be reminded.

The patient version has the following major functions: medical record viewing, health-related data uploading, and remote consultation (Figure 2). This app enables patients to view their medical records documented by physicians, which is significant for those taking multiple drugs simultaneously. As an important part of self-management, patients can record and upload health-related data via the app, including daily recording of symptom changes, resting ventricular rate, rhythm, and blood pressure. Their attending physicians will then examine the data and provide corresponding suggestions. In addition, patients can communicate with physicians more conveniently with the remote consultation function. Bleeding caused by warfarin is partly due to unreasonable medication combinations, which can be avoided by timely communication with clinicians [29]. Furthermore, a reminder alert is sent automatically for forthcoming clinic appointments and a structured follow-up schedule is accessible in the app.

Figure 2. Screenshots of the mobile app for patients.



Intelligent Health Monitoring Devices

In addition to the ACT Plus automatic clotting time tester (Medtronic PLC, Minneapolis, MN, USA), some intelligent health monitoring devices are available in our program: a multicomponent monitor (blood pressure, oxyhemoglobin saturation, and electrocardiograph; TE-4000Y, Beijing Hailiying Medical Technology Ltd, Beijing, China) and a long-term wearable electrocardiograph monitor (BECG1200-A, Thoth Medical Technology Ltd, Suzhou, China) (see Multimedia Appendix 1). These health monitoring devices make remote monitoring, recording, and analysis of patient health status possible.

Procedures

In general, the intervention included three major elements: education, streamlined guideline-based therapy, and periodic follow up. The educational section mainly had two target audiences. The first was patients and their caregivers, incorporating the detriment and self-management of atrial fibrillation, use of supportive health monitoring devices, and other components. The second was general practitioners in communities, aiming to narrow the gap between physicians and guideline-based therapy.

Researchers introduced the study protocol to the eligible patients and their caregivers. In addition, for integrating the management

XSL•FO
RenderX

group, research assistants helped to install the app, and provided instructions on use of the multifunctional service platform, app, and health monitoring devices. A complete medical record was established on the multifunctional service platform before the study officially started. Thereafter, patients were assigned to the community hospital closest to them for further follow up. Structured follow up was arranged once every 2-4 months after discharge, and related examinations were performed every 3-6 months or depending on the condition of patients. During periodic clinical visits, participants were encouraged to record and upload their health-related data via the app and to follow the educational program to improve their disease self-management ability. For patients who were at high risk of severe arrhythmia, some intelligent health monitoring devices were furnished to record and transfer health parameters remotely. Patients and their families could interact with their supervising general practitioners more conveniently via the app if needed.

The general practitioners in community hospitals checked the data submitted by the patients and identified abnormal health data in comparison with previous medical records. In such cases, they would give appropriate treatments or consult specialists at a regional central hospital according to the risk level of the patients. Patients can be transferred to the central hospital through a fast track if necessary. In addition to managing patients, community physicians can also access the latest guideline-based atrial fibrillation therapy through the multifunctional service platform.

Cardiologists of regional central hospitals performed remote ward rounds weekly for patients that were deemed to be complicated or difficult to handle via the dedicated audio-video system by general practitioners, and clinic visit schedules were adjusted according to the condition of the patient when necessary. For patients with complicated conditions, cardiologists would discuss the treatment schedule with neurologists, cardiac surgeons, and other specialists. Furthermore, specialists performed online video seminars regularly for general practitioners and patients, aiming to increase the capacity of general practitioners and compliance of patients to the standardized therapy.

Outcomes and Instruments

Patients' drug adherence was assessed via the Pharmacy Quality Alliance adherence measure [30] at baseline and at 4 months individually. Patients' lifestyles and behaviors associated with the occurrence and progress of atrial fibrillation were collected at baseline and 4 months through interviews with the purpose of evaluating changes in self-management. Additionally, the acceptability, feasibility, and usability of the HCFT-AF intervention for patients were measured via the Perceived Health Web Site Usability Questionnaire [31] at 4 months. This questionnaire consists of three separate portions that evaluate patient satisfaction (eg, "It is easy to find specific information"), ease of use (eg, "I found the HCFT-AF intervention easy to learn"), and usefulness (eg, "Using the HCFT-AF intervention will help me improve my knowledge about health"). All points are rated on a 1 to 7 scale. Responses were averaged for each element and across all points, with higher scores indicating better satisfaction, easier use, higher effectiveness, and greater overall usability of the intervention. The engagement of the HCFT-AF program was roughly estimated by the frequency of patients' account logins and the completion degree of the follow-up schedule. The patients' feedback with the intervention was obtained through self-reported questionnaires at the end of the study.

Statistical Analysis

Demographic traits of patients were collected and are presented as continuous or categorical variables. Continuous variables, verified for normality by the Kolmogorov-Smirnov test, are presented as means (SD) or median (IQR) as appropriate. In addition, the data were compared via a *t* test or Mann-Whitney *U* test, as appropriate. Categorical variables are reported as the absolute number and percentage, which were analyzed by the chi-square test or Fisher exact test. All statistical tests were two-tailed, and $P<.05$ was considered to indicate statistical significance. Statistical analyses were performed on GraphPad Prism version 8.00 for Windows (GraphPad Software, La Jolla, CA, USA).

Results

Participant Characteristics

A total of 73 eligible patients (mean age, 68.42 years; 52% male) were enrolled in this feasibility study in November 2019 with a median follow up of 132 days (IQR 125–138). Of the original cohort, 71 patients completed the 4-month follow up. Two patients dropped out after failing to complete regular clinical follow ups. Basic demographic data were collected at baseline (Table 1). Less than half of the qualified patients received anticoagulant therapy. Fifty-three physicians (mean age, 41.72 years; 58% were male) participated in the routine management of patients. Nearly half of the physicians (25/53, 47%) had more than 10 years of work experience. About 30% (16/53) of the physicians were specialists, including cardiologists (20%, 11/53), neurologists (6%, 3/53), and cardiac surgeons (4%, 2/53).

During the study period, specialists from the regional central hospital conducted 34 online lectures and 168 remote ward rounds in total. In addition, general practitioners consulted the experts in the regional central hospital remotely 102 times. One patient with high-grade atrioventricular block detected by intelligent health monitoring devices was transferred to the regional central hospital via a fast track. Furthermore, 5 patients had bleeding gums and 2 had epistaxis; however, the anticoagulation treatment was not interrupted after appropriate treatment was given. No other relevant serious adverse events were reported.

XSL•FO
RenderX

Table 1. Baseline characteristics of patients in the study (N=73).

Characteristics	Value
Age (years), mean (SD)	68.42 (10.25)
Male, n (%)	38 (52)
Medical history, n (%)	
Hypertension	37 (51)
Diabetes	16 (22)
Congestive heart failure	15 (21)
Previous stroke/TIA[a]	10 (14)
Renal dysfunction	6 (8)
Liver dysfunction	4 (6)
Peripheral vascular disease	4 (6)
Type of atrial fibrillation, n (%)	
Paroxysmal	35 (48)
Persistent	28 (38)
Permanent	10 (14)
Atrial fibrillation treatment, n (%)	
Beta-blocker	34 (47)
Pharmacologic cardioversion	15 (21)
Anticoagulant therapy[b]	28 (38)
Atrial fibrillation ablation	8 (11)
LAAO[c]	1 (1)
Anthropometric data, n (%)	
BMI (kg/m^2)	23.15 (5.49)
Overweight[d]	21 (29)
Current drinker	36 (49)
Current smoker	28 (38)
Daily caregiver, n (%)	
Spouse or children	62 (85)
Relative	8 (11)
Others	3 (4)
CHA2DS2-VASc[e] score, mean (SD)	2.89 (1.71)
HAS-BLED[f] score, mean (SD)	2.32 (1.13)

[a]TIA: transient ischemic attack.

[b]Anticoagulant therapy denotes receiving vitamin K antagonist or nonvitamin K antagonist oral anticoagulant.

[c]LAAO: left atrial appendage occlusion.

[d]Overweight denotes BMI≥25 kg/m^2.

[e]CHA_2DS_2-VASc: congestive heart failure, hypertension, age≥75 (doubled), diabetes mellitus, prior stroke or transient ischemic attack (doubled), vascular disease, age 65-74, female gender.

[f]HAS-BLED: hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile international normalized ratio, elderly, drugs/alcohol concomitantly.

Predicted Drug Adherence to Long-Term Treatment

Predicted drug adherence of patients significantly improved at 4 months (P<.001) (Table 2). More than 90% of the indicated patients received anticoagulant therapy at 4 months, and none of them received antiplatelet therapy unless there was an additional indication.

XSL•FO
RenderX

Table 2. Drug adherence, and amelioration in lifestyle and health behaviors.

Variable	Baseline (N=73)	4 months (N=71)	*P* value
Predicted adherence[a], mean (SD)	6.57 (2.76)	1.45 (1.47)	<.001
Diet, n (%)			
Low-salt, low-fat diet	31 (42)	43 (61)	.04
More fruits or vegetables intake	18 (25)	54 (76)	<.001
Healthy lifestyles, n (%)			
Moderate physical activity[b]	16 (22)	30 (42)	.009
Quitting or reducing alcohol intake	37 (51)	52 (73)	.005
Quitting or reducing smoking	45 (62)	55 (78)	.04
Self-monitoring, n (%)			
Blood pressure	19 (26)	51 (72)	<.001
Heart rate	8 (11)	37 (52)	<.001
Rhythm	5 (7)	34 (48)	<.001

[a]Pharmacy Quality Alliance adherence measures were used to predict possible adherence problems at the following three levels: low risk (0), moderate risk (2-7), and high risk (8+); possible range=0-36.

[b]Moderate physical activity=150 minutes/week of moderate-intensity exercise [32].

Ameliorations in Lifestyles and Healthy Behaviors

The healthy lifestyles and behaviors of patients improved to varying degrees after the 4-month HCFT-AF intervention (Table 2).

Acceptability, Feasibility, and Usability of the HCFT-AF Intervention

The majority of patients provided good feedback on the HCFT-AF intervention, especially related to its usefulness and satisfaction (Table 3). Cronbach α of the satisfaction, ease of use, usefulness, and overall scale was .89, .85, .83, and .90, respectively. The 31 surveyed physicians (10 specialists and 21 general practitioners) also provided a positive appraisal on the program. All specialists agreed that they were liberated from simple primary repetitive work via the hierarchical management program, and could offer more comprehensive and individual care of patients in comparison with traditional outpatient visits. More than 90% (19/21) of the general practitioners claimed that they had gained substantial professional knowledge and experience about atrial fibrillation from this program. They were also willing and able to provide atrial fibrillation patients with standardized guideline-based therapy.

Table 3. Acceptability, feasibility, and usability of patients with the Hospital-Community-Family-Based Telemedicine (HCFT-AF) intervention.

Category[a]	Mean (SD)	Range
Satisfaction	5.21 (1.43)	2.45-7.00
Ease of use	4.76 (1.58)	1.86-7.00
Usefulness	5.45 (1.40)	3.12-7.00
Overall usability of the intervention	5.11 (1.52)	2.68-7.00

[a]Scored on a scale of 1-7; higher scores indicate better satisfaction, easier use, higher effectiveness, and greater overall usability of the intervention.

Engagement of the HCFT-AF Program

Overall, 76% (54/71) of the patients used the software or website more than once a week and accomplished clinic visits as scheduled. In addition, 70% (50/71) of the patients uploaded health-related data periodically during the 4-month study period. It is worth mentioning that more than half (56%, 40/71) of the patients could utilize the mobile app or website by themselves. Learning education materials, communicating with health care physicians, and viewing their own medication prescriptions were the three most frequently used functions.

Discussion

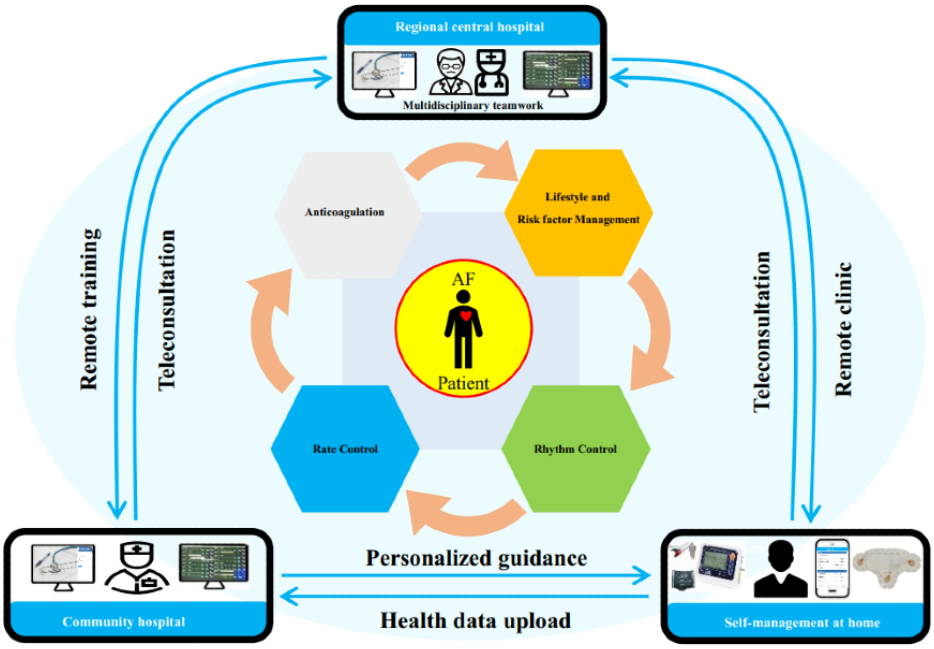
Principal Results

The HCFT-AF program integrates streamlined guideline-based therapy, sustaining education (for patients, their caregivers, and general practitioners), and patient self-management with family involvement via a multifunctional telemedicine platform (Figure 3). This prospective single-arm pre-post design pilot study aimed to assess the feasibility and safety of this program. Preliminary results show that patients and physicians have a high degree of satisfaction and participation. More importantly, the standardized therapy rate, drug adherence, and unhealthy

XSL•FO
RenderX

lifestyles of patients were all ameliorated to varying degrees after a short-term intervention.

Figure 3. The Hospital-Community-Family–based Telemedicine Program to implement integrated atrial fibrillation care.



Comparison With Prior Work

Several large studies have contributed to disclosing the vital epidemiological traits of atrial fibrillation worldwide but also exposed the underuse of guideline-based standardized treatment, especially anticoagulation therapy [1,5,10,11]. Many factors account for this intractable plight. Warfarin's narrow treatment window, multiple drug and dietary interactions, and frequent laboratory monitoring are inherent defects contributing to the poor adherence to anticoagulant therapy [33]. The development of NOACs has overcome these drawbacks, but these drugs are still not covered by health insurance in most countries, including China, with a daily cost that is 160 to 200 times higher than that of warfarin, thereby hindering their long-term availability. In addition to the low drug-use rate, the discontinuation of medication is another problem that plagues clinicians. The RE-LY study reported that the rates of discontinuation at 1 and 2 years were 14.5% and 21.1% for dabigatran, and 10.2% and 16.6% for warfarin, respectively [34]; real-world data are even less optimistic [35].

Periodical clinical follow up is important to guarantee safe and effective anticoagulant therapy [17,36]. However, traditional outpatient visits mostly depend on the abnormality of patients' self-perception. Furthermore, the face-to-face outpatient visit mode prevents doctors from fully communicating with patients and providing effective out-of-hospital management. The HCFT-AF program furnishes patients with dual online and offline consistent follow up. Such facilitated medical visits can help to improve patient compliance with long-term therapy, especially for those with limited mobility or living in rural regions. More than three-quarters (76%, 54/71) of patients accomplished medical visits as scheduled during this 4-month study.

Insufficient awareness of the hazard of atrial fibrillation and excessive worry of bleeding caused by anticoagulation therapy hinder nonspecialists from offering indicated patients with standardized therapy [29,37]. Guideline-based integrated atrial fibrillation management encompasses many aspects such as anticoagulant therapy, reversion to sinus rhythm, and controlling the ventricular rate, which are not easy to master for nonspecialists. However, this can also be simply streamlined in line with the Atrial fibrillation Better Care (ABC) pathway: "A" avoiding stroke with anticoagulants; "B" for better symptom amelioration with symptom-directed treatment via rate or rhythm control; and "C" for cardiovascular and comorbidity risk reduction, comprising lifestyle and risk factors management [38]. Promotion of such a strategy offers an opportunity to improve awareness and diagnosis, while empowering clinicians with straightforward decision-making steps that may align with the therapeutic regimen of generalists and specialists. Integrated management following the ABC pathway could reduce more than two-thirds of the all-cause deaths and half of the composite outcomes (eg, death, major bleeding, ischemic stroke, and myocardial infarction) after an average follow up of 6.2 (SD 3.5) years [39]. Moreover, it can reduce health-related costs significantly [40], demonstrating a clear benefit to optimize the management of patients with atrial fibrillation. However, the best strategy for implementation of such a program is still being explored. A majority of prior studies have been dominated by nurses [17-19] with latent gaps in the better implementation of guidelines [22]. Additionally, these were mostly single-center studies covering a finite region or supported by a single management strategy such as structured telephone follow up [23], web platform [16], or app [24]. The HCFT-AF program, supported by efficient infrastructures, may be an effective carrier for the implementation of integrated atrial fibrillation care.

XSL•FO
RenderX

In addition, perceptions of patients and their caregivers about atrial fibrillation will influence their willingness and ability to obey the treatment recommendations [41]. Several studies have supported that better patient understanding is associated with improved health outcomes [42,43]. As holistic management of atrial fibrillation calls for patients to comply with long-term therapy and transform their unhealthy lifestyles, sometimes without a rapid visible benefit, it is important that they better understand their duties in the treatment. Clinicians are responsible for offering evidence-based therapy, whereas the compliance with the therapy is the responsibility of informed and autonomous patients, which is termed "shared accountability" [44]. A prior study demonstrated that a multifaceted educational intervention significantly improved the proportion of oral anticoagulants utilization, and therefore had the potential to ameliorate the outcome of stroke prevention [45]. The HCFT-AF program provides comprehensive education to patients and their caregivers on the risk factors, treatment, and self-management of atrial fibrillation. The app for patients and intelligent health monitoring devices empowers them to participate in their own disease management, thus aiding in improving adherence to long-term therapies.

Accumulating evidence demonstrates that amelioration in obesity [46], physical fitness [47], and hypertension [48], as well as other risk factors such as alcohol consumption, can reduce the atrial fibrillation burden, often to a degree that is not inferior to that of catheter ablation and other invasive methods [38]. Furthermore, interventions for additional risk factors can strengthen the benefits of disease-oriented therapies such as ablation. A recent scientific statement from the American Heart Association emphasized that lifestyle and risk factor management should be integrated as the fourth pillars in addition to the conventional triangle of atrial fibrillation management [38]. Some previous studies reported that the atrial fibrillation burden and severity can be improved through general lifestyle advice [46], weight intervention [46], or moderate exercise [49]. In general, it is currently well recognized that lifestyle and risk factors modification for atrial fibrillation should be managed as chronic diseases requiring multiple repeated interventions to bring forth long-term successful outcomes. Telemedicine contributing to lifestyle improvement has been confirmed in other domains of medicine [50-56]. On the one hand, telemedicine provides patients and their caregivers with sustaining education on the necessity of lifestyle alterations [50,57]. On the other hand, timely feedback from clinicians enhances patients' motivation to change their unhealthy lifestyles [52]. The HCFT-AF program encourages patients to record and upload their lifestyles through the app, followed by personalized advice from clinicians. Multiple repeated interventions may further motivate patients to change their unhealthy lifestyles.

With the launch of a large-scale atrial fibrillation screening program, the number of atrial fibrillation patients will increase exponentially [58]. Smartphones, computers, and internet-connected devices have become ubiquitous in modern life, which lay a foundation for integrating new methods and novel technologies to provide better care for more patients. A core team (eg, composed of cardiologists) supported by efficient infrastructures serves as an intermediary with other health care specialists, which may achieve optimal management. The overall process comprises a diagnostic assessment, initiation of appropriate guideline-based therapy, sustaining follow up, and education and empowerment of patients and their caregivers in disease self-management. Accordingly, stabilized and sufficiently managed patients can finish subsequent follow up by supported self-management at home or in the community online or offline. If necessary, the patient can be referred to the regional central hospital rapidly. Such integrated management based on telemedicine makes it possible to provide standardized treatment for more patients with limited medical resources.

Limitations

Some limitations of this study need to be discussed. First, given that this was a feasibility study aiming to lay the foundation for subsequent large-scale research, we adopted a convenience sample size without a formal power calculation. Second, previous similar studies have shown that our outcomes of interest showed no significant changes in the control group [24,45,50]; therefore, we conducted a single-arm study to ascertain the feasibility and safety of the HCFT-AF program. Hence, this study may not have been sufficiently powered to determine the relative benefits of the intervention. In terms of lifestyle management, we simply educated patients on the importance of lifestyle amelioration rather than providing them with a detailed weight loss or exercise plan. On the one hand, the improvement of lifestyle via education has been confirmed in several previous studies [45,50]. On the other hand, this mode will broaden the adaptability of our findings. Certainly, there is an additional benefit that some patients will attain from the protocol-driven lifestyle and risk factor management program [53]. Finally, the impact on the major clinical outcomes (eg, stroke, bleeding, death) will be summarized in our ongoing randomized controlled study, which was not the principle purpose of this study.

Conclusions

The findings from this pilot study highlight the major role of the HCFT-AF program in improving the standardized therapy rate, drug adherence, and lifestyles of patients with atrial fibrillation, while further enhancing the guideline adherence of clinicians, thus providing a theoretical basis for eventual clinical benefits. A multidisciplinary team, supported by efficient infrastructures, is conducive to narrow the gap between clinical practice and guidelines. The potential effectiveness of integrated management has already been confirmed; however, more studies are needed to ascertain the best strategy to implement the program, which will allow more patients to benefit from optimal evidence-based therapies.

XSL•FO
RenderX

Acknowledgments

The authors wish to acknowledge all of the clinicians, research assistants, patients, and their caregivers for their participation in the study. This work was funded in part by grants from Jiangsu Provincial Science and Technology Department (project code: BL2013022).

Authors' Contributions

All authors contributed to the study and preparation of this manuscript. JJ and CD were responsible for collecting the data, performing the statistical analysis, and drafting the first version of the manuscript. GX was responsible for confirming the topic, study design, and further polishing of the manuscript. LX, SL, and ZY were responsible for the literature search and editing the manuscript. DY, SL, CK, BY, ZY, and SH were responsible for data collection, figures preparation, final editing, and preparation of the manuscript for submission. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supportive health monitoring devices.
[PNG File , 266 KB - mhealth_v8i10e22137_app1.png]

References

1. Rahman F, Kwan GF, Benjamin EJ. Global epidemiology of atrial fibrillation. Nat Rev Cardiol 2014 Nov;11(11):639-654. [doi: 10.1038/nrcardio.2014.118] [Medline: 25113750]
2. Krijthe BP, Kunst A, Benjamin EJ, Lip GYH, Franco OH, Hofman A, et al. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. Eur Heart J 2013 Sep;34(35):2746-2751 [FREE Full text] [doi: 10.1093/eurheartj/eht280] [Medline: 23900699]
3. Guo Y, Tian Y, Wang H, Si Q, Wang Y, Lip GYH. Prevalence, incidence, and lifetime risk of atrial fibrillation in China: new insights into the global burden of atrial fibrillation. Chest 2015 Jan;147(1):109-119. [doi: 10.1378/chest.14-0321] [Medline: 24921459]
4. Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, et al. Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. Circulation 2004 Aug 31;110(9):1042-1046. [doi: 10.1161/01.CIR.0000140263.20897.42] [Medline: 15313941]
5. Chugh SS, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ, et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. Circulation 2014 Feb 25;129(8):837-847 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.113.005119] [Medline: 24345399]
6. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation: a major contributor to stroke in the elderly. The Framingham Study. Arch Intern Med 1987 Sep;147(9):1561-1564. [Medline: 3632164]
7. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991 Aug;22(8):983-988. [doi: 10.1161/01.str.22.8.983] [Medline: 1866765]
8. López-López JA, Sterne JAC, Thom HHZ, Higgins JPT, Hingorani AD, Okoli GN, et al. Oral anticoagulants for prevention of stroke in atrial fibrillation: systematic review, network meta-analysis, and cost effectiveness analysis. BMJ 2017 Nov 28;359:j5058. [doi: 10.1136/bmj.j5058] [Medline: 29183961]
9. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med 2007 Jun 19;146(12):857-867. [doi: 10.7326/0003-4819-146-12-200706190-00007] [Medline: 17577005]
10. Guo Y, Wang H, Tian Y, Wang Y, Lip GYH. Time Trends of Aspirin and Warfarin Use on Stroke and Bleeding Events in Chinese Patients With New-Onset Atrial Fibrillation. Chest 2015 Jul;148(1):62-72 [FREE Full text] [doi: 10.1378/chest.14-2018] [Medline: 25501045]
11. Mazurek M, Huisman MV, Rothman KJ, Paquette M, Teutsch C, Diener H, GLORIA-AF Investigators. Regional Differences in Antithrombotic Treatment for Atrial Fibrillation: Insights from the GLORIA-AF Phase II Registry. Thromb Haemost 2017 Dec;117(12):2376-2388 [FREE Full text] [doi: 10.1160/TH17-08-0555] [Medline: 29212125]
12. Wei Y, Xu J, Wu H, Zhou G, Chen S, Wang C, et al. Survey of Antithrombotic Treatment in Rural Patients (>60 years) with Atrial Fibrillation in East China. Sci Rep 2018 May 01;8(1):6830. [doi: 10.1038/s41598-018-24878-y] [Medline: 29717155]
13. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur J Cardiothorac Surg 2016 Nov;50(5):e1-e88. [doi: 10.1093/ejcts/ezw313] [Medline: 27663299]

XSL•FO
RenderX

14. NHFA CSANZ Atrial Fibrillation Guideline Working Group, Brieger D, Amerena J, Attia J, Bajorek B, Chan KH, et al. National Heart Foundation of Australia and the Cardiac Society of Australia and New Zealand: Australian Clinical Guidelines for the Diagnosis and Management of Atrial Fibrillation 2018. Heart Lung Circ 2018 Oct;27(10):1209-1266 [FREE Full text] [doi: 10.1016/j.hlc.2018.06.1043] [Medline: 30077228]
15. Lip GYH, Banerjee A, Boriani G, Chiang CE, Fargo R, Freedman B, et al. Antithrombotic Therapy for Atrial Fibrillation: CHEST Guideline and Expert Panel Report. Chest 2018 Nov;154(5):1121-1201. [doi: 10.1016/j.chest.2018.07.040] [Medline: 30144419]
16. Kotecha D, Chua WWL, Fabritz L, Hendriks J, Casadei B, Schotten U, European Society of Cardiology (ESC) Atrial Fibrillation Guidelines Taskforce, the CATCH ME consortiumthe European Heart Rhythm Association (EHRA). European Society of Cardiology smartphone and tablet applications for patients with atrial fibrillation and their health care providers. Europace 2018 Feb 01;20(2):225-233 [FREE Full text] [doi: 10.1093/europace/eux299] [Medline: 29040548]
17. Hendriks JML, Tieleman RG, Vrijhoef HJM, Wijtvliet P, Gallagher C, Prins MH, et al. Integrated specialized atrial fibrillation clinics reduce all-cause mortality: post hoc analysis of a randomized clinical trial. Europace 2019 Dec 01;21(12):1785-1792. [doi: 10.1093/europace/euz209] [Medline: 31390464]
18. Hendriks JML, de Wit R, Crijns HJGM, Vrijhoef HJM, Prins MH, Pisters R, et al. Nurse-led care vs. usual care for patients with atrial fibrillation: results of a randomized trial of integrated chronic care vs. routine clinical care in ambulatory patients with atrial fibrillation. Eur Heart J 2012 Nov;33(21):2692-2699. [doi: 10.1093/eurheartj/ehs071] [Medline: 22453654]
19. Carter L, Gardner M, Magee K, Fearon A, Morgulis I, Doucette S, et al. An Integrated Management Approach to Atrial Fibrillation. J Am Heart Assoc 2016 Jan 25;5(1):e002950. [doi: 10.1161/JAHA.115.002950] [Medline: 26811169]
20. Gallagher C, Elliott AD, Wong CX, Rangnekar G, Middeldorp ME, Mahajan R, et al. Integrated care in atrial fibrillation: a systematic review and meta-analysis. Heart 2017 Dec;103(24):1947-1953. [doi: 10.1136/heartjnl-2016-310952] [Medline: 28490616]
21. Saraswat MK, Carter L, Berrigan P, Sapp JL, Gray C, Fearon A, et al. Integrated Management Approach to Atrial Fibrillation Care: A Cost Utility Analysis. Can J Cardiol 2019 Sep;35(9):1142-1148. [doi: 10.1016/j.cjca.2019.04.016] [Medline: 31472812]
22. McKee G, Kerins M, Hamilton G, Hansen T, Hendriks J, Kletsiou E, et al. Barriers to ESC guideline implementation: results of a survey from the European Council on Cardiovascular Nursing and Allied Professions (CCNAP). Eur J Cardiovasc Nurs 2017 Dec;16(8):678-686. [doi: 10.1177/1474515117710097] [Medline: 28498092]
23. Li PWC, Yu DSF, Yan BBY. A nurse-coordinated integrated care model to support decision-making and self-care in patients with atrial fibrillation: A study protocol. J Adv Nurs 2019 Dec;75(12):3749-3757. [doi: 10.1111/jan.14164] [Medline: 31350778]
24. Guo Y, Chen Y, Lane DA, Liu L, Wang Y, Lip GYH. Mobile Health Technology for Atrial Fibrillation Management Integrating Decision Support, Education, and Patient Involvement: mAF App Trial. Am J Med 2017 Dec;130(12):1388-1396.e6 [FREE Full text] [doi: 10.1016/j.amjmed.2017.07.003] [Medline: 28847546]
25. Gu X, Zhu Y, Zhang Y, Sun L, Bao Z, Shen J, et al. Effect of telehealth interventions on major cardiovascular outcomes: a meta-analysis of randomized controlled trials. J Geriatr Cardiol 2017 Aug;14(8):501-508 [FREE Full text] [doi: 10.11909/j.issn.1671-5411.2017.08.013] [Medline: 29089966]
26. Guo X, Gu X, Jiang J, Li H, Duan R, Zhang Y, et al. A Hospital-Community-Family-Based Telehealth Program for Patients With Chronic Heart Failure: Single-Arm, Prospective Feasibility Study. JMIR Mhealth Uhealth 2019 Dec 13;7(12):e13229 [FREE Full text] [doi: 10.2196/13229] [Medline: 31833835]
27. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
28. International chronic cardiovascular disease management platform. URL: http://www.iccvd.com [accessed 2013-07-01]
29. Fanaroff AC, Ohman EM. Non-Vitamin K Antagonist Oral Anticoagulants in the Treatment of Atrial Fibrillation. Annu Rev Med 2019 Jan 27;70:61-75. [doi: 10.1146/annurev-med-042617-092334] [Medline: 30477393]
30. McHorney CA. The Adherence Estimator: a brief, proximal screener for patient propensity to adhere to prescription medications for chronic disease. Curr Med Res Opin 2009 Jan;25(1):215-238. [doi: 10.1185/03007990802619425] [Medline: 19210154]
31. Nahm E, Resnick B, Mills ME. Development and pilot-testing of the perceived health Web Site usability questionnaire (PHWSUQ) for older adults. Stud Health Technol Inform 2006;122:38-43. [Medline: 17102214]
32. Kraus WE, Powell KE, Haskell WL, Janz KF, Campbell WW, Jakicic JM, 2018 Physical Activity Guidelines Advisory Committee*. Physical Activity, All-Cause and Cardiovascular Mortality, and Cardiovascular Disease. Med Sci Sports Exerc 2019 Jun;51(6):1270-1281 [FREE Full text] [doi: 10.1249/MSS.0000000000001939] [Medline: 31095084]
33. Vinogradova Y, Coupland C, Hill T, Hippisley-Cox J. Risks and benefits of direct oral anticoagulants versus warfarin in a real world setting: cohort study in primary care. BMJ 2018 Jul 04;362:k2505. [doi: 10.1136/bmj.k2505] [Medline: 29973392]

XSL•FO
RenderX

34. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, RE-LY Steering CommitteeInvestigators. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med 2009 Sep 17;361(12):1139-1151. [doi: 10.1056/NEJMoa0905561] [Medline: 19717844]
35. Salmasi S, Loewen PS, Tandun R, Andrade JG, De Vera MA. Adherence to oral anticoagulants among patients with atrial fibrillation: a systematic review and meta-analysis of observational studies. BMJ Open 2020 Apr 08;10(4):e034778. [doi: 10.1136/bmjopen-2019-034778] [Medline: 32273316]
36. Vinereanu D, Lopes RD, Bahit MC, Xavier D, Jiang J, Al-Khalidi HR, IMPACT-AF investigators. A multifaceted intervention to improve treatment with oral anticoagulants in atrial fibrillation (IMPACT-AF): an international, cluster-randomised trial. Lancet 2017 Oct 14;390(10104):1737-1746. [doi: 10.1016/S0140-6736(17)32165-7] [Medline: 28859942]
37. Raparelli V, Proietti M, Cangemi R, Lip GYH, Lane DA, Basili S. Adherence to oral anticoagulant therapy in patients with atrial fibrillation. Focus on non-vitamin K antagonist oral anticoagulants. Thromb Haemost 2017 Jan 26;117(2):209-218. [doi: 10.1160/TH16-10-0757] [Medline: 27831592]
38. Chung MK, Eckhardt LL, Chen LY, Ahmed HM, Gopinathannair R, Joglar JA, American Heart Association ElectrocardiographyArrhythmias CommitteeExercise, Cardiac Rehabilitation,Secondary Prevention Committee of the Council on Clinical Cardiology; Council on Arteriosclerosis, ThrombosisVascular Biology; Council on CardiovascularStroke Nursing;Council on LifestyleCardiometabolic Health. Lifestyle and Risk Factor Modification for Reduction of Atrial Fibrillation: A Scientific Statement From the American Heart Association. Circulation 2020 Apr 21;141(16):e750-e772. [doi: 10.1161/CIR.0000000000000748] [Medline: 32148086]
39. Yoon M, Yang P, Jang E, Yu HT, Kim T, Uhm J, et al. Improved Population-Based Clinical Outcomes of Patients with Atrial Fibrillation by Compliance with the Simple ABC (Atrial Fibrillation Better Care) Pathway for Integrated Care Management: A Nationwide Cohort Study. Thromb Haemost 2019 Oct;119(10):1695-1703. [doi: 10.1055/s-0039-1693516] [Medline: 31266082]
40. Pastori D, Farcomeni A, Pignatelli P, Violi F, Lip GY. ABC (Atrial fibrillation Better Care) Pathway and Healthcare Costs in Atrial Fibrillation: The ATHERO-AF Study. Am J Med 2019 Jul;132(7):856-861. [doi: 10.1016/j.amjmed.2019.01.003] [Medline: 30659810]
41. Clarkesmith DE, Lip GYH, Lane DA. Patients' experiences of atrial fibrillation and non-vitamin K antagonist oral anticoagulants (NOACs), and their educational needs: A qualitative study. Thromb Res 2017 May;153:19-27. [doi: 10.1016/j.thromres.2017.03.002] [Medline: 28314139]
42. Asnani MR, Quimby KR, Bennett NR, Francis DK. Interventions for patients and caregivers to improve knowledge of sickle cell disease and recognition of its related complications. Cochrane Database Syst Rev 2016 Oct 06;10:CD011175 [FREE Full text] [doi: 10.1002/14651858.CD011175.pub2] [Medline: 27711980]
43. Oddone EZ, Gierisch JM, Sanders LL, Fagerlin A, Sparks J, McCant F, et al. A Coaching by Telephone Intervention on Engaging Patients to Address Modifiable Cardiovascular Risk Factors: a Randomized Controlled Trial. J Gen Intern Med 2018 Sep;33(9):1487-1494. [doi: 10.1007/s11606-018-4398-6] [Medline: 29736750]
44. Peterson ED, Ho PM, Barton M, Beam C, Burgess LH, Casey DE, et al. ACC/AHA/AACVPR/AAFP/ANA concepts for clinician-patient shared accountability in performance measures: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. Circulation 2014 Nov 25;130(22):1984-1994 [FREE Full text] [doi: 10.1161/CIR.0000000000000139] [Medline: 25366994]
45. Vinereanu D, Lopes RD, Bahit MC, Xavier D, Jiang J, Al-Khalidi HR, IMPACT-AF investigators. A multifaceted intervention to improve treatment with oral anticoagulants in atrial fibrillation (IMPACT-AF): an international, cluster-randomised trial. Lancet 2017 Oct 14;390(10104):1737-1746. [doi: 10.1016/S0140-6736(17)32165-7] [Medline: 28859942]
46. Abed HS, Wittert GA, Leong DP, Shirazi MG, Bahrami B, Middeldorp ME, et al. Effect of weight reduction and cardiometabolic risk factor management on symptom burden and severity in patients with atrial fibrillation: a randomized clinical trial. JAMA 2013 Nov 20;310(19):2050-2060. [doi: 10.1001/jama.2013.280521] [Medline: 24240932]
47. Pathak RK, Elliott A, Middeldorp ME, Meredith M, Mehta AB, Mahajan R, et al. Impact of CARDIOrespiratory FITness on Arrhythmia Recurrence in Obese Individuals With Atrial Fibrillation: The CARDIO-FIT Study. J Am Coll Cardiol 2015 Sep 01;66(9):985-996 [FREE Full text] [doi: 10.1016/j.jacc.2015.06.488] [Medline: 26113406]
48. Huxley RR, Lopez FL, Folsom AR, Agarwal SK, Loehr LR, Soliman EZ, et al. Absolute and attributable risks of atrial fibrillation in relation to optimal and borderline risk factors: the Atherosclerosis Risk in Communities (ARIC) study. Circulation 2011 Apr 12;123(14):1501-1508 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.110.009035] [Medline: 21444879]
49. Malmo V, Nes BM, Amundsen BH, Tjonna A, Stoylen A, Rossvoll O, et al. Aerobic Interval Training Reduces the Burden of Atrial Fibrillation in the Short Term: A Randomized Trial. Circulation 2016 Feb 02;133(5):466-473. [doi: 10.1161/CIRCULATIONAHA.115.018220] [Medline: 26733609]
50. Rubinstein A, Miranda JJ, Beratarrechea A, Diez-Canseco F, Kanter R, Gutierrez L, GISMAL group. Effectiveness of an mHealth intervention to improve the cardiometabolic profile of people with prehypertension in low-resource urban settings in Latin America: a randomised controlled trial. Lancet Diabetes Endocrinol 2016 Jan;4(1):52-63. [doi: 10.1016/S2213-8587(15)00381-2] [Medline: 26653067]

XSL•FO
RenderX

51. Meeker D, Goldberg J, Kim KK, Peneva D, Campos HDO, Maclean R, et al. Patient Commitment to Health (PACT-Health) in the Heart Failure Population: A Focus Group Study of an Active Communication Framework for Patient-Centered Health Behavior Change. J Med Internet Res 2019 Aug 06;21(8):e12483 [FREE Full text] [doi: 10.2196/12483] [Medline: 31389339]
52. Joiner KL, Nam S, Whittemore R. Lifestyle interventions based on the diabetes prevention program delivered via eHealth: A systematic review and meta-analysis. Prev Med 2017 Jul;100:194-207 [FREE Full text] [doi: 10.1016/j.ypmed.2017.04.033] [Medline: 28456513]
53. Kliemann N, Croker H, Johnson F, Beeken RJ. Development of the Top Tips Habit-Based Weight Loss App and Preliminary Indications of Its Usage, Effectiveness, and Acceptability: Mixed-Methods Pilot Study. JMIR Mhealth Uhealth 2019 May 10;7(5):e12326 [FREE Full text] [doi: 10.2196/12326] [Medline: 31094352]
54. Weerahandi H, Paul S, Quintiliani LM, Chokshi S, Mann DM. A Mobile Health Coaching Intervention for Controlling Hypertension: Single-Arm Pilot Pre-Post Study. JMIR Form Res 2020 May 07;4(5):e13989 [FREE Full text] [doi: 10.2196/13989] [Medline: 32379049]
55. Peiris D, Wright L, News M, Rogers K, Redfern J, Chow C, et al. A Smartphone App to Assist Smoking Cessation Among Aboriginal Australians: Findings From a Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Apr 02;7(4):e12745 [FREE Full text] [doi: 10.2196/12745] [Medline: 30938691]
56. Simeon R, Dewidar O, Trawin J, Duench S, Manson H, Pardo Pardo J, et al. Behavior Change Techniques Included in Reports of Social Media Interventions for Promoting Health Behaviors in Adults: Content Analysis Within a Systematic Review. J Med Internet Res 2020 Jun 11;22(6):e16002 [FREE Full text] [doi: 10.2196/16002] [Medline: 32525482]
57. Ding H, Jayasena R, Chen SH, Maiorana A, Dowling A, Layland J, et al. The Effects of Telemonitoring on Patient Compliance With Self-Management Recommendations and Outcomes of the Innovative Telemonitoring Enhanced Care Program for Chronic Heart Failure: Randomized Controlled Trial. J Med Internet Res 2020 Jul 08;22(7):e17559 [FREE Full text] [doi: 10.2196/17559] [Medline: 32673222]
58. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, Apple Heart Study Investigators. Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. N Engl J Med 2019 Nov 14;381(20):1909-1917. [doi: 10.1056/NEJMoa1901183] [Medline: 31722151]

Abbreviations

ABC: Atrial fibrillation Better Care
HCFT-AF: Hospital-Community-Family-Based Telemedicine
NOAC: non-vitamin K antagonist oral anticoagulant

Edited by G Eysenbach; submitted 05.07.20; peer-reviewed by H Mehdizadeh, D Agarwal; comments to author 16.07.20; revised version received 22.07.20; accepted 26.07.20; published 21.10.20.

Please cite as:
Jiang J, Gu X, Cheng CD, Li HX, Sun XL, Duan RY, Zhu Y, Sun L, Chen FK, Bao ZY, Zhang Y, Shen JH
The Hospital-Community-Family–Based Telemedicine (HCFT-AF) Program for Integrative Management of Patients With Atrial Fibrillation: Pilot Feasibility Study
JMIR Mhealth Uhealth 2020;8(10):e22137
URL: http://mhealth.jmir.org/2020/10/e22137/
doi:10.2196/22137
PMID:33084588

©Jiang Jiang, Xiang Gu, Chen-Di Cheng, Hong-Xiao Li, Xiao-Lin Sun, Ruo-Yu Duan, Ye Zhu, Lei Sun, Fu-Kun Chen, Zheng-Yu Bao, Yi Zhang, Jian-Hua Shen. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 21.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Carbohydrate Counting App Using Image Recognition for Youth With Type 1 Diabetes: Pilot Randomized Control Trial

Jeffrey E Alfonsi[1,2], BEng, MD, FRCPC; Elizabeth E Y Choi[2], BEng; Taha Arshad[3], BSc; Stacie-Ann S Sammott[3], BA; Vanita Pais[3], BASc, RD, CDE; Cynthia Nguyen[4], MPH; Bryan R Maguire[5], MSc; Jennifer N Stinson[4,6], PhD; Mark R Palmert[3,7], MD, PhD

[1]Department of Medicine, Schulich School of Medicine & Dentistry, Western University, London, ON, Canada

[2]Inner Analytics Inc, Toronto, ON, Canada

[3]Division of Endocrinology, Hospital for Sick Children, Toronto, ON, Canada

[4]Peter Gilgan Centre for Research and Learning, The Hospital for Sick Children, Toronto, ON, Canada

[5]Biostatistics Design and Analysis Unit, SickKids Research Institute, Toronto, ON, Canada

[6]Lawrence S Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

[7]Departments of Pediatrics and Physiology, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Jeffrey E Alfonsi, BEng, MD, FRCPC
Department of Medicine
Schulich School of Medicine & Dentistry
Western University
1151 Richmond St
Clinical Skills Building
London, ON, N6A 3K7
Canada
Phone: 1 647 938 3669
Email: j.alfonsi@utoronto.ca

Abstract

Background: Carbohydrate counting is an important component of diabetes management, but it is challenging, often performed inaccurately, and can be a barrier to optimal diabetes management. iSpy is a novel mobile app that leverages machine learning to allow food identification through images and that was designed to assist youth with type 1 diabetes in counting carbohydrates.

Objective: Our objective was to test the app's usability and potential impact on carbohydrate counting accuracy.

Methods: Iterative usability testing (3 cycles) was conducted involving a total of 16 individuals aged 8.5-17.0 years with type 1 diabetes. Participants were provided a mobile device and asked to complete tasks using iSpy app features while thinking aloud. Errors were noted, acceptability was assessed, and refinement and retesting were performed across cycles. Subsequently, iSpy was evaluated in a pilot randomized controlled trial with 22 iSpy users and 22 usual care controls aged 10-17 years. Primary outcome was change in carbohydrate counting ability over 3 months. Secondary outcomes included levels of engagement and acceptability. Change in HbA_{1c} level was also assessed.

Results: Use of iSpy was associated with improved carbohydrate counting accuracy (total grams per meal, P=.008), reduced frequency of individual counting errors greater than 10 g (P=.047), and lower HbA_{1c} levels (P=.03). Qualitative interviews and acceptability scale scores were positive. No major technical challenges were identified. Moreover, 43% (9/21) of iSpy participants were still engaged, with usage at least once every 2 weeks, at the end of the study.

Conclusions: Our results provide evidence of efficacy and high acceptability of a novel carbohydrate counting app, supporting the advancement of digital health apps for diabetes care among youth with type 1 diabetes. Further testing is needed, but iSpy may be a useful adjunct to traditional diabetes management.

Trial Registration: ClinicalTrials.gov NCT04354142; https://clinicaltrials.gov/ct2/show/NCT04354142

(JMIR Mhealth Uhealth 2020;8(10):e22074) doi:10.2196/22074

XSL•FO
RenderX

KEYWORDS

carbohydrate counting; type 1 diabetes; image recognition; youth; digital health applications (apps); mHealth

Introduction

Type 1 diabetes is among the most common chronic diseases of childhood, and its incidence is rising [1]. The management of diabetes in youth is complex and impacted by numerous factors including numeracy skills, education, socioeconomic status, family dynamics, engagement with treatment regimens, and use of technologies such as pumps and continuous sensors. Among these factors, insulin administration remains the cornerstone of type 1 diabetes management, but its optimal dosing is often complicated by the need to count carbohydrates [2-4]. Carbohydrate counting allows individuals with type 1 diabetes to match their insulin doses to planned food consumption, and accurate carbohydrate counting can improve blood glucose control (measured by hemoglobin A_{1c}; HbA_{1c}) [2]. For example, in one study [5] focused on parents of children with type 1 diabetes, more accurate parental carbohydrate counting was associated with 0.8% lower HbA_{1c} values in their children. Among adults with type 1 diabetes, a meta-analysis [6] of 5 studies showed that HbA_{1c} levels improved by an average of 0.6% with improved carbohydrate counting.

Despite its importance, up to two-thirds of individuals with diabetes report having trouble with carbohydrate counting [7]. It has been reported that only a quarter of youth can routinely count carbohydrates within 10 g of the true net carbohydrate value, even for commonly eaten foods [8] and that carbohydrate counting is a barrier to diabetes management [9]. Carbohydrate counting often requires multiple training sessions with experienced dietitians or educators and ongoing efforts by patients and families to maintain competency. Estimating carbohydrate intake can be difficult when the portions of food being consumed are not the same as those listed in an exchange system or on the food label, requiring youth to adjust the carbohydrate count to the appropriate portion size. Accuracy of carbohydrate counting can be further limited by low nutritional literacy and poor numeracy skills [10].

Technologies such as mobile health apps that address these barriers have the potential to ease burden and improve blood glucose control. Unfortunately, most diabetes-related mobile health apps have not undergone formal evaluation and lack evidence of clinical effectiveness, making it difficult for prospective users to assess the value of a particular app to their self-management [11], a situation that is also true within the domain of carbohydrate counting apps.

To help address these gaps, a mobile app was designed and developed to assist youth in counting carbohydrates. The app (iSpy) uses image recognition and artificial intelligence to identify foods and report their carbohydrate content. Here we addressed the question of how iSpy would perform during usability and pilot testing. We hypothesized that it would be well-accepted and that its use would be associated with improved carbohydrate counting accuracy.

Methods

Description of iSpy

The iSpy app (see Multimedia Appendix 1 for screenshots) was developed and evaluated in sequential phases [12,13]. The image recognition algorithm that identifies foods from images uses a convolutional neural network. The interface for iSpy was initially co-designed with intended users including certified diabetes educators, registered dietitians, and individuals (aged 12-75 years) living with type 1 diabetes. Once developed, cycles of refinement were conducted after assessing how users functionally navigated the app. The app was then tested for accuracy on a sample of 200 commonly consumed food items (169 items from the Youth Adolescent Food Frequency Questionnaire [14] and 31 commonly consumed complex items containing 2 or more components) selected by a registered dietitian and diabetes educator. An accuracy test required iSpy to report a carbohydrate content that was within 10 g of the food item's true net (total minus fiber) carbohydrate value [15]. Revisions were made until iSpy was able to achieve this degree of accuracy for ≥90% (180/200) of the items. Current overall accuracy is 94.5% (189/200). The app was then moved into clinical testing, described herein.

Setting

The usability testing and pilot randomized controlled trial were approved by the research ethics board and conducted within the diabetes program at The Hospital for Sick Children. Informed consent was obtained from all participants, and the pilot randomized controlled trial was registered with clinicaltrials.gov (NCT04354142).

Usability Testing Procedures

Inclusion criteria were (1) age 8.0-18.0 years, (2) a diagnosis of type 1 diabetes per Diabetes Canada guidelines [16], (3) use of carbohydrate counting as part of treatment regimen, and (4) fluency in English (iSpy is only available in English). The sole exclusion criterion was cognitive impairments.

Iterative cycles of testing and app refinement (3 cycles) were utilized. Testing consisted of 4 scenario-based tasks that were developed using standardized guidelines [17], a semistructured interview, and app acceptability measured by the 5-point Acceptability E-Scale [13]. The focus for the task was on user performance (ie, ease of use, navigation among screens, functions, errors, and efficiency); the semistructured interview and Acceptability E-Scale were focused on overall satisfaction with the app. Participants were purposively selected to achieve a range of age, gender, and duration of type 1 diabetes. The participants were asked to think aloud during use of the app and dialog was audiorecorded.

Participants were provided with an Android or iOS mobile device, depending on the participant's preference. Scenario-based tasks included use of app features such as photo taking, portion sizing, and food identification. Errors, efficiency (time taken to complete a task), acceptability (ease of use), and

XSL•FO
RenderX

suggestions for improvements were logged, and tasks were classified into 1 of 3 categories (successfully completed, completed with minor issues, incomplete due to usability issues). Following each cycle, refinements were made to the user interface based on problems and recommendations, with the revised interface being evaluated in the subsequent cycle [13]. iSpy was moved to pilot testing (pilot randomized controlled trial) when no further issues were identified in the third cycle.

Pilot Randomized Controlled Trial Procedures

Inclusion criteria were (1) age 10 years-17.0 years (adjusted after usability testing because those under 10 years of age had difficulty navigating the app), (2) ≥6 months since diagnosis with type 1 diabetes, (3) completion of initial carbohydrate counting classes, (4) incorporation of carbohydrate counting into treatment regimen, and (5) access to a smartphone and data plan. Exclusion criteria were (1) cognitive impairment, (2) comorbid physical or psychiatric conditions that might impact ability to use iSpy, (3) diagnosis of a condition that affects dietary exposure, and (4) participation in usability testing.

A convenience sample was enrolled (n=46) and randomly assigned to either usual care (control) or usual care and iSpy (intervention) group using a 2-group randomized block design in blocks of 4 and 6, where the block sizes were not known to the investigator. The randomization schedule was created using SAS (version 9.4; SAS Institute). Data from previous work in our clinic [18] was used to estimate the sample size, indicating that 20 participants per group would be sufficient to detect a mean accuracy difference of 7.1 g in carbohydrate counting (which fit with our aim of assessing accuracy within 10 g), assuming 80% power (β=.2), α=.05, and using a 2-sided paired *t* test; therefore, 23 participants were recruited per group to allow for potential dropout over the 3-month trial.

Duration of diabetes (time since diagnosis) and HbA_{1c} levels were obtained from chart review. Accuracy and efficiency (time taken) of carbohydrate counting were based on a performance task. Participants counted carbohydrates for 10 foods (consisting of 2 foods from each of the 4 main food groups—vegetables and fruit, grain products, milk and alternatives, and meat and alternatives—and in addition, desserts). In each of the 5 food groups, a simple food item (eg, a single item such as an apple) as well as a complex food item (eg, an item containing 2 or more components but with the base food from the selected food group, such as pasta with tomato sauce) were included. Two sets of foods (Diet A and Diet B) of similar difficulty were utilized with half of the participants in each group counting foods from Diet A at baseline and foods from Diet B at 3 months, and vice versa for the other half. This methodology allowed us to control for confounding from participants educating themselves on test items or from any unanticipated differences between test diets. The net carbohydrate value for each food item was determined by either the nutrition label for packaged foods, the United States Department of Agriculture's National Nutrient Database for Standard Reference [19], the Canadian Nutrient File [20], or by our dietitian (VP) who specializes in diabetes care. We chose to utilize the performance task metric to assess carbohydrate counting instead of using tools such as the PedCarbQuiz [21] so that we could assess the effect of iSpy on counting the carbohydrate content of foods as opposed to its effect on domains such as nutrition label reading or insulin dosing, which are part of the PedCarbQuiz.

Additional measures were also collected. At baseline, comfort with technology was assessed. Quality of life, measured by a subset of questions from quality of life for youth [22,23]; self-care, measured by a subset of questions from the Self Care Inventory [24-26]; and patient or parent responsibility, measured by a subset of questions from Diabetes Family Responsibility Questionnaire were also assessed at baseline and 3 months postintervention. We also assessed factors related to usability of the app including fidelity (tracking of technical difficulties, errors within the app); levels of engagement; and acceptability using a 7-item Acceptability E-Scale (5-point scale) [27]. Qualitative feedback was obtained via postintervention, semistructured interviews among all iSpy users.

At the start of the study, participants in the intervention group downloaded the app on their phone, and a demonstration of iSpy and its functionality were provided. iSpy participants were instructed to use the app at their discretion and when they thought its use would be beneficial. We recognized, for example, that participants may know the carbohydrate counts of the food items that they regularly consume. Thus, they may only want to use iSpy occasionally to assess the counts of only some of these food items whereas they may want to use the app more frequently for food items that they do not regularly consume. Given these instructions instead of a recommended number of uses per day, engagement levels were assessed based on frequency of using the app to log foods per week categorized as high (logging ≥2 meals per week), medium (logging ≥1 meal every 2 weeks but <2 times per week), or low (logging <1 meal every 2 weeks). This structure is similar to that used by others to assess app use [28]. In other instructions, iSpy participants were asked to contact the team should they encounter technical difficulties, and they received a phone call 6 weeks postbaseline for general troubleshooting. As this was a pilot study, we strove to encourage the use of iSpy by sending a maximum of 3 automated alerts to participants not accessing iSpy at least once every 2 weeks.

Statistical analysis for the pilot randomized controlled trial was conducted using R (version 3.6.0) statistical software. Descriptive statistics of participant characteristics for the intervention and control groups are presented as means and standard deviations for continuous variables, and counts and proportions for categorical variables. Differences in these characteristics between the intervention and control groups were tested using 2-sided independent *t* tests for continuous variables, and chi-square tests for categorical variables.

Differences between the intervention and control group on the primary outcome variables (accuracy, time taken for counting, and the percentage of food items for which participants estimated the carbohydrate content within 10 g of the true net carbohydrate value), secondary outcomes (quality of life for youth, self-care, and patient or parent responsibility), and HbA_{1c} level at baseline were examined using 2-sided independent *t* tests. Differences in these variables between the intervention and control group at the follow-up visit were assessed using

XSL•FO
RenderX

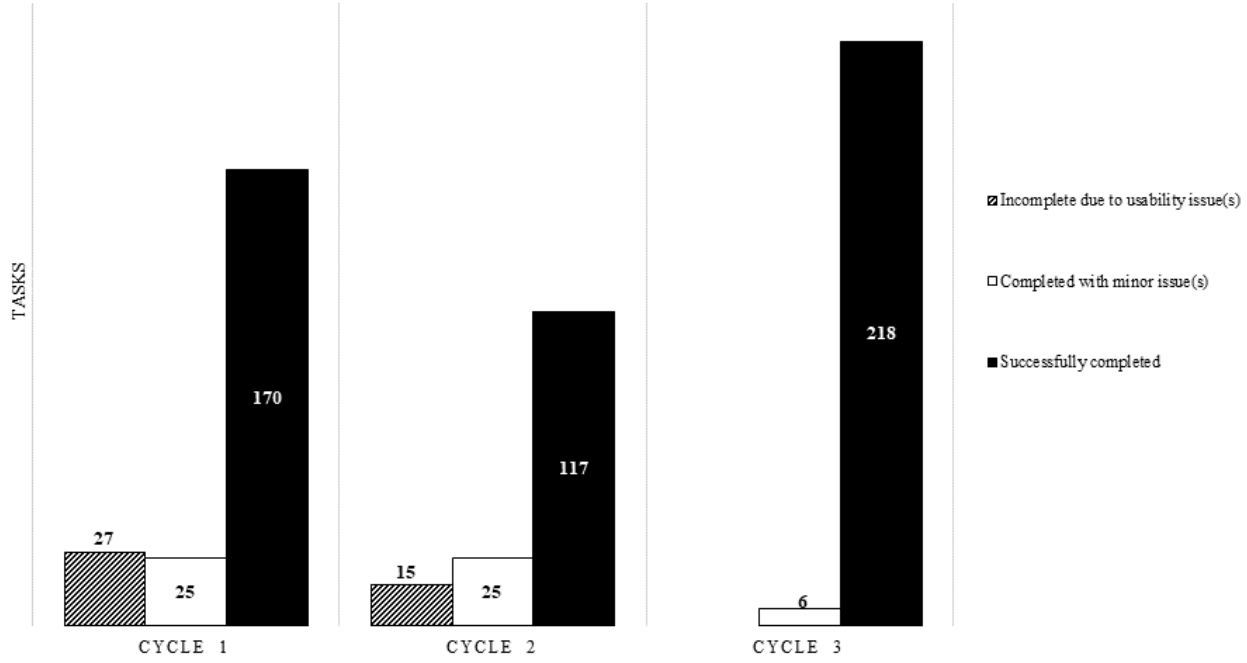
multiple linear regression models, which included the baseline as a covariate. *P* values <.05 were considered to be statistically significant.

Results

Usability Testing

Youth (total: n=16—cycle 1: n=6; cycle 2: n=4, cycle 3: n=6) ranging in age from 8.5 to 17.0 years (mean 13.5, SD 2.6 years) participated in iSpy's iterative usability testing. Scenarios consisting of multiple tasks were used; based on how the participant responded to iSpy or how image recognition classified the food within each scenario, follow-up tasks were required, with the total number of tasks across 4 scenarios varying between 35 and 41 per participant. Errors within each cycle were tracked (Figure 1). In cycle 1, a total of 27 errors preventing successful completion of tasks occurred (mean 4.5 SD 4.4 per participant), representing 12.2% (27/222) of the total tasks. In response, modifications were made such as simplifying the user interface and changing wording so that the app flow was more intuitive. In cycle 2, errors were made on 9.6% of the tasks (15/157). Additional changes were made to the app including simplifying input requirements, making only one action possible at a time, improving graphics, and clarifying instructions. In cycle 3, no errors (0/224, 0%) preventing task completion occurred, and only 2.7% of tasks (6/224) had minor incidents. Acceptability E-Scale scores were positive (mean 4.6, SD 0.7) on domains that included helpfulness in carbohydrate counting and food identification, ease of use, time taken, and overall satisfaction across all 3 cycles of testing. Postcycle 3, minor modifications such as aesthetic changes to the user interface were made prior to pilot testing (pilot randomized controlled trial).

Figure 1. Usability testing errors per cycle representing tasks completed during each cycle of usability testing. The total number of tasks varied per cycle (cycle 1: 222; cycle 2: 157; cycle 3: 224).



Pilot Randomized Controlled Trial

Of the 46 participants who were enrolled and randomly allocated into the 2 arms of the pilot study, 43 participants completed the study (Figure 2). All participants reported being comfortable using computers and smart devices, and there were no significant differences between the 2 groups for any of the baseline characteristics (gender: *P*=.22; age: *P*=.99; duration since diagnosis: *P*=.79; regulation method: *P*=.62; confidence in counting: *P*=.39; Table 1).

At baseline, there was also no difference in carbohydrate counting accuracy or time taken to complete the task between the 2 groups. At the 3-month follow-up visit, the iSpy group displayed a statistically significant increase in carbohydrate counting accuracy (*P*=.008), and a statistically significant decrease in counting errors (*P*=.047) compared to that of the control group (Table 2). None of the secondary outcome variables such as quality of life measures (*P*=.64), self-care measures (*P*=.17), or patient/parent responsibility (*P*=.69), differed between the groups at baseline and 3-month postintervention period. Although not a main outcome variable for this pilot study, HbA_{1c} values were assessed at baseline and at the 3-month follow-up visit, with the iSpy group displaying statistically significant lower HbA_{1c} values (*P*=.03) compared to those of the control group.

XSL•FO
RenderX

Figure 2. Participant flowchart.

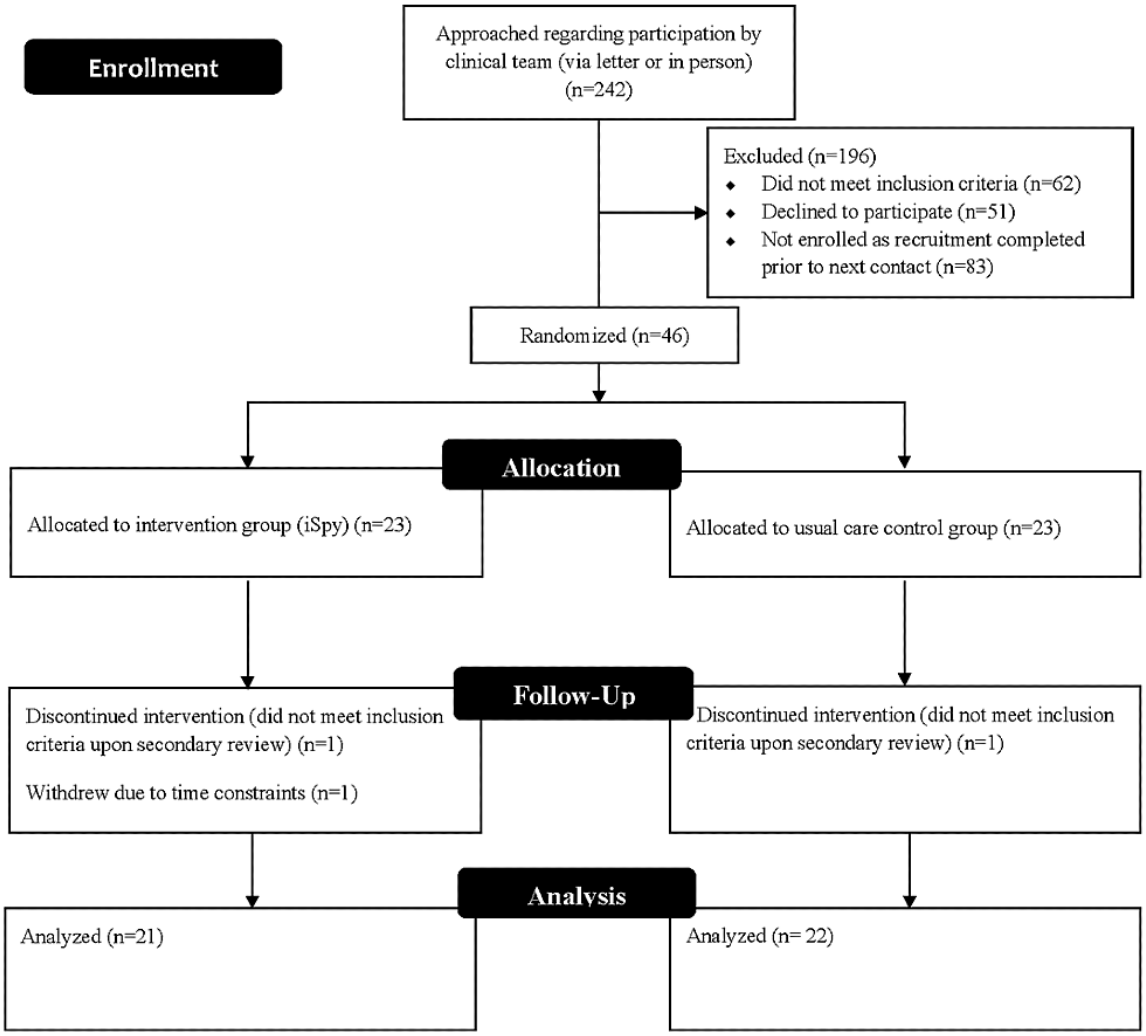


Table 1. Participant characteristics (at baseline).

Characteristics	iSpy intervention (n=22)	Usual care (n=22)	*P* value
Gender, n (%)			.22
Male	11 (50)	16 (73)	
Female	11 (50)	6 (27)	
Age (in years), mean (SD)	13.98 (1.57)	13.98 (1.76)	.99
Duration since diagnosis (in years), mean (SD)	6.08 (4.14)	6.44 (4.45)	.79
Method of insulin regulation, n (%)			.62
Pump	11 (50)	10 (46)	
Other	11 (50)	12 (54)	
Rated confidence in carbohydrate counting skills (out of 10), mean (SD)	6.70 (1.59)	7.24 (2.30)	.39

XSL•FO
RenderX

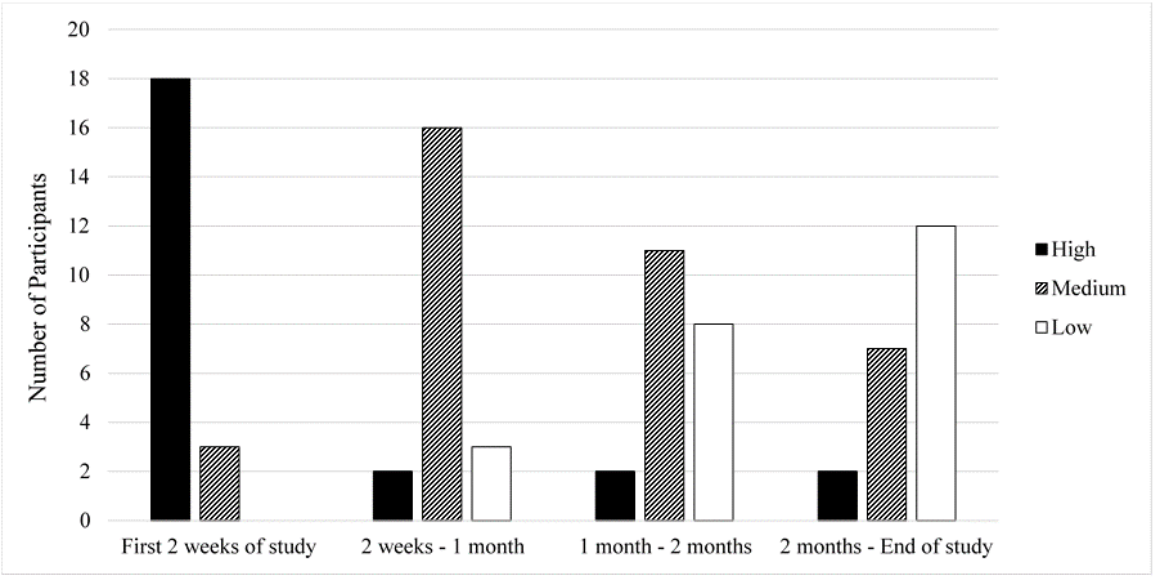
Table 2. Carbohydrate counting and glycemic control outcomes (at baseline and follow-up).

Outcomes	Baseline			Follow-up			Coefficient
	iSpy (n=22)	Usual care (n=22)	*P* value	iSpy (n=21)	Usual care (n=22)	*P* value	
Absolute error (% of total grams), mean (SD)	31.97 (11.36)	32.03 (10.01)	.99	27.45 (10.90)	38.00 (14.74)	.008	–10.479
Errors >10 g (%), mean (SD)	25.00 (14.06)	27.73 (15.10)	.54	21.43 (16.82)	32.27 (16.31)	.047	–9.851
Total time (seconds), mean (SD)	79.95 (23.88)	80.73 (27.82)	.92	74.86 (31.78)	78.23 (44.97)	.79	–2.741
HbA_{1c}[a] (%), mean (SD)	8.41 (1.84)	8.35 (1.32)	.91	8.06 (1.43)	8.80 (1.60)	.03	–0.603

[a]HbA_{1c}: hemoglobin A_{1c}.

No major technical challenges were identified. App engagement was assessed over 4 time periods (first 2 weeks of study, 2 weeks to 1 month, first to second month, and second month to end of study), with 43% (9/21) of participants indicating medium or high use at the end of study (Figure 3). Over the course of the study, a mean of 1.9 (SD 0.94) reminder emails were sent to iSpy users.

Figure 3. Engagement levels with the number of participants in each category displayed for each time frame.



Acceptability E-Scale results were positive (Multimedia Appendix 2). Of the 7 questions asked, iSpy respondents ranked iSpy positively in 6 out of 7 categories. The highest rankings were related to ease of understanding and ease of use. The weakest responses were related to how helpful iSpy was in food identification.

Semistructured interviews were conducted among all iSpy participants, and results mirrored those of the questionnaires. Most participants (18/21, 86%) found the iSpy app fairly to very easy to use. Participants preferred using the photos followed by text features to identify foods with few respondents utilizing the voice function. Participants also valued speed of image recognition results and delays due to misidentification of foods were viewed negatively. Participants provided suggestions for improvement, such as including additional options for portion sizing, refining the identification features so that foods within a complex meal do not need to be added one-by-one, expanding the database of known foods, and including optional reminder notifications about logging foods.

Discussion

Principal Results

With few exceptions [28,29], apps used to facilitate diabetes care have not undergone formal testing [11,30] making it difficult to assess their utility and limiting the advancement of digital health apps for diabetes care. Here we report iterative usability testing and pilot testing of iSpy showing that use of the app was associated with improved carbohydrate counting accuracy and high acceptability and satisfaction scores. Areas for further refinement were also identified such as increased speed and more focus on image and text recognition features.

Carbohydrate counting is an important component of diabetes care [31-33], and use of iSpy was associated with fewer counting errors of >10 g. Although not measured directly, this degree of improved accuracy would theoretically lead to improved postprandial glucose, and the improvement we observed in HbA_{1c} levels may suggest an effect on overall glycemic control, a finding that warrants further study in future trials. Errors of >10 g are considered clinically important [8,15,34], with one

XSL•FO
RenderX

study reporting that children who received prandial insulin boluses based on carbohydrate estimates within 10 g had minimal changes in blood glucose postprandially [34]. On the other hand, when insulin boluses were based on carbohydrate estimates that were off by 20 g, more instances of hypo- and hyperglycemia occurred 2-3 hours after the meal [15].

Comparison With Prior Work

iSpy is not the only app that has been developed to assist patients or caregivers with carbohydrate counting; however, the majority of the other apps, such as MyFitnessPal or Samsung Health, are general nutrition tracking apps. Data are limited, but available studies report conflicting information regarding the difference between output nutrient data from these apps compared to reference data [35,36]. Furthermore, these apps have limited input modalities (generally limited to manual text searching) [37]. Other diabetes-related apps address multiple aspects of diabetes self-management, including tracking of glucose data, physical activity, diet, and insulin doses; such apps may also include assistance with carbohydrate counting. When tested, many of these apps have not demonstrated significant improvements in their primary outcomes, which have mainly been centered on glycemic control [28,38]. It is also worth noting that while these apps are all-encompassing for diabetes self-management, some recommend educational features such as carbohydrate counting to improve their usage [39]. However, apps that have been developed specifically to assist with carbohydrate counting are limited, and few of these have been formally evaluated, with studies having only been conducted over a duration of a few days or weeks and lacking successful comparisons with controls [40-44].

Thus, our phased development of iSpy, along with usability testing and the 3-month pilot study are relatively unique, as are the promising results. It is possible that use of iSpy was associated with more accurate carbohydrate counting because iSpy reinforced a structured approach to carbohydrate counting: identifying each food item, determining the portion size being consumed, and asking about any “hidden” carbohydrates, such as barbeque sauce under a bun. Whether this step-by-step approach to carbohydrate counting with real-time feedback underlies the observed improvement can be tested in subsequent studies.

Participant engagement is an additional measure of an app’s usability. Although use was declining, at the end of the 3-month trial 9/21 (43%) participants were still medium to high users. It is difficult to know how to interpret this degree of usage. One could view this percentage of individuals with use at a minimum of ≥1 meal every 2 weeks as evidence of engagement that is diminishing too rapidly. However, it is our experience that app usage often wanes over time, even when users rate the app quite favorably [28]. It is this expected degree of dropoff that led us to define ongoing usage of at least ≥1 meal every 2 weeks as medium engagement. Moreover, despite the dropoff, iSpy use was associated with improved carbohydrate counting at 3 months, suggesting that use of an educational app may have long lasting impact even after the period of high use has ended. Nevertheless, it will be important to consider options to further improve engagement such as push alert notifications, reminders, and ensuring ease of data entry.

Although randomized controlled trials are considered the gold standard for evaluating efficacy, there is concern that such trials may not be optimal for the assessment of apps. Thus, when considering future testing of iSpy and other apps, one must acknowledge the long timeline for recruitment and conduct of such trials within a rapidly and continually evolving technology-based environment [28,45]. User testing and product revision often occurs on shorter timelines, necessitating consideration of adaptive clinical trials that allow for continual modifications while data are being collected [46]. Furthermore, evaluation of potential barriers to incorporating app use into ongoing clinical care should also be assessed as a component of such trials. Assessing and addressing topics such as workflow integration and patient (or family)–provider communication will be needed to continue to support effective advancement of digital health [28,45].

Limitations

While our results are encouraging, we acknowledge that our studies may have had some limitations. The studies were conducted at a single tertiary pediatric center, and the results may not be generalizable. A larger trial and wider clinical implementation study is an important next step to verify our findings. In addition, although based on databases of commonly consumed foods [14], the number of foods recognized by iSpy is not all-encompassing. The database was not identified as a limiting factor by our participants, but we will continue to expand iSpy’s ability to recognize foods eaten around the world among different cultures. Though no differences were found between the intervention and control groups for baseline technology familiarity and use, we did not acquire detailed information about other factors that can influence care such as education level, socioeconomic status data, family dynamics, or details of treatment regimen, all of which could have accounted for some differences. Finally, we did not provide text reminders to the control subjects in this pilot study. Although these reminders were brief texts that occurred at most 3 times over the trial, it is feasible that they could have motivated change and thus represent a confounding variable affecting our results.

Conclusion

Carbohydrate counting remains a challenge for youth with type 1 diabetes and their families, and errors in counting can have clinical impact. We have developed and conducted rigorous pilot testing of an app designed to assist youth with carbohydrate counting. The data suggest that use of iSpy is associated with improved carbohydrate counting and that usability and acceptability of the app is quite positive. Further testing is now warranted to verify these pilot data and determine if the app can indeed improve blood glucose control and help decrease the burden of living with type 1 diabetes.

XSL•FO
RenderX

Acknowledgments

The authors thank the individuals who participated in this study, Sarah Lloyd (clinical research project assistant) and Yalinie Kulandaivelu (graduate student at the time) from the iOuch Lab for their assistance with app usability testing, and Supriya Parikh and Derek Stephens for assistance with statistical analyses. This study was financially supported by a grant from the Physicians' Services Incorporated Foundation.

Authors' Contributions

MRP and JNS supervised the study. MRP, EEYC, VP, JNS, CN, and JEA designed research and gained research funding. MRP, TA, VP, and S-ASS conducted the research. VP, MRP, and S-ASS were responsible for validating and balancing the carbohydrate counting task food lists. TA, S-ASS , and BRM analyzed data. MRP, TA, BRM, JEA, and EEYC wrote the manuscript. S-ASS, CN, VP, and JNS helped with revisions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

EEYC and JEA are the inventors of iSpy and currently own rights to its intellectual property; thus, it is feasible that they could benefit financially should iSpy become a commercial product. This conflict was acknowledged and included in relevant study documentation. Moreover, the duality of interest was mitigated by having data collection and entry conducted by 2 individuals (S-ASS, TA) with no conflicts of interest and who are independent from the inventors. Data analysis was conducted by an independent, institution-based statistician (BRM), also without conflict. The remaining authors declare no conflict of interest.

Multimedia Appendix 1
Screenshots of the primary features of the iSpy app.
[PNG File , 1038 KB - mhealth_v8i10e22074_app1.png]

Multimedia Appendix 2
Acceptability e-scale questionnaire results.
[DOCX File , 14 KB - mhealth_v8i10e22074_app2.docx]

References

1. Chapter 5: Diabetes in Canada: facts and figures from a public health perspective – youth and children. Government of Canada Public Health Agency of Canada. URL: https://www.canada.ca/en/public-health/services/chronic-diseases/reports-publications/diabetes/diabetes-canada-facts-figures-a-public-health-perspective/chapter-5.html [accessed 2020-04-27]
2. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, Cheng AYY. Canadian Diabetes Association 2013 clinical practice guidelines for the prevention and management of diabetes in Canada. Introduction. Can J Diabetes 2013 Apr;37 Suppl 1:S1-S3. [doi: 10.1016/j.jcjd.2013.01.009] [Medline: 24070926]
3. Gökşen D, Atik Altınok Y, Ozen S, Demir G, Darcan S. Effects of carbohydrate counting method on metabolic control in children with type 1 diabetes mellitus. J Clin Res Pediatr Endocrinol 2014;6(2):74-78 [FREE Full text] [doi: 10.4274/Jcrpe.1191] [Medline: 24932599]
4. Brazeau AS, Mircescu H, Desjardins K, Leroux C, Strychar I, Ekoé JM, et al. Carbohydrate counting accuracy and blood glucose variability in adults with type 1 diabetes. Diabetes Res Clin Pract 2013 Jan;99(1):19-23. [doi: 10.1016/j.diabres.2012.10.024] [Medline: 23146371]
5. Mehta SN, Quinn N, Volkening LK, Laffel LMB. Impact of carbohydrate counting on glycemic control in children with type 1 diabetes. Diabetes Care 2009 Jun;32(6):1014-1016 [FREE Full text] [doi: 10.2337/dc08-2068] [Medline: 19244089]
6. Bell KJ, Barclay AW, Petocz P, Colagiuri S, Brand-Miller JC. Efficacy of carbohydrate counting in type 1 diabetes: a systematic review and meta-analysis. Lancet Diabetes Endocrinol 2014 Mar;2(2):133-140. [doi: 10.1016/S2213-8587(13)70144-X] [Medline: 24622717]
7. Ahola AJ, Mäkimattila S, Saraheimo M, Mikkilä V, Forsblom C, Freese R, FinnDIANE Study Group. Many patients with Type 1 diabetes estimate their prandial insulin need inappropriately. J Diabetes 2010 Sep;2(3):194-202 [FREE Full text] [doi: 10.1111/j.1753-0407.2010.00086.x] [Medline: 20923484]
8. Bishop FK, Maahs DM, Spiegel G, Owen D, Klingensmith GJ, Bortsov A, et al. The Carbohydrate Counting in Adolescents With Type 1 Diabetes (CCAT) Study. Diabetes Spectrum 2009 Jan 01;22(1):56-62. [doi: 10.2337/diaspect.22.1.56]
9. Lancaster BM, Pfeffer B, McElligott M, Ferguson AT, Miller M, Wallace D, et al. Assessing treatment barriers in young adults with type 1 diabetes. Diabetes Res Clin Pract 2010 Dec;90(3):243-249. [doi: 10.1016/j.diabres.2010.07.003] [Medline: 21035226]
10. White RO, Wolff K, Cavanaugh KL, Rothman R. Addressing health literacy and numeracy to improve diabetes education and care. Diabetes Spectr 2010 Oct 02;23(4):238-243 [FREE Full text] [doi: 10.2337/diaspect.23.4.238] [Medline: 21297890]
11. Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. Pediatr Diabetes 2013 Jun 30;14(4):231-238 [FREE Full text] [doi: 10.1111/pedi.12034] [Medline: 23627878]

XSL•FO
RenderX

12. Danaher BG, Seeley JR. Methodological issues in research on web-based behavioral interventions. Ann Behav Med 2009 Aug;38(1):28-39 [FREE Full text] [doi: 10.1007/s12160-009-9129-0] [Medline: 19806416]
13. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 Sep 16;321(7262):694-696 [FREE Full text] [doi: 10.1136/bmj.321.7262.694] [Medline: 10987780]
14. Rockett HR, Breitenbach M, Frazier AL, Witschi J, Wolf AM, Field AE, et al. Validation of a youth/adolescent food frequency questionnaire. Prev Med 1997;26(6):808-816. [doi: 10.1006/pmed.1997.0200] [Medline: 9388792]
15. Smart CE, King BR, McElduff P, Collins CE. In children using intensive insulin therapy, a 20-g variation in carbohydrate amount significantly impacts on postprandial glycaemia. Diabet Med 2012 Jul;29(7):e21-e24. [doi: 10.1111/j.1464-5491.2012.03595.x] [Medline: 22268422]
16. Diabetes Canada Clinical Practice Guidelines Expert Committee. Definition, classification and diagnosis of diabetes, prediabetes and metabolic syndrome. In: Sellers E, editor. Diabetes Canada 2018 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada. Winnipeg: Canadian Journal of Diabetes; Apr 2018:S10-S15.
17. Usability testing. Usability.gov.: U.S. General Services Administration Technology Transformation Services URL: https://www.usability.gov/how-to-and-tools/methods/usability-testing.html [accessed 2020-04-25]
18. Pais V, Patel BP, Ghayoori S, Hamilton JK. Counting carbs to be in charge: a comparison of an internet-based education module with in-class education in adolescents with type 1 diabetes. Clinical Diabetes 2020 Oct (forthcoming).
19. National database for standard reference release 28. The United States Department of Agriculture. URL: https://data.nal.usda.gov/dataset/usda-national-nutrient-database-standard-reference-legacy-release [accessed 2018-03-01]
20. Canadian nutrient file. Government of Canada. URL: https://www.canada.ca/en/health-canada/services/food-nutrition/healthy-eating/nutrient-data/canadian-nutrient-file-2015-download-files.html [accessed 2018-03-01]
21. Koontz MB, Cuttler L, Palmert MR, O'Riordan M, Borawski EA, McConnell J, et al. Development and validation of a questionnaire to assess carbohydrate and insulin-dosing knowledge in youth with type 1 diabetes. Diabetes Care 2010 Mar;33(3):457-462 [FREE Full text] [doi: 10.2337/dc09-0390] [Medline: 20007940]
22. Ingersoll GM, Marrero DG. A modified quality-of-life measure for youths: psychometric properties. Diabetes Educ 1991;17(2):114-118. [doi: 10.1177/014572179101700219] [Medline: 1995281]
23. Skinner TC, Hoey H, McGee HM, Skovlund SE, Hvidøre Study Group on Childhood Diabetes. A short form of the Diabetes Quality of Life for Youth questionnaire: exploratory and confirmatory analysis in a sample of 2,077 young people with type 1 diabetes mellitus. Diabetologia 2006 Apr;49(4):621-628. [doi: 10.1007/s00125-005-0124-0] [Medline: 16525844]
24. La Greca AM, Swales T, Klemp S, Madigan S. Self care behaviors among adolescents with diabetes. 1988 Apr 30 Presented at: Ninth Annual Sessions of the Society of Behavioural Medicine; April 27-30, 1988; Boston, MA p. A42.
25. Greco P, La Greca AM, Ireland A, Wick P, Freeman C, Agramonte R. Assessing adherence in IDDM: A comparison of two methods. Miami, FL: Diabetes 40(Suppl 1); 1990:A165.
26. Lewin AB, LaGreca AM, Geffken GR, Williams LB, Duke DC, Storch EA, et al. Validity and reliability of an adolescent and parent rating scale of type 1 diabetes adherence behaviors: the Self-Care Inventory (SCI). J Pediatr Psychol 2009 Oct;34(9):999-1007 [FREE Full text] [doi: 10.1093/jpepsy/jsp032] [Medline: 19423660]
27. Wu W, Johnson R, Schepp KG, Berry DL. Electronic self-report symptom and quality of life for adolescent patients with cancer: a feasibility study. Cancer Nurs 2011;34(6):479-486. [doi: 10.1097/NCC.0b013e31820a5bdd] [Medline: 21372703]
28. Goyal S, Nunn CA, Rotondi M, Couperthwaite AB, Reiser S, Simone A, et al. A mobile app for the self-management of type 1 diabetes among adolescents: a randomized controlled trial. JMIR Mhealth Uhealth 2017 Jun 19;5(6):e82 [FREE Full text] [doi: 10.2196/mhealth.7336] [Medline: 28630037]
29. 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]
30. Veazie S, Winchell K, Gilbert J, Paynter R, Ivlev I, Eden K, et al. Mobile Applications for Self-Management of Diabetes. AHRQ Comparative Effectiveness Technical Briefs. Rockville, Maryland: Agency for Healthcare Research and Quality; 2018 May. URL: https://www.ncbi.nlm.nih.gov/books/NBK518944/pdf/Bookshelf_NBK518944.pdf [accessed 2020-09-30]
31. Tascini G, Berioli MG, Cerquiglini L, Santi E, Mancini G, Rogari F, et al. Carbohydrate counting in children and adolescents with type 1 diabetes. Nutrients 2018 Jan 22;10(1) [FREE Full text] [doi: 10.3390/nu10010109] [Medline: 29361766]
32. Gokosmanoglu F, Onmez A. Influence of flexible insulin dosing with carbohydrate counting method on metabolic and clinical parameters in type 1 diabetes patients. Open Access Maced J Med Sci 2018 Aug 20;6(8):1431-1434 [FREE Full text] [doi: 10.3889/oamjms.2018.256] [Medline: 30159070]
33. Yamada E, Okada S, Nakajima Y, Bastie CC, Tagaya Y, Osaki A, et al. Effect of carbohydrate counting using bolus calculators on glycemic control in type 1 diabetes patients during continuous subcutaneous insulin infusion. J Diabetes Investig 2017 Jul;8(4):496-500 [FREE Full text] [doi: 10.1111/jdi.12604] [Medline: 27896954]
34. Smart CE, Ross K, Edge JA, Collins CE, Colyvas K, King BR. Children and adolescents on intensive insulin therapy maintain postprandial glycaemic control without precise carbohydrate counting. Diabet Med 2009 Mar;26(3):279-285. [doi: 10.1111/j.1464-5491.2009.02669.x] [Medline: 19317823]

XSL•FO
RenderX

35. Fallaize R, Zenun Franco R, Pasang J, Hwang F, Lovegrove JA. Popular nutrition-related mobile apps: an agreement assessment against a UK reference method. JMIR Mhealth Uhealth 2019 Feb 20;7(2):e9838 [FREE Full text] [doi: 10.2196/mhealth.9838] [Medline: 30785409]
36. Fadhil A. Comparison of Self-monitoring Feedback Data from Electronic Food and Nutrition Tracking Tools. Cornell University: Computer Science - Computers and Society. Ithaca, New York: Cornell University; 2019 Apr 17. URL: https://arxiv.org/abs/1904.08376 [accessed 2020-09-30]
37. Franco RZ, Fallaize R, Lovegrove JA, Hwang F. Popular nutrition-related mobile apps: a feature assessment. JMIR Mhealth Uhealth 2016 Aug 01;4(3):e85 [FREE Full text] [doi: 10.2196/mhealth.5846] [Medline: 27480144]
38. Skrøvseth SO, Årsand E, Godtliebsen F, Joakimsen RM. Data-driven personalized feedback to patients with type 1 diabetes: a randomized trial. Diabetes Technol Ther 2015 Jul;17(7):482-489 [FREE Full text] [doi: 10.1089/dia.2014.0276] [Medline: 25751133]
39. Jeffrey B, Bagala M, Creighton A, Leavey T, Nicholls S, Wood C, et al. Mobile phone applications and their use in the self-management of Type 2 Diabetes Mellitus: a qualitative study among app users and non-app users. Diabetol Metab Syndr 2019;11:84 [FREE Full text] [doi: 10.1186/s13098-019-0480-4] [Medline: 31636719]
40. Mookherji S, Mehl G, Kaonga N, Mechael P. Unmet Need: Improving mHealth Evaluation Rigor to Build the Evidence Base. J Health Commun 2015;20(10):1224-1229. [doi: 10.1080/10810730.2015.1018624] [Medline: 26042660]
41. Rhyner D, Loher H, Dehais J, Anthimopoulos M, Shevchik S, Botwey RH, et al. Carbohydrate estimation by a mobile phone-based system versus self-estimations of individuals with type 1 diabetes mellitus: a comparative study. J Med Internet Res 2016 May 11;18(5):e101 [FREE Full text] [doi: 10.2196/jmir.5567] [Medline: 27170498]
42. 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]
43. Rollo ME, Ash S, Lyons-Wall P, Russell AW. Evaluation of a mobile phone image-based dietary assessment method in adults with type 2 diabetes. Nutrients 2015 Jun 17;7(6):4897-4910 [FREE Full text] [doi: 10.3390/nu7064897] [Medline: 26091234]
44. Rossi MC, Nicolucci A, Lucisano G, Pellegrini F, Di Bartolo P, Miselli V, et al. Impact of the "Diabetes Interactive Diary" telemedicine system on metabolic control, risk of hypoglycemia, and quality of life: a randomized clinical trial in type 1 diabetes. Diabetes Technol Ther 2013 Aug;15(8):670-679. [doi: 10.1089/dia.2013.0021] [Medline: 23844569]
45. Pham Q, Wiljer D, Cafazzo JA. Beyond the randomized controlled trial: a review of alternatives in mHealth clinical trial methods. JMIR Mhealth Uhealth 2016 Sep 09;4(3):e107 [FREE Full text] [doi: 10.2196/mhealth.5720] [Medline: 27613084]
46. Thorlund K, Haggstrom J, Park JJ, Mills EJ. Key design considerations for adaptive clinical trials: a primer for clinicians. BMJ 2018 Mar 08;360:k698. [doi: 10.1136/bmj.k698] [Medline: 29519932]

Abbreviations

HbA_{1c}: hemoglobin A_{1c}

Edited by SB Tsai; submitted 02.07.20; peer-reviewed by R Lal, M Haller; comments to author 29.07.20; revised version received 06.09.20; accepted 14.09.20; published 28.10.20.

Please cite as:
Alfonsi JE, Choi EEY, Arshad T, Sammott SAS, Pais V, Nguyen C, Maguire BR, Stinson JN, Palmert MR
Carbohydrate Counting App Using Image Recognition for Youth With Type 1 Diabetes: Pilot Randomized Control Trial
JMIR Mhealth Uhealth 2020;8(10):e22074
URL: http://mhealth.jmir.org/2020/10/e22074/
doi:10.2196/22074
PMID:33112249

©Jeffrey E Alfonsi, Elizabeth E Y Choi, Taha Arshad, Stacie-Ann S Sammott, Vanita Pais, Cynthia Nguyen, Bryan R Maguire, Jennifer N Stinson, Mark R Palmert. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 28.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Adoption of Mobile Health Apps in Dietetic Practice: Case Study of Diyetkolik

Gorkem Akdur[1*], BSc, MSc, MA; Mehmet Nafiz Aydin[1*], BSc, MA, PhD; Gizdem Akdur[2*], BEng, MSc, PhD

[1]Department of Management Information Systems, Kadir Has University, Istanbul, Turkey

[2]Centre for Research in Public Health and Community Care, University of Hertfordshire, Hatfield, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Gorkem Akdur, BSc, MSc, MA
Department of Management Information Systems
Kadir Has University
Kadir Has St., Cibali / Fatih
Istanbul, 34083
Turkey
Phone: 90 5360662866
Email: gorkemakdur@gmail.com

Abstract

Background: Dietetics mobile health apps provide lifestyle tracking and support on demand. Mobile health has become a new trend for health service providers through which they have been shifting their services from clinical consultations to online apps. These apps usually offer basic features at no cost and charge a premium for advanced features. Although diet apps are now more common and have a larger user base, in general, there is a gap in literature addressing why users intend to use diet apps. We used Diyetkolik, Turkey's most widely used online dietetics platform for 7 years, as a case study to understand the behavioral intentions of users.

Objective: The aim of this study was to investigate the factors that influence the behavioral intentions of users to adopt and use mobile health apps. We used the Technology Acceptance Model and extended it by exploring other factors such as price-value, perceived risk, and trust factors in order to assess the technology acceptance of users.

Methods: We conducted quantitative research on the Diyetkolik app users by using random sampling. Valid data samples gathered from 658 app users were analyzed statistically by applying structural equation modeling.

Results: Statistical findings suggested that perceived usefulness (P<.001), perceived ease of use (P<.001), trust (P<.001), and price-value (P<.001) had significant relationships with behavioral intention to use. However, no relationship between perceived risk and behavioral intention was found (P=.99). Additionally, there was no statistical significance for age (P=.09), gender (P=.98), or previous app use experience (P=.14) on the intention to use the app.

Conclusions: This research is an invaluable addition to Technology Acceptance Model literature. The results indicated that 2 external factors (trust and price-value) in addition to Technology Acceptance Model factors showed statistical relevance with behavioral intention to use and improved our understanding of user acceptance of a mobile health app. The third external factor (perceived risk) did not show any statistical relevance regarding behavioral intention to use. Most users of the Diyetkolik dietetics app were hesitant in purchasing dietitian services online. Users should be frequently reassured about the security of the platform and the authenticity of the platform's dietitians to ensure that users' interactions with the dietitians are based on trust for the platform and the brand.

(JMIR Mhealth Uhealth 2020;8(10):e16911) doi:10.2196/16911

KEYWORDS

mHealth; technology acceptance; user acceptance; mobile apps; diet apps; Technology Acceptance Model; TAM; dietetics

XSL•FO
RenderX

Introduction

Background

The advancement of information and communication technologies has urged the health care industry to develop new solutions to bring better health functions to individuals by reducing costs and time as well as enhancing convenience for both service providers and users [1]. Over the last decade, considerable health services have been partially or wholly shifted to mobile phones. Since the concept of mobile health (mHealth) was first coined in 2005 [2,3], mHealth has incorporated innovative medical services, such as mobile dietetics services. Most mHealth services consist of innovative solutions such as web-based consultation systems with physicians, web-based health conferences, health-relevant data, medical inspection results via portable and wearable gadgets, and smartphone-based apps [3].

Diet apps are a subset of mHealth and typically offer services including physical activity, calorie, water intake, and weight tracking; weight goal setting; meal recipes; and meal planning [4]. Due to the excessive number of diet apps that promote physical health and well-being, it is crucial to determine whether these apps are valid and what they provide for their users [5]. Generally, users pay attention to the number of downloads, average ratings, and reviews before downloading diet apps [6]. An important feature of Diyetkolik is its dietician consultation service. Diyetkolik allows dietitians to join the platform as experts and to offer their customized services to users.

Diyetkolik was founded in 2011, and as of 2020, it has a base of around 1.5 million users [7]. It is Turkey's most widely used online dietetics platform [8]. The platform has a large database of food that is specific to Turkey. Diyetkolik has features such as calorie checker, calorie tracking, water tracking, exercise tracking, content, calculations, and reports. Users are offered 3 types of packages: users with a basic (free) membership can access Diyetkolik's food database, track calorie intake, and read blog posts; users with a standard diet package are assigned noncustomized diet plans and can ask a limited number of questions to a dietitian; and users with a premium diet package can benefit from customized diet plans that are prepared by dietitians. Dietitians can create expert accounts on Diyetkolik, publish nutrition-related blog posts, and connect with clients who have subscribed to a premium membership. The business strategy of Diyetkolik is business to business to consumer, which is a novel approach in dietitian services in Turkey [9].

Previous reviews [10] have found that the use of nutrition and diet mHealth apps can change nutrition health behavior and improve diet. According to Carter et al [11], app users show greater adherence, retention, and weight loss compared to those shown by other nutrition and diet site users. Yet, what is unclear is why people choose to use mHealth apps and how they use them [12]. Generally, users choose any specific mobile app based on such factors as reliability, ease of use, quality, usefulness, aesthetics, trust, and recommendations by others [13]. Hence, engineers, managers, and app developers must acknowledge what end users demand. Many previous studies [14-19] have focused on mHealth apps, but there are only a few studies [20-24] that identify diet mHealth apps from the user's perspective. The objective of this research was to analyze the influencing factors for user acceptance of a diet mHealth app and investigate and test the importance of external constructs affecting users' behavioral intention to use. Only a few studies [3,25-27] have analyzed end-user perspectives on diet app acceptance in non-European Union countries; this study offers additional insights on mHealth user acceptance in Turkey.

Theoretical Framework and Research Hypotheses

One of the major influential models that has been used to assess the acceptance of technologies is the Technology Acceptance Model (TAM) [28], originally developed by Davis [29] in 1989. Davis [30] has stated that regulating circumstances to obtain and adopt any technology are based on some presumptions, and the major key factors are perceived ease of use and perceived usefulness [30]. According to McFarland and Hamilton [31], extended models can increase the explanation and forecasting of acceptance. Considering innovative technologies and their effects, Thompson et al [32] discussed that perceived ease of use and perceived usefulness are not the only suitable factors to determine technology adoption. For particular settings, using different variables from other information technology acceptance models could enhance the specificity and explanatory power of the extended models [30]. For this reason, instead of using the original TAM, or its well-known extensions such as the Technology Acceptance Model 2 (TAM2) [33,34] and the Unified Theory of Acceptance and Use of Technology (UTAUT), we adopted the most relevant and complementary external factors to extend TAM and explain user adoption of diet apps [35].

Perceived usefulness is one of the main factors determining, and main indicator of, behaviors of any kind of technology usage [29,30]. Davis defined *perceived usefulness* as "the degree to which a person believes that using a particular system would enhance his or her job performance" [29]; in this study, we integrated and adapted this definition of perceived usefulness—the degree to which a person believes that using a particular system would enhance their health and well-being to make it more user-centric. Our second construct was *perceived ease of use*, which is described by Davis as "the degree to which a person believes that using a particular system would be free from effort [29]," and we wholly integrated this definition into our model.

The major goal of the TAM framework is to understand the behavioral intention of users (acceptance) toward and their actual use of a technology [36,37]. Davis suggested that future technology acceptance studies should focus on other principles that influence usefulness, ease of use, and user acceptance [37]. Thus, price-value, trust, and perceived risk factors were used to extend the TAM framework in this study. These factors are crucial to exploring the perspective of users in the mHealth domain [3,38,39].

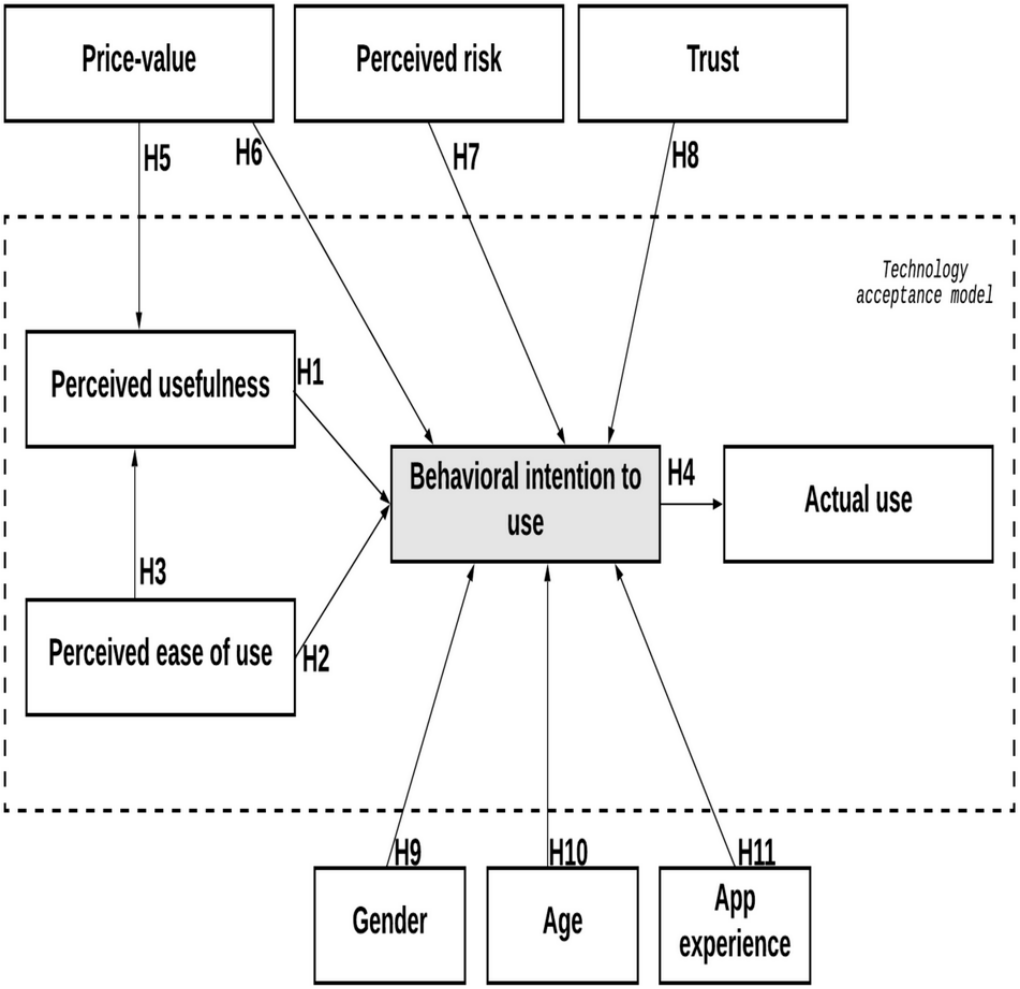
The *price-value* factor is related to the subscription fees of the service offered by the mobile health services. It has a direct effect on user acceptance; therefore, we adopted price-value as an external factor. We investigated the direct relationship of user satisfaction on the price that the users pay and whether the

XSL•FO
RenderX

service matched their needs [38,40,41]. Moreover, *trust* is a powerful external factor since trust plays a major role in attracting new users and retaining existing users. If a person believes in the service administered by an mHealth app, they are likely to adopt the service. According to the TAM framework, a person's beliefs can encourage their assertiveness and intention to adopt the technology [3,42,43]. Hence, we integrated trust as another external factor in this study.

The external factor *perceived risk* can be defined as a "person's perception of uncertainty in the use of mHealth services and its severity in terms of consequences [39]." There are some risks posed by mHealth apps, such as privacy invasion, legal, and performance risks [44]. Researchers have shown that privacy and perceived risk can predict user adoption [45,46], hence we used perceived risk in our extended model. Finally, some control variables were added to enhance the explanatory potential of the study. Gender, age, and previous mHealth app experience were selected as covariates. Considering these factors, hypotheses and the research framework were established; they are illustrated in Figure 1 and are summarized in Table 1.

Figure 1. Theoretical framework.



XSL•FO
RenderX

Table 1. Hypotheses list.

Label	Hypotheses
H1	Perceived usefulness positively affects the intention to use the app.
H2	Perceived ease of use positively affects the intention to use the app.
H3	Perceived ease of use positively affects the perceived usefulness of the app.
H4	Behavioral intention to use positively affects the actual use of the app.
H5	Price-value factor positively affects the perceived usefulness.
H6	Price-value factor positively affects the intention to use the app.
H7	Perceived risk is negatively associated with one's adoption intention toward the app.
H8	Trust positively affects the intention to use the app.
H9	Gender difference plays a significant role in the intention to use the app.
H10	Age has a significant role in the intention to use the app.
H11	Previous mHealth app use experience of users has a significant role in the intention to use the app.

Methods

Study Design

Multiple-itemed scales were used to quantify each construct. All questions were carefully selected and adapted from the questionnaires used in previous studies [1,3,29,43,47-50]. We adjusted most items to make them more suitable for an mHealth context (please see Multimedia Appendix 1 for references for each item). Perceived usefulness was measured with 4 items (PU1-PU4), perceived ease of use was measured with 4 items (PEOU1-PEOU4), behavioral intention to use was measured with 5 items (BI1-BI5), price-value was measured with 3 items (PV1-PV3), perceived risk was measured with 6 items (PR1-PR6), and trust was measured with 3 items (T1-T3).

All of these items were rated by users in a 32-question online survey using a 5-point Likert scale (1, strongly disagree; 3, indecisive; 5, strongly agree). Questions were shuffled and one attention check (trap) question was added to eliminate bias responses. The aim of the attention check question was to ensure that the users were focused. One of the perceived usefulness items was stated as the exact opposite of another perceived usefulness item. In this way, we attempted to minimize the risk of collecting less reliable and inaccurate data from users who were not being attentive [51]. Demographic questions were asked at the beginning of the survey: gender, age group, education level, usage frequency of the app (which directly answers the actual use factor), previous mobile-health app experience, main aim for the use of the app (users were allowed to choose more than one answer from the pool), and membership type (free 1-month standard membership, 3-month standard membership, premium-dietitian service). The survey was prepared in the Turkish language, which is the native language of the app and its users. The survey was sent to Diyetkolik managers to be placed on their survey platform.

Data Collection

Data were collected with an online survey method on a Typeform web-based platform [52]. This platform is used by Diyetkolik to gather feedback from its users. The survey was distributed randomly to active users who subscribed to Diyetkolik's mailing lists. Out of 100,000 active users, 50,000 users who checked their weekly emails from Diyetkolik, randomly received the survey link via email. A brief explanation of the project was given in the email and indicated that participation in the survey was voluntary and completely anonymous (please see Multimedia Appendix 1 for the English version and Multimedia Appendix 2 for the Turkish version of the survey).

The data collection phase took 4 weeks between June 11, 2019 and July 10, 2019. Responses were collected from 840 Diyetkolik users. In order to increase the reliability and accuracy, surveys that had missing entries and inattentive responses were eliminated. Hence, the data count was reduced to 658 from the original 840.

Statistical Procedure

Data were analyzed using SmartPLS 3 (SmartPLS GmbH) and SPSS (version 24.0; IBM Corp) software. Data were gathered in an Excel (Microsoft Inc) file and the shuffled items were ordered categorically. Partial least square, a component-based structural equation modeling technique, was used to examine the research model and test the first 8 hypotheses. Partial least square methods are particularly convenient for complex and large research models to test both reflective and formative constructs. Hence, partial least square-based structural equation modeling was better suited to this research than covariance-based structural equation modeling techniques [39]. For the last 3 hypotheses and for demographics, analysis of variance (ANOVA) was used (at the significance levels $P<.01$ and $P<.05$).

Analysis of the Measurement Model in Structural Equation Modeling

To assess construct validity and internal accuracy, we used discriminant and convergent validity measures that complemented each other [1]. Cronbach α and composite reliability measures must be around 0.70 for each factor to have a valid internal consistency. Additionally, the average variance extracted and component loadings should be more than 0.50 to have acceptable convergent and construct validity [53].

XSL•FO
RenderX

Calculating the standardized outer loadings of main variables can explain individual variance level, convergent validity, and individual manifest reliability. Thus, items that have more than 0.70 outer loadings count as highly satisfactory. Items which have a value of 0.50 or below should be eliminated to increase the composite reliability. Items valued between 0.40 to 0.70 should be reviewed before dropping out [54].

Composite reliability and Cronbach α values play an important role in analyzing internal consistency. A common rule of thumb indicates that values 0.6 or higher are adequate for exploratory purposes [55].

Discriminant Validity Measurement

Discriminant validity was checked using the Fornell-Larcker criterion (correlation analysis). According to this criterion, an average variance extracted value above 0.50 is acceptable, and an average variance extracted value above 0.70 shows significant validity.

Another method to check validity is by analyzing the R^2 values. It is a goodness-of-fit measure for linear regression models. It explains the estimated variance of the construct's relationship with independent variables and higher R^2 values show significantly better model fits [56]. The value of R^2 of underlying variables should be greater than 0.26 [57].

Hypothesis Testing

To identify the relationships between the variables in the structural model, path coefficient (β) and *t* values were tested by bootstrapping with 1000 subsamples. The relationships among factors are statistically proven with *P* values lower than .01 and with *t* values larger than 1.96.

Results

Demographic Characteristics

The demographic characteristics of respondents (n=658) are presented in Table 2. The majority of the respondents were female (438/658, 66.6%) and between the ages of 18-41 years (496/658, 75.4%). Most of the respondents had attained a bachelor's degree or higher (362/658, 55.0%). Two-thirds of the respondents did not disclose their membership type, and almost one-third of the respondents disclosed that their membership type was basic (211/658, 32.1%).

Analysis of the Measurement Model in Structural Equation Modeling

In the first round, the items BI5 (0.403), PV3 (0.594), PR1 (–0.772), PR2 (0.330), PR4 (0.556), and PR6 (0.450) were removed from our partial least square structural equation model. Only the items that had outer loadings greater than 0.6 remained, and the final data analysis table was constructed.

Table 3 shows the final measurement model with the mean values. According to the results, internal consistency is satisfied; perceived risk, price-value, and trust (the 3 external variables) had mean values of approximately 3 indicating that the average respondent was indecisive about perceived risk, trust, and price-value.

Discriminant Validity Measurement

The model passed the validity test with satisfactory results. All variables in Table 4 show discriminant validity with cross loadings above 0.70 (actual use: 1; behavioral intention: 0.732; perceived ease of use: 0.739; price-value: 0.796; perceived risk: 0.873; perceived usefulness: 0.789; trust: 0.873).

For this study, 2 out of 3 R^2 values satisfied the criteria (behavioral intention: R^2=0.635; perceived usefulness: R^2=0.434; perceived ease of use: R^2<0.26).

XSL•FO
RenderX

Table 2. Demographics of sample data.

Variable	Participants (n=658), n (%)
Gender	
Male	163 (24.8)
Female	438 (66.6)
Prefer not to say	57 (8.7)
Age (years)	
18-25	169 (25.7)
26-33	168 (25.5)
34-41	159 (24.1)
42-49	112 (17.0)
50-59	34 (5.2)
60 +	16 (2.4)
Educational qualifications	
Primary school graduate	9 (1.4)
High school graduate	91 (13.8)
Bachelor’s degree	247 (37.5)
University student	102 (15.5)
Post-graduate degree and higher	115 (17.5)
2-year degree graduate	94 (14.3)
App experience	
Yes	314 (47.7)
No	344 (52.3)
Membership type	
1-month standard subscription	5 (0.8)
3-month standard subscription	2 (0.3)
Premium service	1 (0.2)
Basic (free) membership	211 (32.1)
Prefer not to say	446 (67.8)

XSL•FO
RenderX

Table 3. The measurement model with mean values.

Factors	Loadings	Mean	Average variance extracted	Composite reliability	Cronbach α
Perceived usefulness			0.623	0.868	.798
PU1	0.710	3.69			
PU2	0.806	3.73			
PU3	0.774	3.89			
PU4	0.859	3.70			
Perceived ease of use			0.546	0.780	.580
PEOU1	0.631	4.05			
PEOU3	0.695	3.28			
PEOU4	0.870	3.59			
Price-value			0.634	0.772	.450
PV1	0.905	3.12			
PV2	0.669	3.05			
Perceived risk			0.762	0.865	.688
PR3	0.884	2.84			
PR5	0.861	3.03			
Trust			0.762	0.906	.884
T1	0.852	3.33			
T2	0.854	3.22			
T3	0.912	3.36			
Behavioral intention			0.535	0.820	.708
BI1	0.678	3.40			
BI2	0.633	2.55			
BI3	0.752	3.53			
BI4	0.846	3.28			

Table 4. Correlation matrix of the square root of the average variance extracted for discriminant validity.

Variable	Variable						
	Actual use	Behavioral intention	Perceived ease of use	Price-value	Perceived risk	Perceived usefulness	Trust
Actual use	1	0.322	0.312	0.21	–0.218	0.332	0.117
Behavioral intention	0.322	0.732	0.697	0.684	–0.248	0.61	0.583
Perceived ease of use	0.312	0.697	0.739	0.647	–0.254	0.641	0.481
Price-value	0.21	0.684	0.647	0.796	–0.202	0.536	0.48
Perceived risk	–0.218	–0.248	–0.254	–0.202	0.873	–0.366	–0.247
Perceived usefulness	0.332	0.61	0.641	0.536	–0.366	0.789	0.462
Trust	0.117	0.583	0.481	0.48	–0.247	0.462	0.873

Hypothesis Testing

The results of partial least square model for the first 8 hypotheses of the study are illustrated in Table 5. Perceived risk's direct effect on behavioral intention (β=0, *t*=0.01, *P*=.99) was not accepted. The direct effects of perceived ease of use on perceived usefulness (β=0.506, *t*=12.98, *P*<.001), perceived ease of use on behavioral intention (β=0.293, *t*=6.85, *P*<.001), behavioral intention on actual use (β=0.322, *t*=9.10, *P*<.001), and price-value on behavioral intention (β=0.303, *t*=7.67, *P*<.001) were significant.

XSL•FO
RenderX

Table 5. Results of partial least square for H1 to H8.

Hypothesis	Path	β	*t* value	*P* value	Result
H1	Perceived usefulness→behavioral intention	0.156	4.28	<.001	Accepted
H2	Perceived ease of use→behavioral intention	0.293	6.85	<.001	Accepted
H3	Perceived ease of use→perceived usefulness	0.506	12.98	<.001	Accepted
H4	Behavioral intention→actual use	0.322	9.10	<.001	Accepted
H5	Price-value→perceived usefulness	0.209	5.03	<.001	Accepted
H6	Price-value→behavioral intention	0.303	7.67	<.001	Accepted
H7	Perceived risk→behavioral intention	0	0.01	.99	Not accepted
H8	Trust→behavioral intention	0.225	7.09	<.001	Accepted

Gender had no statistically significant role on the behavioral intention to use the app (mean 3.21, SD 0.873; $F_{1,599}$=0 *P*=.98. In addition, age groups showed no association between age and behavioral intention (mean 3.19, SD 0.872; $F_{5,652}$=1.931, *P*=.09). Finally, previous app use was not significant (mean 3.19, SD 0.872; $F_{1,656}$=2.137, *P*=.14). Hence, H9, H10, and H11 were all rejected (Multimedia Appendix 3).

Discussion

General

Considering the vast influence of information and communication technologies in our daily lives, this research aimed to extend TAM framework and to understand the main factors that influenced Diyetkolik users' acceptance of mobile dietitian services. Identified factors were tested for the variables' hypothesized positive associations with user acceptance of the diet app. The findings showed that perceived ease of use, perceived usefulness, price-value, and trust had a positive impact on the app acceptance at the behavioral level. We only demonstrated that high behavioral intention of users translated into higher usage frequency of the app (actual use). Previous studies [21,29,33,56,58-62] on TAM had also shown similar results. Our study's results were consistent with Ajzen and Fishbein's theory [62,63] and also in congruence with findings [29] that suggested that behavioral intention was a good predictor of actual technology use. However, perceived risk did not show any significant relationship with behavioral intention to use (*P*=.99). Users used mHealth apps and benefit from the services if they perceived the app to be useful, easy to control, and trustworthy. Moreover, the analysis of age groups and previous app use data showed that there were no statistically meaningful differences (age: *P*=.09; use: *P*=.14) in behavioral intention to use. Our results also revealed there was no major role of gender in the behavioral intention to use (*P*=.98). This conflicts with findings from another study [64] that showed that men perceived mobile diet apps to be more beneficial in managing their lifestyle and eating habits.

We predict that most respondents who did not respond to the membership question had basic membership, because the marketing and sales managers of Diyetkolik have stated that the majority of users own free accounts. Disclosing the subscription type will be made compulsory in any future research, because it could positively affect our analysis to identify the respondent type clearly. By doing this, we can differentiate our questions for different segmentations instead of focusing on all respondents as a homogenous population.

The findings from the survey's price-value questions suggested that users were ambivalent about the benefits of premium membership (the subscription type that provides dietitian services to the users). Having dietitian services straight from a mobile device, instead of a physical consultation, seems eminently practical for some respondents, but other respondents were indecisive about consulting a dietitian online through the platform. Most users have vigorously opposed or were hesitant about purchasing health services online. Some users were also indecisive about the reliability of the information that dietitians provide on the Diyetkolik platform. These might stem from issues in trust concerning the accreditation and the authenticity of the platform's dietitians.

Two external factors (trust and price-value) that were used in this study in addition to the TAM factors (perceived ease of use and perceived usefulness) improved our understanding of user acceptance of an mHealth app and showed relevance with behavioral intention to use. The third external factor (perceived risk) did not provide any explanation regarding intention to use, but it can be divided into further categories (legal concerns, privacy risks, etc), and its statistical relationship with behavioral intention can be retested in future research. Therefore, based on our findings, we propose an extended TAM framework.

Limitations

This study had some limitations. It was conducted in one country (Turkey), and the findings might not be generalizable to related apps in European Union countries or in North America. Demographic variables such as sociocultural differences with regard to diet would create differences between different countries. Future research could study different country settings and compare cross-country data. Although the investigation had a good representative sample, it was conducted online which can cause some problems regarding the attention of respondents. We tried to minimize this effect to a degree by including a trap question.

Other data collection methods might be considered for future studies, for example, inviting previous or occasionally active Diyetkolik users to take part in the study. The insights from former users could be important to understanding the factors

XSL•FO
RenderX

behind technology abandonment, resistance, or rejection; there may be close links with behavioral intention to use.

Due to the number of users preferring not to disclose their type of membership, we were unable to identify additional insights from comparing user perceptions about different subscriptions. Future research can leverage digital trace data of user actions on the mobile app and can use data analytics to examine further hypotheses.

Conclusion

Measuring the user acceptance of mHealth by using an extended TAM model was the primary aim of this study. Our research showed that the perceived ease of use and the perceived usefulness of an mHealth app were associated with the behavioral intention of users to adopt and continue using the app. A mHealth-specific extended TAM model that was developed showed that price-value and trust factors directly affected the behavioral intention to use the app. However, perceived risk and behavioral intention were not associated, and gender, age groups, and previous app use experience had no direct or moderating roles on the intention to use the app.

Despite the aforementioned limitations, the results and findings of the study contribute to mHealth app design and the development of dietary and nutrition mHealth apps. The success of apps is highly dependent on the user's willingness to adopt them. On a practical level, the insights from this research can benefit nutrition and diet app developers and managers to increase the adoption and actual use of apps. Furthermore, the study proposes a theoretical extension to the literature. TAM is a generic framework which is often used by researchers to study the user acceptance of technologies, however, there are only a few studies [22,44,65] that have explored diet-based mHealth apps. Hence, our expanded TAM framework lays out additional constructs and contributes to technology acceptance literature. We believe that further research will use this model and possibly identify new factors that could be useful in studying user acceptance of diet apps.

As a recommendation, managers might find it useful to focus on promoting the perceived usefulness of the service in order to increase the uptake of premium services by the users. The average respondent in our Diyetkolik case study was hesitant about purchasing health services online, based on the app's price-value. If the benefits of the premium packages are communicated to the users in innovative ways, and if the users are frequently reassured about the security of the app and the authenticity of the platform's experts (eg, Diyetkolik's dietitians), then they might feel more inclined to purchase premium services. This recommendation can be generalizable to other mHealth apps where the user concerns are similar.

Acknowledgments

The authors would like to thank the Diyetkolik's head of marketing and sales and Diyetkolik's management team for their support. There is no funding to declare.

Conflicts of Interest

None declared.

Multimedia Appendix 1
English version of survey questions.
[PDF File (Adobe PDF File), 185 KB - mhealth_v8i10e16911_app1.pdf]

Multimedia Appendix 2
Turkish version of survey questions.
[PDF File (Adobe PDF File), 354 KB - mhealth_v8i10e16911_app2.pdf]

Multimedia Appendix 3
Hypotheses 9, 10, and 11 one-way ANOVA results.
[PDF File (Adobe PDF File), 99 KB - mhealth_v8i10e16911_app3.pdf]

References

1. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
2. Schuster L, Tossan V, Drennan J. Consumer acceptance of mhealth services: a comparison of behavioral intention models. Services Marketing Quarterly 2017 Apr 21;38(2):115-128. [doi: 10.1080/15332969.2017.1289791]
3. Deng Z, Hong Z, Ren C, Zhang W, Xiang F. What predicts patients' adoption intention toward mhealth services in china: empirical study. JMIR Mhealth Uhealth 2018 Aug 29;6(8):e172 [FREE Full text] [doi: 10.2196/mhealth.9316] [Medline: 30158101]

XSL•FO
RenderX

4. Whisner C, Porter A, Pecor N, Maciejewski R. A survey of personal visualization in mhealth nutrition apps. Arizona State University Faculty Profile. 2015. URL: http://rmaciejewski.faculty.asu.edu/papers/2015/PersonalVis15.pdf [accessed 2020-09-24]
5. Franco RZ, Fallaize R, Lovegrove JA, Hwang F. Popular nutrition-related mobile apps: a feature assessment. JMIR Mhealth Uhealth 2016 Aug 01;4(3):e85 [FREE Full text] [doi: 10.2196/mhealth.5846] [Medline: 27480144]
6. Schumer H, Amadi C, Joshi A. Evaluating the dietary and nutritional apps in the Google Play Store. Healthc Inform Res 2018 Jan;24(1):38-45 [FREE Full text] [doi: 10.4258/hir.2018.24.1.38] [Medline: 29503751]
7. Sikca Sorulan Sorular [in Turkish]. Diyetkolik. URL: https://www.diyetkolik.com/landing-faq/ [accessed 2020-09-24]
8. Diyetkolik home page [in Turkish]. Diyetkolik. URL: https://www.diyetkolik.com/ [accessed 2020-09-24]
9. What is Diyetkolik? How it works? Diyetkolik. URL: https://www.diyetkolik.com/biz-kimiz/ [accessed 2019-07-19]
10. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. J Am Med Inform Assoc 2013 May 01;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
11. 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 Apr 15;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
12. Ernsting C, Dombrowski SU, Oedekoven M, O Sullivan JL, Kanzler M, Kuhlmey A, et al. Using smartphones and health apps to change and manage health behaviors: a population-based survey. J Med Internet Res 2017 Apr 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
13. Handel MJ. mHealth (mobile health)-using apps for health and wellness. Explore (NY) 2011;7(4):256-261. [doi: 10.1016/j.explore.2011.04.011] [Medline: 21724160]
14. Ewais S, Alluhaidan A. Classification of stress management mHealth apps based on Octalysis framework. 2015 Presented at: 21st Americas Conference on Information Systems; August 13-15; Puerto Rico.
15. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: a systematic review of prevalent and innovative methods. Am J Prev Med 2016 Dec;51(6):1051-1059. [doi: 10.1016/j.amepre.2016.07.009] [Medline: 27659122]
16. Higgins JP. Smartphone applications for patients' health and fitness. Am J Med 2016 Jan;129(1):11-19. [doi: 10.1016/j.amjmed.2015.05.038] [Medline: 26091764]
17. Aungst TD, Clauson KA, Misra S, Lewis TL, Husain I. How to identify, assess and utilise mobile medical applications in clinical practice. Int J Clin Pract 2014 Feb;68(2):155-162. [doi: 10.1111/ijcp.12375] [Medline: 24460614]
18. Larson RS. A path to better-quality mhealth apps. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e10414 [FREE Full text] [doi: 10.2196/10414] [Medline: 30061091]
19. Akter S, Ray P. mhealth - an ultimate platform to serve the unserved. Yearb Med Inform 2018 Mar 07;19(01):94-100. [doi: 10.1055/s-0038-1638697]
20. Chen J, Lieffers J, Bauman A, Hanning R, Allman-Farinelli M. The use of smartphone health apps and other mobile health (mHealth) technologies in dietetic practice: a three country study. J Hum Nutr Diet 2017 Aug;30(4):439-452. [doi: 10.1111/jhn.12446] [Medline: 28116773]
21. Okumus B, Bilgihan A, Ozturk AB. Factors affecting the acceptance of smartphone diet applications. 2015 Nov 03;25(6):726-747. [doi: 10.1080/19368623.2016.1082454]
22. Lee HE, Cho J. What motivates users to continue using diet and fitness apps? application of the uses and gratifications approach. Health Commun 2017 Dec;32(12):1445-1453. [doi: 10.1080/10410236.2016.1167998] [Medline: 27356103]
23. Okumus B, Ali F, Bilgihan A, Ozturk AB. Psychological factors influencing customers' acceptance of smartphone diet apps when ordering food at restaurants. International Journal of Hospitality Management 2018 Jun;72:67-77. [doi: 10.1016/j.ijhm.2018.01.001]
24. Wang Q, Egelandsdal B, Amdam GV, Almli VL, Oostindjer M. Diet and physical activity apps: perceived effectiveness by app users. JMIR Mhealth Uhealth 2016 Apr 07;4(2):e33 [FREE Full text] [doi: 10.2196/mhealth.5114] [Medline: 27056639]
25. Huang G, Ren Y. Linking technological functions of fitness mobile apps with continuance usage among Chinese users: moderating role of exercise self-efficacy. Comput Human Behav 2020 Feb;103:151-160. [doi: 10.1016/j.chb.2019.09.013]
26. Jospe MR, Fairbairn KA, Green P, Perry TL. Diet app use by sports dietitians: a survey in five countries. JMIR Mhealth Uhealth 2015 Jan 22;3(1):e7 [FREE Full text] [doi: 10.2196/mhealth.3345] [Medline: 25616274]
27. Cho J, Quinlan MM, Park D, Noh GY. Determinants of adoption of smartphone health apps among college students. Am J Health Behav 2014 Nov;38(6):860-870. [doi: 10.5993/AJHB.38.6.8] [Medline: 25207512]
28. Hu PJ, Chau PY, Sheng ORL, Tam KY. Examining the technology acceptance model using physician acceptance of telemedicine technology. Journal of Management Information Systems 2015 Dec 02;16(2):91-112. [doi: 10.1080/07421222.1999.11518247]
29. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]

XSL•FO
RenderX

30. Davis FD. A technology acceptance model for empirically testing new end-user information systems: theory and results. Massachusetts Institute of Technology. 1986. URL: http://hdl.handle.net/1721.1/15192 [accessed 2020-09-24]
31. McFarland DJ, Hamilton D. Adding contextual specificity to the technology acceptance model. Comput Human Behav 2006 May;22(3):427-447. [doi: 10.1016/j.chb.2004.09.009]
32. Thomson R, Compeau D, Higgins C. Intentions to use information technologies: an integrative model. Journal of Organizational and End User Computing 2006;18(3):25-46. [doi: 10.4018/joeuc.2006070102]
33. Venkatesh V, Morris G, Davis F. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
34. Venkatesh V, Davis F. A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
35. Agarwal R, Prasad J. A conceptual and operational definition of personal innovativeness in the domain of information technology. Information Systems Research 1998 Jun;9(2):204-215. [doi: 10.1287/isre.9.2.204]
36. Blagoeva KT, Mijoska M. Applying TAM to study online shopping adoption among youth in the Republic of Macedonia. 2017 Presented at: Management International Conference; May 24-27, 2017; Venice, Italy p. 543 URL: http://www.hippocampus.si/ISBN/978-961-7023-71-8/167.pdf
37. Momani A, Mamoun JM. The evolution of technology acceptance theories. International Journal of Contemporary Computer Research 2017 Mar;1(1):51. [doi: 10.4018/ijcbpl.2017040101]
38. West JH, Belvedere LM, Andreasen R, Frandsen C, Hall PC, Crookston BT. Controlling your "app"etite: how diet and nutrition-related mobile apps lead to behavior change. JMIR Mhealth Uhealth 2017 Jul 10;5(7):e95 [FREE Full text] [doi: 10.2196/mhealth.7410] [Medline: 28694241]
39. Luo X, Li H, Zhang J, Shim J. Examining multi-dimensional trust and multi-faceted risk in initial acceptance of emerging technologies: An empirical study of mobile banking services. Decision Support Systems 2010 May;49(2):222-234. [doi: 10.1016/j.dss.2010.02.008]
40. Hoj TH, Covey EL, Jones AC, Haines AC, Hall PC, Crookston BT, et al. How do apps work? an analysis of physical activity app users' perceptions of behavior change mechanisms. JMIR Mhealth Uhealth 2017 Aug 03;5(8):e114 [FREE Full text] [doi: 10.2196/mhealth.7206] [Medline: 28778846]
41. Yuan S, Ma W, Kanthawala S, Peng W. Keep using my health apps: discover users' perception of health and fitness apps with the UTAUT2 model. Telemed J E Health 2015 Sep;21(9):735-741. [doi: 10.1089/tmj.2014.0148] [Medline: 25919238]
42. Hoque MR, Bao Y, Sorwar G. Investigating factors influencing the adoption of e-Health in developing countries: a patient's perspective. Inform Health Soc Care 2017 Jan;42(1):1-17. [doi: 10.3109/17538157.2015.1075541] [Medline: 26865037]
43. Chong AY. Understanding mobile commerce continuance intentions: an empirical analysis of chinese consumers. Journal of Computer Information Systems 2015 Dec 10;53(4):22-30. [doi: 10.1080/08874417.2013.11645647]
44. Macdonald AJ. Acceptance and continuance factors associated with mobile medical app use: a qualitative case study of diabetes apps. Northcentral University.: ProQuest Dissertations Publishing; 2017. URL: https://www.proquest.com/docview/1936019458 [accessed 2020-09-24]
45. Kim DJ, Ferrin DL, Rao HR. A trust-based consumer decision-making model in electronic commerce: the role of trust, perceived risk, and their antecedents. Decision Support Systems 2008 Jan;44(2):544-564. [doi: 10.1016/j.dss.2007.07.001]
46. Alaiad A, Zhou L. The determinants of home healthcare robots adoption: an empirical investigation. Int J Med Inform 2014 Nov;83(11):825-840. [doi: 10.1016/j.ijmedinf.2014.07.003] [Medline: 25132284]
47. Ming-Syan C, Jiawei H, Yu P. Data mining: an overview from a database perspective. IEEE Trans Knowl Data Eng 1996;8(6):866-883. [doi: 10.1109/69.553155]
48. Kim TG, Lee JH, Law R. An empirical examination of the acceptance behaviour of hotel front office systems: an extended technology acceptance model. Tourism Management 2008 Jun;29(3):500-513. [doi: 10.1016/j.tourman.2007.05.016]
49. Venkatesh, Thong, Xu. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Quarterly 2012;36(1):157-178. [doi: 10.2307/41410412]
50. Cilliers L, Viljoen KL, Chinyamurindi WT. A study on students' acceptance of mobile phone use to seek health information in South Africa. Health Inf Manag 2018 May 07;47(2):59-69. [doi: 10.1177/1833358317706185] [Medline: 28537211]
51. Liu M, Wronski L. Trap questions in online surveys: results from three web survey experiments. International Journal of Market Research 2018 Feb 05;60(1):32-49. [doi: 10.1177/1470785317744856]
52. What is Typeform? Typeform. 2020. URL: https://help.typeform.com/hc/en-us/articles/360038717092-What-is-Typeform- [accessed 2019-10-01]
53. Kedah Z, Ismail Y, Ahasanul AKM, Ahmed S. Key success factors of online food ordering services: an empirical study. Malaysian Management Review 2015;50(2):19-36 [FREE Full text]
54. Memon AH, Rahman IA. SEM-PLS analysis of inhibiting factors of cost performance for large construction projects in Malaysia: perspective of clients and consultants. Scientific World Journal 2014;2014:165158-165159 [FREE Full text] [doi: 10.1155/2014/165158] [Medline: 24693227]
55. Garson GD. Partial Least Squares: Regression and Structural Equation Models. Asheboro, NC: Statistical Associates Publishing; 2016.

XSL•FO
RenderX

56. Peker C. An analysis of the main critical factors that affect the acceptance of technology in hospital management systems. The Middle East Technical University. 2010. URL: https://etd.lib.metu.edu.tr/upload/12612453/index.pdf [accessed 2020-09-24]
57. Wassertheil S, Cohen J. Statistical power analysis for the behavioral sciences. Biometrics 1970 Sep;26(3):588. [doi: 10.2307/2529115]
58. Okumus B, Bilgihan A. Proposing a model to test smartphone users' intention to use smart applications when ordering food in restaurants. JHTT 2014 Mar 11;5(1):31-49. [doi: 10.1108/jhtt-01-2013-0003]
59. Surendran P. Technology acceptance model: a survey of literature. International Journal of Business and Social Research 2012;2(4):175-178.
60. Zhao Y, Ni Q, Zhou R. What factors influence the mobile health service adoption? a meta-analysis and the moderating role of age. International Journal of Information Management 2018 Dec;43:342-350. [doi: 10.1016/j.ijinfomgt.2017.08.006]
61. Tao D. Intention to use and actual use of electronic information resources: further exploring Technology Acceptance Model (TAM). AMIA Annu Symp Proc 2009 Nov 14;2009:629-633 [FREE Full text] [Medline: 20351931]
62. Vallerand RJ, Deshaies P, Cuerrier J, Pelletier LG, Mongeau C. Ajzen and Fishbein's theory of reasoned action as applied to moral behavior: a confirmatory analysis. Journal of Personality and Social Psychology 1992 Jan;62(1):98-109. [doi: 10.1037/0022-3514.62.1.98]
63. Ajzen I, Fishbein M. Understanding Attitudes And Predicting Social Behavior. New Jersey: Prentice-Hall; 1980.
64. Zhang X, Guo X, Lai K, Guo F, Li C. Understanding gender differences in m-health adoption: a modified theory of reasoned action model. Telemed J E Health 2014 Jan;20(1):39-46. [doi: 10.1089/tmj.2013.0092] [Medline: 24161004]
65. Scheibe M, Reichelt J, Bellmann M, Kirch W. Acceptance factors of mobile apps for diabetes by patients aged 50 or older: a qualitative study. Med 2 0 2015 Mar 02;4(1):e1 [FREE Full text] [doi: 10.2196/med20.3912] [Medline: 25733033]

Abbreviations

ANOVA: analysis of variance
mHealth: mobile health
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by E Borycki; submitted 05.11.19; peer-reviewed by E Stringer, J Moon; comments to author 31.03.20; revised version received 25.05.20; accepted 23.08.20; published 02.10.20.

Please cite as:
Akdur G, Aydin MN, Akdur G
Adoption of Mobile Health Apps in Dietetic Practice: Case Study of Diyetkolik
JMIR Mhealth Uhealth 2020;8(10):e16911
URL: https://mhealth.jmir.org/2020/10/e16911
doi:10.2196/16911
PMID:33006566

©Gorkem Akdur, Mehmet Nafiz Aydin, Gizdem Akdur. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 02.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Challenges in Acceptance and Compliance in Digital Health Assessments During Pregnancy: Prospective Cohort Study

Katharina Brusniak[1*]; Hannah Maria Arndt[1*]; Manuel Feisst[2], MSc; Kathrin Haßdenteufel[1]; Lina Maria Matthies[1], Dr med; Thomas Maximilian Deutsch[1], Dr med; Hannes Hudalla[3]; Harald Abele[4,5], Prof Dr med; Markus Wallwiener[1], Prof Dr med; Stephanie Wallwiener[1], Prof Dr med

[1]Department of Gynecology and Obstetrics, University Hospital Heidelberg, Heidelberg, Germany

[2]Institute of Medical Biometry and Informatics, Heidelberg University, Heidelberg, Germany

[3]Department of Neonatology, University Hospital Heidelberg, Heidelberg, Germany

[4]Department of Women's Health, University Hospital Tübingen, Tübingen, Germany

[5]Section of Midwifery Science, Institute for Health Sciences, University Hospital Tübingen, Tübingen, Germany

[*]these authors contributed equally

Corresponding Author:
Stephanie Wallwiener, Prof Dr med
Department of Gynecology and Obstetrics
University Hospital Heidelberg
Im Neuenheimer Feld 440
Heidelberg, 69120
Germany
Phone: 49 6221 5637551
Email: Stephanie.Wallwiener@med.uni-heidelberg.de

Abstract

Background: Pregnant women are increasingly using mobile apps to access health information during the antenatal period. Therefore, digital health solutions can potentially be used as monitoring instruments during pregnancy. However, a main factor of success is high user engagement.

Objective: The aim of this study was to analyze engagement and factors influencing compliance in a longitudinal study targeting pregnant women using a digital health app with self-tracking.

Methods: Digitally collected data concerning demographics, medical history, technical aspects, and mental health from 585 pregnant women were analyzed. Patients filling out ≥80% of items at every study visit were considered to be highly compliant. Factors associated with high compliance were identified using logistic regression. The effect of a change in mental and physical well-being on compliance was assessed using a one-sample *t* test.

Results: Only 25% of patients could be considered compliant. Overall, 63% left at least one visit blank. Influential variables for higher engagement included higher education, higher income, private health insurance, nonsmoking, and German origin. There was no relationship between a change in the number of physical complaints or depressive symptoms and study dropout.

Conclusions: Maintaining high engagement with digital monitoring devices over a long time remains challenging. As cultural and socioeconomic background factors had the strongest influence, more effort needs to be directed toward understanding the needs of patients from different demographic backgrounds to ensure high-quality care for all patients. More studies need to report on compliance to disclose potential demographic bias.

(JMIR Mhealth Uhealth 2020;8(10):e17377) doi:10.2196/17377

KEYWORDS

eHealth; compliance; pregnancy; digital assessments

XSL•FO
RenderX

Introduction

Pregnant women make up a significant proportion of the world's population. With an average age of 30 years, pregnant women represent a generation of patients eager to experience new technologies and extend medical care in the digital sector [1-3]. The use of pregnancy apps among expectant mothers is high; such apps can therefore either be used as educational devices or as monitoring instruments during pregnancy, based on regular assessments of patient-reported outcomes (PROs), as well as psychological and physical symptoms [4,5].

A PRO is defined as "a measurement based on a report that comes directly from the patient (i.e., study subject) about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else" [6]. Many factors such as body mass index (BMI), physical symptoms, or even depressive symptoms can be assessed using PROs or validated questionnaires. Self-reporting of information concerning delivery or prepregnancy weight were reported to be of high accuracy [7,8].

Furthermore, digital monitoring devices could offer a benefit to women at risk for preterm birth due to a reduction of adverse fetal outcomes and costs [9]. Electronic health solutions offer not only the possibility of self-tracking, but also health education. Digital self-tracking of uterine activity has already been shown to prolong pregnancy and improve outcomes for the baby. Online education on smoking and nutrition can also have significant impact on patients' behavior [9,10]. Thus, digital solutions during pregnancy have the potential to improve efficiency and quality of care [11]. With the growing demand of easily accessible educational health information, tailored interventions, and more personalization, digital health apps offer unique opportunities and may be beneficial to treatment compliance [1,12]. In addition to that, there have been reports of higher use of pregnancy apps among women with depression, a history of chronic illness, or other risk factors for adverse outcomes, including smoking [3,13,14].

Thus, electronic health (eHealth) apps seem to be the perfect fit for obstetrics, with great potential for modifying the structure of perinatal care [15]. Several online information platforms or pregnancy apps for expectant mothers and their partners exist at present, providing information as well as online coaching, with good rates of compliance and success in lifestyle interventions [16]. However, up to now, there have been no digital prevention programs routinely integrated into antenatal care.

To achieve successful integration into routine care, high patient engagement, also referred to as compliance, remains a key factor. Compliance is defined as the "the consistency and accuracy with which a patient follows the regimen prescribed by a physician or other health care professional" [17]. In any digital health program, compliance is a key challenge and an essential factor for a successful outcome, as low compliance can threaten the validity of a study [18]. A substantial number of patients stop using apps before the completion of a program [19]. Several studies from different fields reported that compliance declined through the course of their study [20,21]. However, the definition of compliance is often not clarified or can vary greatly among different studies according to the study format. In addition to that, not all studies report on compliance, but rather dropout rates of presumably official study dropouts.

In addition to that, people from lower socioeconomic backgrounds or of minority ethnicities have been reported to have lower compliance rates [18,20-22]. There have also been reports that patients with poorer lifestyle and health profiles are less compliant [18].

To reach those marginal groups of patients at a presumed risk of adverse pregnancy outcomes, some mobile apps have already received broad acceptance as a way of low-threshold, interactive care [23]. In contrast, apps may not effectively engage "hard-to-reach" groups, such as women of low income and those with lower levels of education [24]. Nonetheless, incentives or reminders can be implemented to effectively increase compliance [25,26].

Therefore, this study aimed to examine compliance and its influencing factors among pregnant women in a prospective cohort study. This analysis focused on characterizing women with high engagement and investigating potential influencing factors to understand patients' engagement in digital apps as a key factor for successfully integrating eHealth in routine care settings.

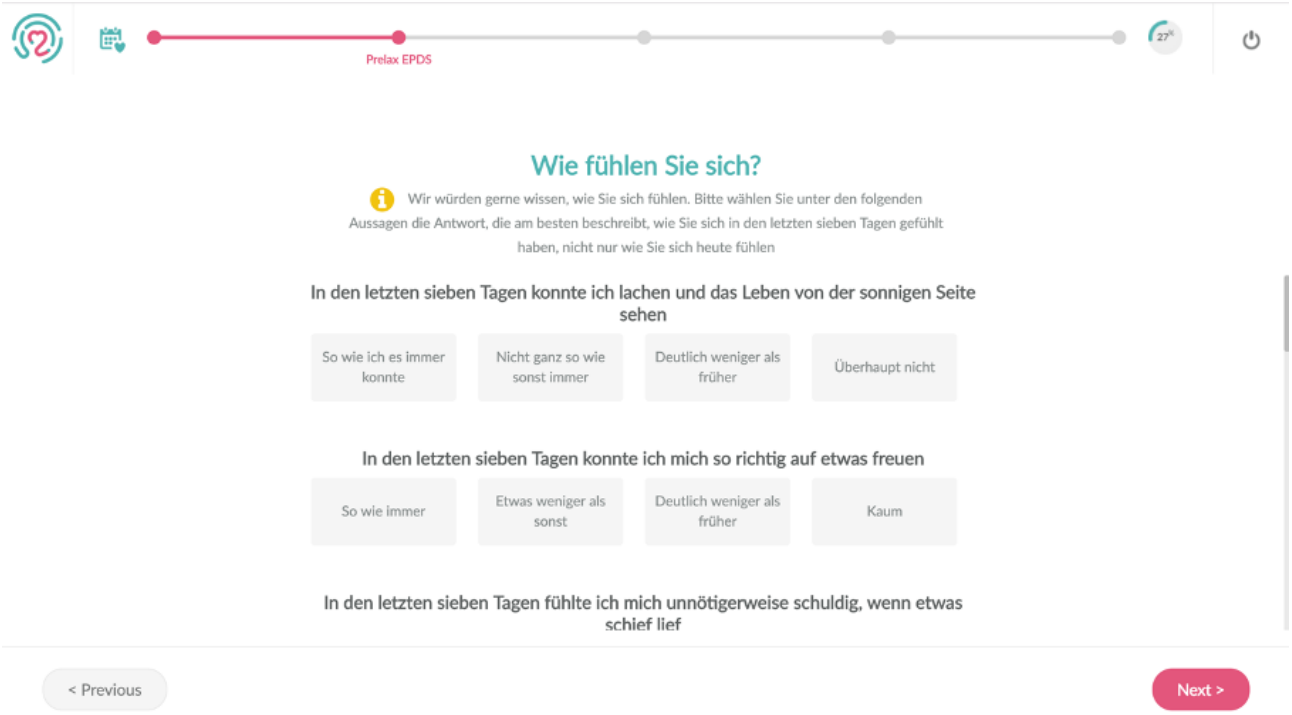
Methods

Participants and Study Design

This exploratory bicentric trial was conducted prospectively between October 2016 and September 2018 at two German university hospitals (University Hospitals of Heidelberg and Tuebingen). It was designed as a longitudinal, bicentric trial that included several questionnaires, which were delivered via an online platform called PiiA (Patient-informiert-interaktiv-Arzt; Figure 1) [27]. Originally, the study design included the randomization of the participants into two groups, with the intervention group having access to an educational pregnancy guidebook in addition to the survey. However, due to the low usage of the app, we refrained from a direct group comparison and treated both groups as one study cohort.

XSL•FO
RenderX

Figure 1. Screenshot of the PiiA platform. PiiA: Patient-informiert-interaktiv-Arzt (patient-informs-interactive-physician).



The online study visits and the corresponding individual questionnaires were scheduled monthly until delivery, followed by 3 postnatal visits finishing at 6 months postpartum. In addition to validated questionnaires, various pregnancy-related symptoms and complications were surveyed as part of a self-tracking app. The questionnaires included in this analysis are shown in Table 1.

Table 1. Structure of the prenatal section.

Characteristics and questionnaires	Baseline	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
Gestational age (weeks)	19	20	24	28	32	36
Socioeconomic questionnaire	✓					
Previous medical history	✓					
Physical symptoms		✓	✓	✓	✓	✓
Digital evaluation	✓			✓		
Edinburgh Postnatal Depression Scale		✓	✓	✓	✓	✓
State-Trait Anxiety Inventory			✓		✓	

Participants were enrolled during outpatient prenatal checkups or, in individual cases, during a hospital stay, and were eligible for participation if they were older than 18 years, they had a sufficient level of German language proficiency, they had internet access, and gestational age was between 19 and 27 weeks. Exclusion criteria were inability to understand the content of the study, as well as multiples or chromosomal aberrations and genetic conditions or fetal abnormalities of different or unknown origin.

All study-related contacts took place in addition to standard antenatal care. The study was entirely conducted in the German language. Ethics approval was granted by the Ethical Committee of the University of Heidelberg (S-158/2016) and the University of Tuebingen (062/2017BO2).

Initial Support and Online Tutorial

After giving informed consent, all patients received a tablet and were introduced to the platform by trained staff who were available for further questions. The baseline visit was filled out onsite. The platform also provided a tutorial that could be accessed by the participant at any time. After the first visit, patients continued the study at home using their own preferred devices. At first login, patients were asked to provide their current gestational week as a trigger for all following visits and reminders, ensuring the exact start of new visits on their respective gestational week.

Measurements

Demographics, Anamnesis, and Physical Complaints

The socioeconomic questionnaire (SEQ) is a self-designed questionnaire that encompasses several items related to

XSL•FO
RenderX

demographic characteristics such as age, origin, and education. For retrieving the participants' medical history, we used a self-designed questionnaire that included a series of obstetric questions on parity, preexisting conditions, and medical conditions in previous pregnancies. The questionnaire contains a selection of common complications and medical diagnoses. Questions concerning technical abilities and preferences were also administered. Furthermore, a questionnaire with a selection of pregnancy-related PRO symptoms was available at every visit as part of a self-tracking app.

Edinburgh Postnatal Depression Scale

Depressive symptoms were assessed using the Edinburgh Postnatal Depression Scale (EPDS) at every visit through an overall summary score. The EPDS was originally developed by Cox et al [28] and translated into German by Bergant et al [29]. The EPDS offers high sensitivity and specificity in predicting depressive disorders [29] and has proven to accomplish this in the prenatal and postnatal period [29,30].

State-Trait Anxiety Inventory

The State-Trait Anxiety Inventory (STAI) is a two-part questionnaire, each consisting of a summary score of 20 items, to evaluate anxiety as a temporary condition (state, STAI-S) and as a personality characteristic (trait, STAI-T) [31]. We used the German version of the questionnaire [31]. The STAI has proven to be a valid instrument for assessing anxiety in pregnant women [32]. We implemented the STAI at two visits.

Patient Engagement and Compliance Monitoring

A range of engagement strategies were employed, designed to encourage participants to continue the study. Participants were reminded via email 2 days prior to as well as 3 and 5 days after the scheduled visit date. The rate of completion for each assessment was calculated, measured by the amount of completed questions divided by all available questions at a specific visit. This rate was assessed weekly for each patient that had been reminded in the previous week. If the completion rate was ≤80%, email reminders were sent and escalated to follow-up phone calls if patients did not respond. Patients who officially dropped out of the study were considered as noncompliant patients and were included in the analysis.

Statistical Analysis

We used the programming language R (Version 3.5.1; R Foundation for Statistical Computing) for all of our analyses [33]. Socioeconomic and obstetric data as well as completion rates were analyzed descriptively by calculating mean scores and standard deviations as well as absolute and relative frequencies. We regarded a completion rate of ≥80% per visit throughout the study as compliant.

First, we performed univariate logistic regression to examine the effect of unique items on compliance. Multivariate logistic regression was then applied to identify influencing factors on compliance ≥80%, using all previously significant variables.

Second, we evaluated whether changes in mental or physical well-being prompted an abrupt discontinuance of study participation. Thereupon, we performed a one-sample *t* test on each change in EPDS score and the number of physical complaints between the last two visits before dropout (decline of completion rate ≥80% followed by no further activity). This analysis included both official dropouts as well as patients concluding the study without further contact or notice. *P* values ≤.05 were considered significant. Since this is an exploratory study, no adjusting for multiplicity was performed and *P* values have to be interpreted in a descriptive sense.

Results

Sample Characteristics

This analysis is based on 585 participants pregnant with singletons. In total, 41 patients (7.0%) actively decided to terminate the study before completion, mostly due to personal issues concerning time management, difficulties related to pregnancy, or family reasons. Overall, 319 (54.5%) participants stopped processing the online visits without officially withdrawing their participation in the study and did not respond to further contact attempts by the study staff. The sample characteristics are presented in Table 2.

XSL•FO
RenderX

Table 2. Demographic characteristics of the study population.

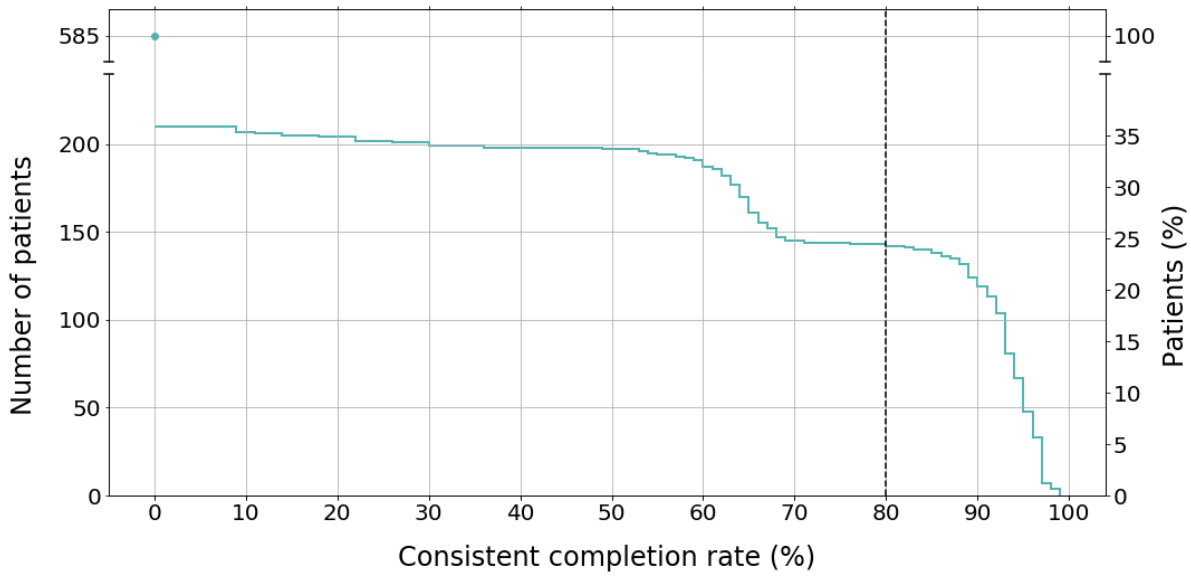
Variable	Frequency
Age (years), median (Q1-Q3)	33 (29-36)
BMI, n (%)	
<25	333 (60.5)
25-30	106 (19.3)
>30	111 (20.2)
Smoker, n (%)	
Current or former	207 (37.4)
Never	347 (62.6)
Origin, n (%)	
German	470 (83.9)
Other	90 (16.1)
Education, n (%)	
University entrance qualification or higher	239 (42.7)
Lower than university entrance qualification	321 (57.3)
Net monthly income per household (€), n (%)	
<1000 (<US $1184)	75 (13.8)
1000-2000 (US $1184-$2367)	176 (32.4)
2000-3000 (US $2367-$3551)	113 (20.8)
>3000 (>US $3551)	180 (33.1)
Current employment, n (%)	
Yes	469 (83.9)
No	90 (16.1)
Health insurance, n (%)	
Public	435 (77.7)
Private	125 (22.3)
First time pregnancy, n (%)	
No	220 (39.7)
Yes	334 (60.3)

Compliance Evaluation: Descriptive Compliance Characteristics

When applying the definition of a completion rate of ≥80% per visit throughout the study as compliant participation, n=148 patients could be considered as compliant during the prenatal stage of the study, which corresponds to only about 25% of all enrolled participants (Figure 2). Another 63% left at least one visit blank.

XSL•FO
RenderX

Figure 2. Relationship between completion rate continuity and number of participants.



As depicted in Table 3, the desired completion rate of 80% to 100% was met by a high number of patients at baseline but decreased significantly in the following assessments. The analysis showed a spike for completion rates of 80% to 100% and 0% to 19%. The range in between these two extremes is evenly distributed and low at every visit analyzed, indicating a tendency to either fill out the entire questionnaire repertoire or not to start the visit altogether.

Table 3. Completion rates per visit in relation to the number of participants per visit.

Compliance per visit	Baseline, n (%)	Visit 1, n (%)	Visit 2, n (%)	Visit 3, n (%)	Visit 4, n (%)	Visit 5, n (%)
0%-19%	25 (4.3)	56 (9.6)	216 (36.9)	256 (43.8)	284 (48.5)	321 (54.8)
20%-39%	2 (0.3)	11 (1.9)	6 (1.0)	2 (0.3)	0 (0.0)	8 (1.4)
40%-59%	11 (1.9)	18 (3.1)	11 (1.9)	4 (0.7)	2 (0.3)	8 (1.4)
60%-79%	20 (3.4)	31 (5.3)	17 (2.9)	6 (1.0)	10 (1.7)	40 (6.8)
80%-100%	527 (90.1)	469 (80.2)	335 (57.3)	317 (54.2)	289 (49.4)	208 (35.6)

Influencing Factors for High Compliance in Online-Based Surveys During Pregnancy

We performed logistic regression analyses on characteristics that could potentially influence compliance (Table 4). Items referring to socioeconomic status appeared to have significant influence; univariate logistic regression analysis showed a positive relationship between higher education (university entrance qualification or higher, P=.007), public health insurance (P=.01), and origin (other versus German, P<.001), as well as a negative correlation with smoking (P=.008, current or former smoker versus never smoked). Further medical factors, such as a history of previous abortions or BMI did not show significant correlation with the outcome. Neither computer skills nor the personal review of the user-friendliness of the website had an effect on compliance. Further correlations are depicted in Table 4.

In general, mental health characteristics had no influence on the patients' compliance. To a small extent, trait anxiety seemed to have an effect, as the STAI-T score at Visit 2 showed a possible negative relationship between trait anxiety and high compliance (P=.08). The EPDS score at baseline was not significantly related to the outcome (P=.78). Additionally, the intention to deliver via Cesarean section at Visit 1 presented significant negative influence (P=.04). The questionnaire characteristics are portrayed in Table 4.

XSL•FO
RenderX

Table 4. Univariate logistic regression on high compliance.

Variable	Effect (95% CI)	*P* value
Socioeconomic characteristics		
Age	0.016 (–0.02 to 0.054)	.41
BMI	0.003 (–0.027 to 0.031)	.84
Smoker	–0.551 (–0.967 to –0.148)	.008
Origin	1.333 (0.667 to 2.116)	<.001
Education	0.521 (0.143 to 0.901)	.007
Income	0.542 (–0.065 to 1.188)	.09
Health insurance	0.56 (0.127 to 0.986)	.01
Employment	–0.189 (–0.75 to 0.34)	.5
Relationship status	0.262 (–0.154 to 0.695)	.23
Anamnesis		
Previous birth	–0.234 (–0.614 to 0.15)	.23
Previous miscarriage	0.061 (–0.359 to 0.471)	.77
Maternal diseases	–0.198 (–0.875 to 0.425)	.55
Complications	0.034 (–0.346 to 0.41)	.86
Desire for Cesarian birth	–0.742 (–1.509 to –0.069)	.04
Sport or birth course	0.199 (–0.221 to 0.565)	.4
Technical details		
Technical skills	0.127 (–0.372 to 0.653)	.63
User friendly	0.068 (–0.05 to 0.19)	.26
Medium	–0.187 (–0.499 to 0.12)	.24
Patient report outcomes		
Edinburgh Postnatal Depression Scale result from baseline	–0.007 (–0.055 to 0.04)	.78
State-Trait Anxiety Inventory (Trait) result from Visit 2	–0.017 (–0.036 to 0.002)	.08
State-Trait Anxiety Inventory (State) result from Visit 2	–0.019 (–0.041 to 0.004)	.11

Based on the observed correlations, we performed multivariate logistic regression on the factors influencing high compliance in online-based surveys during pregnancy using the variables with significant influence (P<.05) from our univariate analysis (Table 4). We determined that German origin (P=.006) and smoking (P=.02) were the remaining statistically significant variables of our model (Table 5).

Table 5. Multivariate logistic regression on high compliance.

Variable	Effect (95% CI)	*P* value
Smoker	–0.538 (–1.001 to –0.088)	.02
Origin	1.099 (0.371 to 1.951)	.006
Education	0.225 (–0.235 to 0.663)	.35
Health insurance	0.248 (–0.268 to 0.755)	.34
Desire for Cesarian birth	–0.597 (–1.381 to 0.1)	.11

Psychometric Implications of Study Dropout

No significant influence of a change in the EPDS score (P=.64) or in the number of physical complaints (P=.2) on study termination was found.

Discussion

Principal Findings

In this study, we analyzed compliance patterns and determining factors for study compliance among pregnant women in a digital, web-based setting.

XSL•FO
RenderX

Only approximately 25% of patients were highly engaged in the digital app, defined by a completion rate of ≥80% at every visit.

Definitions of compliance vary from number of compliant patients, number of completed visits, to transmission rates; each definition only fits a particular study structure [26,34-36]. We aimed to implement a definition of compliance that included only complete data sets. Therefore, we considered an individual completion rate ≥80% at all visits as compliant. Similar values have been used or reported in other studies, so we considered it to be an appropriate margin [35,37-39].

About 25% of the participants met the aforementioned criteria for high compliance. This compliance rate is comparable to similar studies, which reported compliance rates from 21.8% to 35.6% among pregnant women [40,41]. However, other pregnancy-related studies showed higher compliance ranging from 41% to 92% [22,42,43]. In contrast, at 7%, the number of official dropouts in our study is lower than in comparable studies [16,44,45].

The range between the two extremes of compliance was evenly distributed, indicating patients either accomplished full compliance (80% to 100%) or did not fill out anything (0% to 19%). Technical questions, content-related questions, or difficulties in filling out the online assessment did not seem to play a role in participants' compliance. However, noncompliant patients may feel overwhelmed with the assessment and may not even start it because of the aforementioned challenges. In this study, we actively encouraged the patients to participate via reminders and asked noncompliant patients over the phone whether they were having problems with the study content or the online portal. In future studies, even more focus must be placed on motivating the patients to simply start the study visit, as they are then more likely to complete the entire visit.

We found a steadily decreasing compliance rather than an abrupt dropout of participants, which could have been provoked by a specific factor that all participants were exposed to at the same time (ie, a shutdown of the website). Other studies have also reported continuously declining compliance rates over time [20,21]. The steadily declining compliance rate can be attributed to a decreasing motivation instead of a low level of initiative or motivation at the beginning of the study, since the number of study participants at baseline was relatively high.

As our key finding, this study revealed that pregnant women who highly engaged with the digital platform had a distinctive profile. We found higher compliance rates in patients with presumed advantageous social standing, which was determined by responses that indicated higher education, higher income, nonsmoking status, private insurance, and being of German origin. Higher socioeconomic advantages such as a certain ethnicity or higher education have previously been reported to be associated with higher compliance [18,20-22]. In Germany, immigrants are still more likely to have a lower socioeconomic status and lower education [46,47]. Hence, our finding reflects a known imbalance for health-related aspects. Before the implementation of a digital risk assessment tool, the needs and expectations of these groups need to be assessed to offer high-quality digital health care to all patients.

Previous and current smoking was also associated with lower compliance. An influence of smoking rate on low compliance has previously been observed in adolescents [48]. Smoking could indicate a limited perception for personal health risks and hence explain a lower interest in health-related studies [49]. A poorer lifestyle and health profile have also been reported to be associated with lower compliance [18].

However, the impact of multicollinearity cannot be completely discounted in socioeconomic analyses. A higher level of education often goes hand in hand with a higher income, a lower likelihood of smoking, and private insurance, which is only available to higher income citizens in Germany [50,51].

Factors influencing compliance in medical research during pregnancy are therefore mostly static in nature and not necessarily pregnancy-specific. Individual solutions may have to be found to improve compliance with digital health solutions.

We also found that a request for a Cesarean section, which was expressed in the second trimester, was associated with lower compliance. Such patients often have misconceptions and a lack of accurate knowledge about different modes of delivery [52,53]. As seen above, our study sample presents with higher levels of education, which may emphasize a lack of accurate information among those with lower education concerning this matter.

Preexisting computer skills or the accessing device (smartphone versus computer) did not show significant influence on user compliance. However, these questions were only implemented at Visit 3, offering limited insight into technical difficulties.

We also evaluated psychometric factors on compliance. The EPDS at baseline and STAI-S at Visit 2 did not have a significant impact on compliance, indicating no detectable correlation between depressive symptoms or state-anxiety. Likewise, Wright et al [21] reported that anxiety, depression, or quality of life did not have an effect on compliance. However, other studies observed higher retention of patients with anxiety or positive affect [22,48], demonstrating the contrasting stance of literature on this topic. Nonetheless, we found a weak link between higher STAI-T scores at Visit 2 and lower compliance (P=.08), suggesting a higher burden of the study in patients with preexisting anxious tendencies.

In the multivariate regression model, German origin and being a nonsmoker were revealed to be the influential factors contributing to high compliance. Given the literature and the aforementioned factors on demographic and socioeconomic influence on study compliance [20,48], this result stresses the potential preexisting catering of medical apps to patients of higher socioeconomic standing. As our findings are in accordance with previous research, future studies need to focus on examining the needs of socially underprivileged groups and immigrants in regard to medical studies and online apps. In addition, studies should generally report on the demographic profile of noncompliers as noncompliance can generally cause a bias in results [54]. Customizing educational health apps to patients with lower educational background or offering a service in additional languages may assist in reaching these hard-to-reach groups.

XSL•FO
RenderX

A change in depressive symptoms or an increase in physical complaints had no effect on the abrupt termination of the study, which corresponds to the stance that a poor psychological condition does not promote high compliance [21]. We believe acute dropouts without further notice are a significant problem for studies like this, which was our motivation for an examination of this matter. Therefore, the compliance-determining factors appear to be static characteristics rather than influenceable developments or conditions. We recommend focusing on examining the needs and expectations of participants with a lower socioeconomic background or those of non-German origin to provide all patients with high-quality eHealth solutions.

Limitations

First of all, with country of origin being the strongest influencing variable, we are faced with the problem of the ambiguous meaning of that variable. We presented this question as "origin" and offered different countries of origin that are common in Germany. We focused on whether or not the participant considered their country of origin to be Germany, which can reflect socioeconomic disadvantages and possibly a lack of pregnancy-related apps in languages other than German [46,47]. In addition to that, the study was only conducted in German, which excluded some demographic groups. However, as a sufficient level of the German language was included in the criteria of eligibility, the influence of a language barrier was minimized. This variable reflects possible disadvantages of immigrants regarding medical studies.

Even though we mainly enrolled women with low-risk pregnancies, the study enrolment took place in a university hospital, which may lead to selection bias of patients with a higher risk for adverse physical or mental health outcomes. Additionally, most of the participants live in the areas surrounding Heidelberg and Tuebingen, which are known to be regions of higher income [55]. This is reinforced by the relatively high number of patients with private health insurance in our study population [56].

Moreover, as with any digital tool, our platform experienced some technical difficulties, mostly due to individual browser settings. These problems were minor and were solved quickly when reported. However, we cannot completely eliminate the possibility of technical difficulties that were not brought to our attention by the participants and may have caused a decline in compliance.

Conclusions

In conclusion, our results demonstrate that, despite standardized reminders and motivational support, maintaining a high rate of compliance in digital health assessments with pregnant patients over a longer period of time remains challenging. As we observed static characteristics such as socioeconomic background and German origin to be the most influential factors, we propose seeking new ways to reach these groups (e.g., by providing surveys in additional languages and targeting the individuality of every single patient). Further research needs to be directed toward a better understanding of the needs and expectations of socially disadvantaged groups with regard to digital apps.

Conflicts of Interest

None declared.

References

1. Goetz M, Müller M, Matthies LM, Hansen J, Doster A, Szabo A, et al. Perceptions of Patient Engagement Applications During Pregnancy: A Qualitative Assessment of the Patient's Perspective. JMIR mHealth uHealth 2017 May 26;5(5):e73 [FREE Full text] [doi: 10.2196/mhealth.7040] [Medline: 28550005]
2. Geburten. Destatis - Statistisches Bundesamt. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Geburten/_inhalt.html [accessed 2020-05-28]
3. Shen C, Wang MP, Chu JT, Wan A, Viswanath K, Chan SSC, et al. Health App Possession Among Smartphone or Tablet Owners in Hong Kong: Population-Based Survey. JMIR mHealth uHealth 2017 Jun 05;5(6):e77. [doi: 10.2196/mhealth.7628] [Medline: 28583905]
4. Wallwiener S, Müller M, Doster A, Laserer W, Reck C, Pauluschke-Fröhlich J, et al. Pregnancy eHealth and mHealth: user proportions and characteristics of pregnant women using Web-based information sources-a cross-sectional study. Arch Gynecol Obstet 2016 Apr 15;294(5):937-944. [doi: 10.1007/s00404-016-4093-y] [Medline: 27084763]
5. Mo Y, Gong W, Wang J, Sheng X, Xu DR. The Association Between the Use of Antenatal Care Smartphone Apps in Pregnant Women and Antenatal Depression: Cross-Sectional Study. JMIR mHealth uHealth 2018 Nov 29;6(11):e11508 [FREE Full text] [doi: 10.2196/11508] [Medline: 30497996]
6. Patient-Reported Outcomes (PROs) in Medical Device Decision Making. URL: https://www.fda.gov/about-fda/cdrh-patient-engagement/patient-reported-outcomes-pros-medical-device-decision-making [accessed 2020-05-28]
7. Elliott JP, Desch C, Istwan NB, Rhea D, Collins AM, Stanziano GJ. The Reliability of Patient-Reported Pregnancy Outcome Data. Population Health Management 2010 Feb;13(1):27-32. [doi: 10.1089/pop.2009.0008]
8. Han E, Abrams B, Sridhar S, Xu F, Hedderson M. Validity of Self-Reported Pre-Pregnancy Weight and Body Mass Index Classification in an Integrated Health Care Delivery System. Paediatr Perinat Epidemiol 2016 Mar 09;30(4):314-319. [doi: 10.1111/ppe.12286] [Medline: 26961120]

XSL•FO
RenderX

9. Lanssens D, Vandenberk T, Thijs IM, Grieten L, Gyselaers W. Effectiveness of Telemonitoring in Obstetrics: Scoping Review. J Med Internet Res 2017 Sep 27;19(9):e327 [FREE Full text] [doi: 10.2196/jmir.7266] [Medline: 28954715]
10. Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth Platform for Pregnancy on Nutrition and Lifestyle of the Reproductive Population: A Survey. JMIR mHealth uHealth 2016 May 27;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197]
11. Dudenhausen JW. Primary prevention of preterm birth. J Perinat Med 2014 Jul;42(4):431-433. [doi: 10.1515/jpm-2014-0036] [Medline: 24995401]
12. Drewes C, Kirkovits T, Schiltz D, Schinkoethe T, Haidinger R, Goldmann-Posch U, et al. EHealth Acceptance and New Media Preferences for Therapy Assistance Among Breast Cancer Patients. JMIR Cancer 2016 Sep 14;2(2):e13 [FREE Full text] [doi: 10.2196/cancer.5711] [Medline: 28410189]
13. Mo Y, Gong W, Wang J, Sheng X, Xu DR. The Association Between the Use of Antenatal Care Smartphone Apps in Pregnant Women and Antenatal Depression: Cross-Sectional Study. JMIR mHealth uHealth 2018 Nov 29;6(11):e11508 [FREE Full text] [doi: 10.2196/11508] [Medline: 30497996]
14. Krishnamurti T, Davis AL, Wong-Parodi G, Fischhoff B, Sadovsky Y, Simhan HN. Development and Testing of the MyHealthyPregnancy App: A Behavioral Decision Research-Based Tool for Assessing and Communicating Pregnancy Risk. JMIR mHealth uHealth 2017 Apr 10;5(4):e42 [FREE Full text] [doi: 10.2196/mhealth.7036]
15. van den Heuvel JF, Groenhof TK, Veerbeek JH, van Solinge WW, Lely AT, Franx A, et al. eHealth as the Next-Generation Perinatal Care: An Overview of the Literature. J Med Internet Res 2018 Jun 05;20(6):e202 [FREE Full text] [doi: 10.2196/jmir.9262] [Medline: 29871855]
16. Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth Platform for Pregnancy on Nutrition and Lifestyle of the Reproductive Population: A Survey. JMIR mHealth uHealth 2016;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
17. Compliance. The Free Dictionary by Farlex. 2020. URL: https://medical-dictionary.thefreedictionary.com/compliance [accessed 2020-06-15]
18. Tolonen H, Dobson A, Kulathinal S. Effect on Trend Estimates of the Difference between Survey Respondents and Non-respondents: Results from 27 Populations in the WHO MONICA Project. Eur J Epidemiol 2005 Nov;20(11):887-898. [doi: 10.1007/s10654-005-2672-5]
19. Pedersen DH, Mansourvar M, Sortsø C, Schmidt T. Predicting Dropouts From an Electronic Health Platform for Lifestyle Interventions: Analysis of Methods and Predictors. J Med Internet Res 2019 Sep 04;21(9):e13617 [FREE Full text] [doi: 10.2196/13617] [Medline: 31486409]
20. Land SR, Ritter MW, Costantino JP, Julian TB, Cronin WM, Haile SR, et al. Compliance With Patient-Reported Outcomes in Multicenter Clinical Trials: Methodologic and Practical Approaches. JCO 2007 Nov 10;25(32):5113-5120. [doi: 10.1200/jco.2007.12.1749]
21. Wright E, Selby P, Crawford M, Gillibrand A, Johnston C, Perren T, et al. Feasibility and Compliance of Automated Measurement of Quality of Life in Oncology Practice. JCO 2003 Jan 15;21(2):374-382. [doi: 10.1200/jco.2003.11.044]
22. Lara MA, Navarro C, Navarrete L, Le HN. Retention rates and potential predictors in a longitudinal randomized control trial to prevent postpartum depression. Salud Mental 2010;33(5):429-436 [FREE Full text]
23. Krishnamurti T, Davis AL, Wong-Parodi G, Fischhoff B, Sadovsky Y, Simhan HN. Development and Testing of the MyHealthyPregnancy App: A Behavioral Decision Research-Based Tool for Assessing and Communicating Pregnancy Risk. JMIR mHealth uHealth 2017 Apr 10;5(4):e42 [FREE Full text] [doi: 10.2196/mhealth.7036] [Medline: 28396302]
24. Shen C, Wang MP, Chu JT, Wan A, Viswanath K, Chan SSC, et al. Health App Possession Among Smartphone or Tablet Owners in Hong Kong: Population-Based Survey. JMIR mHealth uHealth 2017 Jun 05;5(6):e77 [FREE Full text] [doi: 10.2196/mhealth.7628] [Medline: 28583905]
25. Morgan AJ, Rapee RM, Bayer JK. Increasing response rates to follow-up questionnaires in health intervention research: Randomized controlled trial of a gift card prize incentive. Clinical Trials 2017 Apr 08;14(4):381-386. [doi: 10.1177/1740774517703320]
26. Homko CJ, Deeb LC, Rohrbacher K, Mulla W, Mastrogiannis D, Gaughan J, et al. Impact of a Telemedicine System with Automated Reminders on Outcomes in Women with Gestational Diabetes Mellitus. 2012 Jul;14(7):624-629. [doi: 10.1089/dia.2012.0010]
27. Wallwiener M, Matthies L, Simoes E, Keilmann L, Hartkopf AD, Sokolov AN, et al. Reliability of an e-PRO Tool of EORTC QLQ-C30 for Measurement of Health-Related Quality of Life in Patients With Breast Cancer: Prospective Randomized Trial. J Med Internet Res 2017 Sep 14;19(9):e322 [FREE Full text] [doi: 10.2196/jmir.8210] [Medline: 28912116]
28. Cox JL, Holden JM, Sagovsky R. Detection of Postnatal Depression. Br J Psychiatry 1987;150(6):782-786. [doi: 10.1192/bjp.150.6.782]
29. Bergant AM, Nguyen T, Heim K, Ulmer H, Dapunt O. Deutschsprachige Fassung und Validierung der Edinburgh postnatal depression scale. Dtsch med Wochenschr 2008 Mar 25;123(03):35-40. [doi: 10.1055/s-2007-1023895]

XSL•FO
RenderX

30. Rubertsson C, Börjesson K, Berglund A, Josefsson A, Sydsjö G. The Swedish validation of Edinburgh Postnatal Depression Scale (EPDS) during pregnancy. Nordic Journal of Psychiatry 2011 Jul 05;65(6):414-418. [doi: 10.3109/08039488.2011.590606]
31. Laux L, Glanzmann P, Schaffner P, Spielberger C. Das State-Trait-Angstinventar. Beltz Test. Weinheim 1981.
32. Grant K, McMahon C, Austin M. Maternal anxiety during the transition to parenthood: A prospective study. Journal of Affective Disorders 2008 May;108(1-2):101-111. [doi: 10.1016/j.jad.2007.10.002]
33. R Core Team. R: A Language and Environment for Statistical Computing. The R Foundation. Vienna, Austria: R Foundation for Statistical Computing URL: https://www.R-project.org
34. Osoba D, Zee B. Completion rates in health-related quality-of-life assessment: approach of the National Cancer Institute of Canada Clinical Trials Group. Statist Med 1998 Mar 15;17(5-7):603-612. [doi: 10.1002/(sici)1097-0258(19980315/15)17:5/7<603::aid-sim807>3.0.co;2-d]
35. Basch E, Iasonos A, Barz A, Culkin A, Kris MG, Artz D, et al. Long-Term Toxicity Monitoring via Electronic Patient-Reported Outcomes in Patients Receiving Chemotherapy. JCO 2007 Dec 01;25(34):5374-5380. [doi: 10.1200/jco.2007.11.2243]
36. Homko CJ, Santamore WP, Whiteman V, Bower M, Berger P, Geifman-Holtzman O, et al. Use of an Internet-Based Telemedicine System to Manage Underserved Women with Gestational Diabetes Mellitus. 2007 Jun;9(3):297-306. [doi: 10.1089/dia.2006.0034]
37. Atherton PJ, Burger KN, Pederson LD, Kaggal S, Sloan JA. Patient-reported outcomes questionnaire compliance in Cancer Cooperative Group Trials (Alliance N0992). Clin Trials 2016 Dec;13(6):612-620 [FREE Full text] [doi: 10.1177/1740774516655101] [Medline: 27365012]
38. Cuffe S, Hon H, Tobros K, Espin-Garcia O, Brhane Y, Harland L, et al. Cancer patients' acceptability of incorporating an epidemiology questionnaire within a clinical trial. Clin Trials 2015 Jun;12(3):237-245. [doi: 10.1177/1740774514568689] [Medline: 25633805]
39. Judson TJ, Bennett AV, Rogak LJ, Sit L, Barz A, Kris MG, et al. Feasibility of Long-Term Patient Self-Reporting of Toxicities From Home via the Internet During Routine Chemotherapy. JCO 2013 Jul 10;31(20):2580-2585. [doi: 10.1200/jco.2012.47.6804]
40. Homko CJ, Deeb LC, Rohrbacher K, Mulla W, Mastrogiannis D, Gaughan J, et al. Impact of a telemedicine system with automated reminders on outcomes in women with gestational diabetes mellitus. Diabetes Technol Ther 2012 Jul;14(7):624-629 [FREE Full text] [doi: 10.1089/dia.2012.0010] [Medline: 22512287]
41. Homko CJ, Santamore WP, Whiteman V, Bower M, Berger P, Geifman-Holtzman O, et al. Use of an internet-based telemedicine system to manage underserved women with gestational diabetes mellitus. Diabetes Technol Ther 2007 Jun;9(3):297-306. [doi: 10.1089/dia.2006.0034] [Medline: 17561800]
42. Renault KM, Nørgaard K, Nilas L, Carlsen EM, Cortes D, Pryds O, et al. The Treatment of Obese Pregnant Women (TOP) study: a randomized controlled trial of the effect of physical activity intervention assessed by pedometer with or without dietary intervention in obese pregnant women. American Journal of Obstetrics and Gynecology 2014 Feb;210(2):e1-e9. [doi: 10.1016/j.ajog.2013.09.029]
43. Smuts CM, Huang M, Mundy D, Plasse T, Major S, Carlson SE. A Randomized Trial of Docosahexaenoic Acid Supplementation During the Third Trimester of Pregnancy. 2003;101(3):469-479. [doi: 10.1097/00006250-200303000-00011]
44. Kenter R, Warmerdam L, Brouwer-Dudokdewit C, Cuijpers P, van SA. Guided online treatment in routine mental health care: an observational study on uptake, drop-out and effects. BMC Psychiatry 2013 Jan;13:43 [FREE Full text] [doi: 10.1186/1471-244X-13-43] [Medline: 23368894]
45. Hildingsson I, Haines H, Karlström A, Nystedt A. Presence and process of fear of birth during pregnancy—Findings from a longitudinal cohort study. Women and Birth 2017 Oct;30(5):e242-e247. [doi: 10.1016/j.wombi.2017.02.003]
46. Armutsgefährdungsquote nach Migrationshintergrund und ausgewählten Merkmalen. Desatis - Statistisches Bundesamt. 2018. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/migrationshintergrund-armutsgefaehrdung.html [accessed 2020-05-18]
47. Bevölkerung in Privathaushalten nach Migrationshintergrund und höchstem allgemeinem Schulabschluss. Destatis - Statistisches Bundesamt. 2018. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/migrationshintergrund-schulabschluss.html [accessed 2020-05-18]
48. Sokolovsky AW, Mermelstein RJ, Hedeker D. Factors Predicting Compliance to Ecological Momentary Assessment Among Adolescent Smokers. 2013 Oct 04;16(3):351-358. [doi: 10.1093/ntr/ntt154]
49. Weinstein ND. Accuracy of smokers' risk perceptions. Ann Behav Med 1998;20(2):135-140. [doi: 10.1007/BF02884459] [Medline: 9989319]
50. Villalbí JR, Salvador J, Cano-Serral G, Rodríguez-Sanz MC, Borrell C. Maternal smoking, social class and outcomes of pregnancy. Paediatr Perinat Epidemiol 2007 Sep;21(5):441-447. [doi: 10.1111/j.1365-3016.2007.00845.x]
51. Wer kann sich privat versichern? PKV-Verband der Privaten Krankenversicherung. URL: https://www.pkv.de/themen/krankenversicherung/so-funktioniert-die-pkv/wer-kann-sich-privat-versichern/ [accessed 2020-06-08]
52. Sanavi FS, Rakhshani F, Ansari-Moghaddam A, Edalatian M. Reasons for Elective Cesarean Section amongst Pregnant Women: A Qualitative Study. J Reprod Infertil 2012 Oct;13(4):237-240 [FREE Full text] [Medline: 23926552]

XSL•FO
RenderX

53. McCourt C, Weaver J, Statham H, Beake S, Gamble J, Creedy DK. Elective Cesarean Section and Decision Making: A Critical Review of the Literature. Birth 2007 Mar;34(1):65-79. [doi: 10.1111/j.1523-536x.2006.00147.x]
54. O'Neill RT, Temple R. The Prevention and Treatment of Missing Data in Clinical Trials: An FDA Perspective on the Importance of Dealing With It. Clin Pharmacol Ther 2012 Feb 08;91(3):550-554. [doi: 10.1038/clpt.2011.340]
55. Nettoeinkommen und verfügbares Nettoeinkommen privater Haushalte im Monat nach Bundesländern. Statista. URL: https://de.statista.com/statistik/daten/studie/5758/umfrage/verfuegbares-nettoeinkommen-nach-bundeslaendern/ [accessed 2019-07-15]
56. GKV-Spitzenverband 2019. GKV-Spitzenverband. URL: https://www.gkv-spitzenverband.de/gkv_spitzenverband/presse/zahlen_und_grafiken/zahlen_und_grafiken.jsp [accessed 2019-07-28]

Abbreviations

eHealth: electronic health
EPDS: Edinburgh Postnatal Depression Scale
PiiA: “Patient-informiert-interaktiv-Arzt” (patient-informs-interactive-physician)
PRO: patient-reported outcome
SEQ: socioeconomic questionnaire
STAI: State-Trait Anxiety Inventory

Edited by G Eysenbach, L Buis; submitted 10.12.19; peer-reviewed by MW Haun, G Tzilos Wernette, E Ding; comments to author 18.05.20; revised version received 30.07.20; accepted 10.08.20; published 14.10.20.

Please cite as:
Brusniak K, Arndt HM, Feisst M, Haßdenteufel K, Matthies LM, Deutsch TM, Hudalla H, Abele H, Wallwiener M, Wallwiener S
Challenges in Acceptance and Compliance in Digital Health Assessments During Pregnancy: Prospective Cohort Study
JMIR Mhealth Uhealth 2020;8(10):e17377
URL: https://mhealth.jmir.org/2020/10/e17377
doi:10.2196/17377
PMID:33052134

©Katharina Brusniak, Hannah Maria Arndt, Manuel Feisst, Kathrin Haßdenteufel, Lina Maria Matthies, Thomas Maximilian Deutsch, Hannes Hudalla, Harald Abele, Markus Wallwiener, Stephanie Wallwiener. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 14.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Effects of a Brief Biofeedback Breathing Session Delivered Through the BioBase App in Facilitating Employee Stress Recovery: Randomized Experimental Study

Olga Chelidoni[1], BSc, MSc, PhD; David Plans[2,3,4], BA, MSc, PhD; Sonia Ponzo[4], BSc, MSc, PhD; Davide Morelli[4,5], BSc, MSc, PhD; Mark Cropley[6], BSc, PhD

[1]Evolution, Behaviour and Environment, School of Life Sciences, University of Sussex, Brighton, United Kingdom
[2]Initiative in the Digital Economy, Department of Science, Innovation, Technology, and Entrepreneurship, University of Exeter, Exeter, United Kingdom
[3]Social Cognition Lab, Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom
[4]BioBeats Group Ltd, London, United Kingdom
[5]Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Oxford, United Kingdom
[6]Faculty of Health and Medical Sciences, School of Psychology, University of Surrey, Guildford, United Kingdom

Corresponding Author:
Mark Cropley, BSc, PhD
Faculty of Health and Medical Sciences
School of Psychology
University of Surrey
Stag Hill
Guildford, GU2 7XH
United Kingdom
Phone: 44 01483686928
Email: mark.cropley@surrey.ac.uk

Abstract

Background: Recovery from stress is a predictive factor for cardiovascular health, and heart rate variability (HRV) is suggested to be an index of how well people physiologically recover from stress. Biofeedback and mindfulness interventions that include guided breathing have been shown to be effective in increasing HRV and facilitating stress recovery.

Objective: This study aims to assess the effectiveness of a brief app-based breathing intervention (BioBase) in enhancing physiological recovery among employees who were induced to cognitive and emotional stress.

Methods: In total, we recruited 75 full-time employees. Interbeat (RR) intervals were recorded continuously for 5 min at baseline and during cognitive and emotional stress induction. The session ended with a 5-min recovery period during which participants were randomly allocated into 3 conditions: app-based breathing (BioBase), mindfulness body scan, or control. Subjective tension was assessed at the end of each period.

Results: Subjective tension significantly increased following stress induction. HRV significantly decreased following the stress period. In the recovery phase, the root mean square of successive RR interval differences (P=.002), the percentage of successive RR intervals that differed by >50 ms (P=.008), and high frequency (P=.01) were significantly higher in the BioBase breathing condition than in the mindfulness body scan and the control groups.

Conclusions: Biofeedback breathing interventions digitally delivered through a commercially available app can be effective in facilitating stress recovery among employees. These findings contribute to the mobile health literature on the beneficial effects of brief app-based breathing interventions on employees' cardiovascular health.

(JMIR Mhealth Uhealth 2020;8(10):e19412) doi:10.2196/19412

KEYWORDS

breathing; biofeedback; smartphone; heart rate variability; recovery; mindfulness; stress; mobile phone

XSL•FO
RenderX

Introduction

Work environments have become more complex, and in the current scenario, there are incredible levels of job demands on employees. Increasing job demands are leading to unanimously negative trends toward occupational stress levels [1], which have been shown to result in damaging effects on employees' health and well-being. High occupational stress levels are strongly associated with cardiovascular disease [2,3], which is one of the leading causes of morbidity and mortality [4].

Work stress may affect cardiovascular health through the imbalance between two major branches of the autonomic nervous system (ANS). The ANS consists of 2 antagonistic components: the sympathetic system, responsible for energy mobilization (fight or flight response), and the parasympathetic system, associated with restorative function (rest and digest status) [5]. Under optimal conditions, these systems operate in balance. However, under extended periods of work stress, the sympathetic system is overactivated, producing prolonged stress responses and eventually delaying stress recovery [6]. In the long term, if prolonged activation persists, it is thought to cause *wear and tear* on the physiological systems as higher levels of sympathetic activation lead to exhaustion [7]. What seems to be key for improving cardiovascular health is not the reactivity to stress but how quickly physiological recovery from stress occurs [8].

Heart rate variability (HRV) is a prominent biomarker of cardiovascular health that reflects the ability to maintain balance between sympathetic and parasympathetic activity in the presence of a stressor [9]. HRV refers to the time variation between interbeat (RR) intervals and HRV has been suggested to be an index of self-regulatory strength [10] and affective stability [11]. Higher HRV is a signal of good physiological adaptation, whereas lower HRV is an indicator of abnormal and insufficient adaptation of the ANS [12]. Individuals with naturally lower HRV have been shown to take longer to recover following exposure to experimentally induced stressors [13]. Low HRV has also been found to be a predictor of cardiac events [14], hypertension [15], and increased morbidity and mortality [16].

Given that HRV is a crucial index of somatic health and self-regulation, it is imperative to investigate possible ways to improve it, and growing evidence suggests that biofeedback is a simple and efficient way to increase HRV [17]. Biofeedback involves processes that enable individuals to learn how to change their physiological activity to improve health and performance [18]. One of these processes is to help individuals change their physiological responses by directly teaching them how to breath at a specific rate [19]. Respiratory sinus arrhythmia (RSA) is the mechanism through which breathing increases HRV. RSA occurs when heart rate fluctuations are synchronized with breathing; heart rate slightly increases during inhalation and slightly decreases during exhalation [20]. By teaching individuals breathing maneuvers, they learn to maximize and achieve RSA [19]. With breathing and heart rate at corresponding rates, the baroreflex, an ANS reflex responsible for controlling important bodily functions such as blood pressure (BP) [21], is strengthened. Thus, RSA not only increases HRV but also strengthens the baroreflex [22] and improves the control of physiological stress responses.

Previous research has shown that biofeedback interventions can be effective in increasing HRV. Hassett et al [23] showed that 10 biofeedback training sessions produced a significant increase in HRV among patients with fibromyalgia. Large and acute HRV increases were also observed after a 10-week biofeedback training among depressed adults [24]. Similar HRV increases were observed after 6 weeks of biofeedback training and maintained for 12 weeks after program completion in patients with coronary artery disease [25]. Empirical evidence suggests that biofeedback interventions can also result in significant improvements in self-reported stress and anxiety [26]. A study conducted among medical staff showed that a 28-day biofeedback training program led to significant decreases in self-reported work stress [27], and similar training has been successfully used as a stress reduction method for hospital nurses [28].

Furthermore, even brief biofeedback sessions can have positive effects in reducing stress reactivity in healthy populations. For example, it has been shown that participants who practiced 30 min of slow breathing demonstrated increased HRV during stress induction [29]. It has also been shown that a single session of biofeedback breathing can increase HRV [30], whereas a single 15-min breathing session can improve HRV in response to laboratory-induced stress and during recovery [31].

Beneficial effects on HRV can also be found following mindfulness-based interventions. Mindfulness is the practice of paying attention to the present moment in a nonjudgmental way by disregarding distracting thoughts and focusing attention on breathing [32]. Being trained to direct attention to a single focus such as breathing, while inhibiting any intrusive thoughts, can help individuals stay connected to their bodily senses [33]. A body scan is a common mindfulness practice that typically involves guided instructions where listeners attend to various parts of their body and their breathing, gently observing these areas and allowing other thoughts to recede [34]. Ditto et al [34] showed that participants practicing the body scan displayed significantly greater increases in HRV and larger BP decreases compared with those engaging in other relaxing activities. Shearer et al [35] also showed that students who underwent a one-day mindfulness training course that included breathing exercises demonstrated higher HRV in anticipation and during a cognitive challenge. More importantly, brief body scan sessions have also been shown to significantly reduce emotional exhaustion symptoms in health care workers [36].

At present, the mass popularity of smartphone devices opens possibilities for the use of mobile apps as a useful platform for delivering these self-help interventions. Their high accessibility and ease of use offer great potential for mobile health growth. Guided breathing can now be incorporated in mobile apps and delivered at a relatively low cost [37]. For example, previous research has shown that a biofeedback app-based game was more effective in decreasing arousal during stress when compared with conventional relaxation techniques [38], and an 8-week app-based mindfulness meditation program has also

XSL•FO
RenderX

been shown to reduce stress in college students compared with controls [39].

Mobile apps have been used to modify stress responses, and recently there has been a surge in apps that aim to decrease workplace stress [40]. This is because many employees confront demanding tasks requiring high levels of cognitive effort on a daily basis. If tight deadlines and time pressure are added, then employees experience heightened levels of stress that elicit physiological arousal. It is also common for certain work situations involving high workload, heated interactions with supervisors, or discrepancies with colleagues to elicit negative emotions such as anger. These negative emotions, if not regulated efficiently, can trigger negative ruminative thoughts that, if persistent, can cause prolonged physiological activation [41], which in the long term can increase cardiovascular disease risk [16]. Therefore, it is imperative to find practical ways to help employees physiologically recover from the stress that work demands cause.

Most people who need such interventions are often those who are the most time deprived or busy. It is therefore important to develop parsimonious and easy-to-use interventions but also interventions that are evidence based. BioBase (BioBeats Ltd) is a commercially available mobile app used to modify breathing. Recently, breathing exercises delivered through BioBase were shown to produce significant HRV increases after stress exposure [42]. However, its effects on employee HRV have not been tested. This study therefore aims to examine the effects of a brief breathing session delivered through the BioBase app against a mindfulness body scan on HRV after stress exposure in a group of full-time employees. Participants were exposed to cognitive and emotional stress in the laboratory and then randomly allocated to 1 of 3 conditions: BioBase breathing app, mindfulness body scan, or a control condition.

It is intuitive that biofeedback and mindfulness share some degree of resemblance regarding the mechanisms they target; biofeedback improves breathing whereas mindfulness trains the individuals to focus their attention on bodily sensations such as breathing. Biofeedback is an active process that requires participants to achieve RSA by actively following specific audio and visual instructions and results in HRV changes very quickly [22]. However, a mindfulness body scan is a passive process involving audio instructions that require less active participant engagement, as its efficacy depends on participants' attention levels and is generally considered to be effective only after regular practice [43]. Despite the abovementioned differences between biofeedback and mindfulness interventions, no study to date has investigated their effectiveness in comparison with one another and a control condition. Hence, the aim of this study was to assess the impact of a biofeedback intervention, namely guided breathing, in reducing stress, indexed by HRV recovery, in comparison with a mindfulness intervention and a control condition. As biofeedback requires participants to focus their attention on both audio and visual instructions, it was predicted that—following stress exposure—individuals in the BioBase breathing condition would demonstrate greater HRV recovery, relative to those in the mindfulness body scan condition.

Methods

Participants

In total, 107 individuals volunteered to participate. Of these 107 participants, 15 did not meet the eligibility criteria; 6 were receiving psychotherapeutic treatment, 3 had a heart condition, 1 took prescribed medication, and 5 were working less than 30 hours per week. Of the remaining 92, the first 75 individuals were tested, as this sample size has revealed moderate-to-strong effect sizes in the past [42]. The final sample consisted of 75 full-time working adults (48 females; age range 18-62 years; mean age 32.32, SD 10.46 years). Participants' demographics such as work hours per week, level of education, years of work experience, and type of occupation are shown in Table 1. Single items were included asking participants if they were smokers, as frequent tobacco use has been associated with increased sympathetic activity resulting in higher heart rate (HR) [44]. Levels of fitness were also assessed as more fit individuals tended to have lower HR [45]. Smoking and fitness status for all participants is shown in Table 1. All participants were given a £10 (US $13) gift voucher for their participation. The University of Surrey Ethics Committee granted a favorable ethical opinion for this project (reference: UEC 2018 119 FHMS).

XSL•FO
RenderX

Table 1. Participant demographics (N=75).

Characteristics	Values, n (%)
Work hours per week	
31-40	65 (87)
41-50	9 (12)
>50	1 (1)
Education	
No formal education	1 (1)
Secondary education	13 (17)
University education	61 (82)
Work experience (years)	
Up to 1	33 (44)
1-5	26 (35)
>5	16 (21)
Occupation	
Marketing or advertising	10 (13)
Recruitment or Human Resources	9 (12)
Teaching or education	9 (12)
Accounting or banking or finance	7 (10)
Business or consulting or management	6 (8)
Engineering	5 (7)
Health care	5 (7)
Retail	5 (7)
Research	4 (5)
Other	15 (19)
Smoking status	
Nonsmoker	52 (69)
Past smoker	14 (19)
Current smoker	9 (12)
Fitness status	
Several times a week	34 (45)
Once or twice a week	25 (33)
Once or twice a month	8 (11)
Less than once a month	5 (7)
Never	3 (4)

Design

This study was a randomized laboratory-based experiment that examined HRV changes following stress exposure. Initially, participants underwent a baseline period, followed by a stressor period, which included 2 cognitive tasks and 2 emotion-eliciting film clips. They were then randomly assigned to 1 of the 3 conditions: BioBase app, mindfulness body scan, or control. HRV was assessed throughout the experiment.

Cognitive Stress

Cognitive performance-based tasks require high cognitive effort, and they increase physiological activation (HR) [46]. The use of nonverbal tasks also avoids the effect of cardiovascular confounds such as speaking, and its effects on respiration during a verbal stress task can mask the acute effect of the physiological stress response on HRV [47]. Therefore, in this study, we administered 2 tasks to induce cognitive stress. The first was the Continuous Performance Task, a neuropsychological performance-based task that assesses sustained attention. In this task, participants had to attend a series of letters presented on

XSL•FO
RenderX

the screen and respond by pressing the space bar whenever an "X" was presented. Letters were presented in 9 blocks of 20 letters, 5 of which were "X" in each block. All letters were displayed for 250 ms each. The interstimulus intervals were 1 second, and participants were given a *practice round* before commencing to make sure they understood when they needed to respond. The task lasted approximately 8 min and was adapted from a previous study [48]. The second task we used was the stop signal paradigm of Go/No-Go, which assesses inhibition. Participants were presented with a series of images. About 75% of these images were acting as Go *cues*, and participants needed to respond (press the space bar) as fast as possible. Occasionally the series of Go *cues* got interrupted by images that were acting as No-Go *cues*, and participants were instructed not to respond. Owing to the frequency of the Go *cues* (≥75%) being greater than the No-Go *cues*, a tendency to respond is created while participants should effortfully inhibit their responses to the No-Go *cues*. Participants responded to 2 blocks of 80 trials each, and this task was adapted from a previous study [49]. Participants were told to answer as quickly as possible in both tasks, which were programmed using E-PRIME 2.0(Psychology Software Tools).

Emotional Stress

Emotional stress was induced by asking participants to watch 2 emotion-eliciting film clips, a common technique to induce emotions in the laboratory while maintaining ecological validity [50]. The film clips were derived from the FilmStim database [51], which includes videos for emotion elicitation purposes. Participants watched 2 anger-eliciting film clips because anger has been found to induce the greatest physiological activation in HR among negative emotions [52]. The first anger-eliciting film clip was a 93-second excerpt from the film "American History X" which was found to produce a high arousal score (5.84/7) and the maximum negative affect score among databases (2.73/5). The second was a 43-second excerpt from the film "The Piano," producing a high arousal score (5.67/7) and very high negative affect score (2.49/5) [51]. Sound from both clips was removed to ensure that participants' attention was focused on the images [53].

BioBase App

The BioBase app is a commercially available app that provides guided breathing that can be modified according to participants' HR. In this study, participants in the BioBase breathing condition performed a 5-min guided breathing exercise based on the Papworth-Benson breathing method, which focuses on diaphragmatic breathing to use the full lung capacity and aims to decrease breathing rate, which will then trigger vagus nerve stimulation, leading to a desired relaxation response [42]. The relaxation phase occurs when HR and breathing synchronize or become resonant. Participants in this condition followed visual and audio instructions on a smartphone without being given feedback on their breathing. Instructions guided participants to breathe at 6 breaths per minute for 5 min. This is a common breathing rate that results in great HRV increases, which can occur within the first few minutes of practice [54].

Mindfulness Body Scan

Participants in the mindfulness condition were asked to perform a body scan for 5 min by following audio instructions [55]. This required participants to focus on the audio guiding their attention on their bodily and breathing sensations and encouraging nonjudgmental acceptance of thoughts and feelings experienced at that moment.

Self-Report Measures

Mindfulness

Participants completed the Five Factor Mindfulness Questionnaire (FFMQ) [56] before the experiment. The FFMQ is a 24-item self-report measure that assesses 5 distinct but related facets of mindfulness and consists of 5 different subscales: nonreactivity to inner experience, observing or noticing, acting with awareness, describing, and nonjudging of experience. Participants responded on a 5-point Likert scale ranging from 1 (never or very rarely true) to 5 (very often or always true).

Fatigue

Fatigue makes it difficult to maintain attention and has been found to impair performance in attention tasks [57]. To ensure that participants would be able to focus on the cognitive stressor task, we further controlled for fatigue by using the Samn-Perelli Fatigue Checklist [58] before testing started. This single item assesses state fatigue by asking participants to rate their fatigue levels on a 7-point Likert scale (1=fully alert to 7=completely exhausted, unable to function effectively, ready to drop).

Sleepiness

We also controlled for sleepiness levels as sleep deprivation is known to degrade inhibitory control [59]. The Stanford Sleepiness Scale [60] was used to assess subjective perceptions of daytime sleepiness/alertness by asking participants how sleepy/alert they felt before the experiment on a 7-point Guttman scale item (1=feeling active and vital, alert, wide awake to 7=lost struggle to remain awake).

Mood

The Visual Analog Scale (VAS) for Moods [61] was used as a brief measure to repeatedly assess participants' mood changes after the completion of each period. Participants were asked to indicate how tense they felt after the baseline period, the cognitive stressor, the emotional stressor, and the recovery period. Participants were asked, "At this particular moment, how tense do you feel?" and they answered by moving a marker up and down a 100-mm line representing a continuum from 0 (extremely tense) to 100 (not tense at all). Higher scores indicate higher levels of tension.

Heart Rate Variability

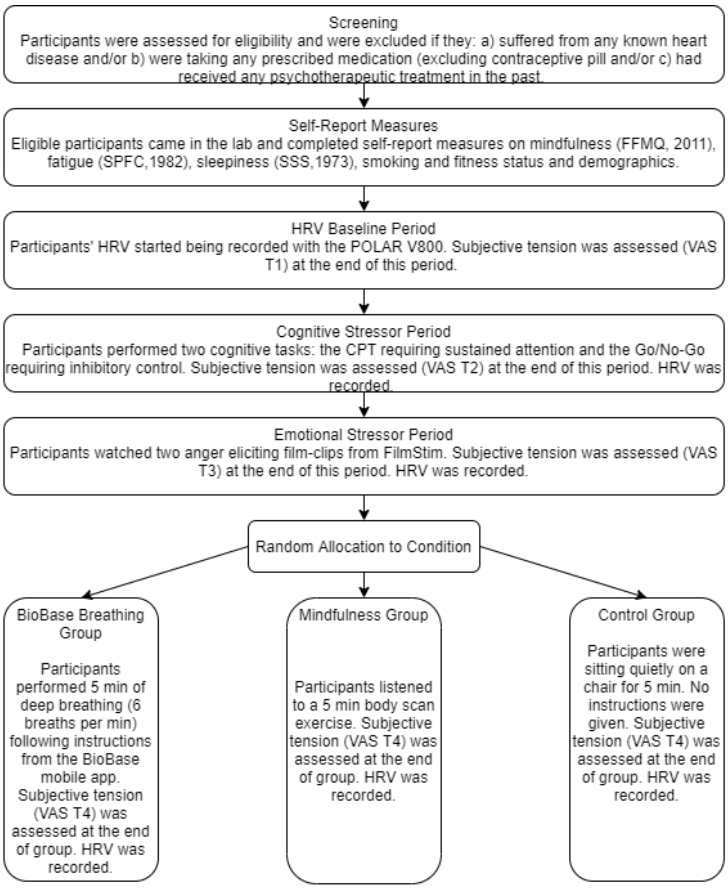
A Polar V800 HR monitor with a Polar H7 HR Sensor (Polar Electro OY) was used to record RR intervals at a sampling frequency of 1000 Hz [62]. The HR sensor was attached to a chest band, which sends a signal via Bluetooth to a smartwatch. RR intervals were continuously recorded throughout the experiment.

XSL•FO
RenderX

Procedure

The study was conducted in a laboratory at the University of Surrey. Before attending the laboratory session, participants gave their informed consent and completed a short health screening survey. Participants who reported having any heart condition or taking any prescribed medication (not including contraceptive pill) were excluded from participation as some pathological heart-related conditions such as arrhythmia and various types of medication can confound the interpretation of HRV values [63]. Those who reported having received psychotherapeutic treatment in the past were also excluded as we wanted to ensure that the conditions were novel to participants. In the laboratory, participants completed self-report measures on their demographics and control variables reported in the *Self-Report Measures* section. They were then given instructions on how to place the HR chest band. The experimenter left the room to allow privacy for participants to place the chest band. Once the chest band was safely attached, and HR readings were successfully synchronized with the smartwatch, participants were asked to quietly remain seated on a chair for 5 min; this was the baseline period. Once the 5 min had passed, they were asked to indicate how tense they felt by completing the VAS. Following this, they were asked to complete 2 cognitive tasks on the computer. Participants were given clear instructions on how to complete the tasks and were told that they measured attention speed and they had to respond as quickly as possible. To ensure that they understood the instructions, they completed a short practice block for which they were given feedback on their performance before each task. Once participants completed the tasks, they were asked to indicate how tense they felt by completing the VAS. They were then told that they had to watch 2 short film clips on the computer. They were not aware of the film clips' emotional valence. Once the film clips were viewed, participants were again asked to indicate how tense they felt on the VAS. Participants were then randomly allocated to 3 different recovery conditions: BioBase breathing, mindfulness body scan, and control. Block randomization was used to allocate participants into 3 groups, and web-based software was used to generate the group allocation. Participants in the control condition were asked to remain seated silently on the chair. They were not given any specific instructions regarding their breathing rates. At the end of the 5-min recovery phase, participants across all 3 conditions were asked to indicate how tense they felt on the VAS. Finally, they were told that the experiment had ended and the HR recording was turned off. Participants were asked to remove the chest band. The experimenter left the room and returned once participants were ready, thanked them for their participation, and gave them the debrief statement. Figure 1 shows the flowchart of the procedure.

Figure 1. Procedure flow chart. CPT: Continuous Performance Task; FFMQ: Five Factor Mindfulness Questionnaire; HRV: Heart Rate Variability; SPFC: Samn-Perelli Fatigue Checklist; SSS: Stanford Sleepiness Scale; VAS: Visual Analogue Scale.

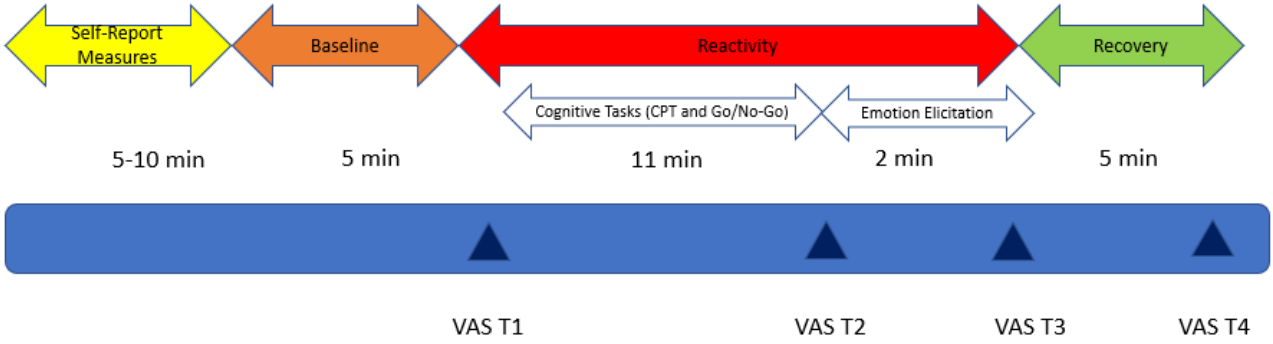


XSL•FO
RenderX

Data Analysis

The HRV analysis was conducted using ARTiiFACT software [64]. General guidelines [65] considered 5-min excerpts as an accepted short-term recording for group comparison and a minimum of 250 seconds is suggested for organizational research [66]. For this study, RR excerpts were grouped into 4 epochs (baseline, cognitive stressors, emotional stressors, and recovery). The baseline period was analyzed in 5-min segments; the cognitive stressor, in 11 min; the emotional stressor, in 2 min segments; and the recovery period, in 5-min segments across all participants. For the data analysis, values from the cognitive and emotional stressor periods were averaged and therefore analyzed as one total stressor period. Following previous research [42,67], we also accounted for individual differences in stress responsivity that could potentially influence the findings by creating a stress reactivity change score (derived from calculating the HRV difference between the total stressor period and the baseline period). To reduce HRV artifacts (owing to movement, sneezing, or loss of signal), recordings were taken while participants were sitting with knees at a 90^{o} angle, both feet flat on the floor without speaking or making any movements [68]. Artifacts were identified using the Berntson detection algorithm. Flagged RR intervals were visually inspected, and once confirmed as artifacts, they were corrected by applying cubic and/or linear spline interpolation, where necessary [69]. For this study, we used 3 HRV parameters. The first parameter was the root mean square of successive RR interval differences (RMSSD), which is a time-domain measure considered to reflect vagally mediated HRV. The second parameter was percentage of successive RR intervals that differ by more than 50 ms (pNN50), which reflects the percentage of successive RR intervals that differ >50 ms. The third parameter was the high frequency (HF), a frequency-domain parameter that represents breathing-related changes in HR, reflects parasympathetic contribution to HRV, and is regarded as an RSA measure [70]. Before analyses, the data were screened for outliers and normality. All HRV variables were positively skewed; therefore, data were log-transformed (lg10) because this transformation type can correct for positive skewness and kurtosis, unequal variances, and lack of linearity [71]. On the basis of their z-scores, 3 outliers were excluded as they were above 3 SDs. Statistical analyses were performed using SPSS 25. Figure 2 shows the experiment timeline.

Figure 2. Experiment timeline. CPT: Continuous Performance Task; VAS: Visual Analogue Scale.



Results

Participant Characteristics

One-way analysis of variances (ANOVAs) were conducted to detect any between-group differences (BioBase, mindfulness, and control) regarding baseline HRV. There were no significant differences between groups regarding baseline HRV as reflected on root mean square of successive RR interval differences (RMSSD; $F_{2,72}$=0.42; P=.65; η^2=0.01), pNN50 ($F_{2,69}$=1.42; P=.24; η^2=0.04), *and HF* ($F_{2,72}$=2.12; P=.33; η^2=0.05). We also tested for between-group differences in demographic variables such as age, gender, work hours, and external factors such as smoking, fitness status, and trait mindfulness to determine whether they had a systematic effect on HRV. Participant characteristics are presented in Multimedia Appendix 1. One-way ANOVAs on levels of fatigue ($F_{2,72}$=1.40; P=.25; η^2=0.03) and sleepiness/alertness ($F_{2,72}$=1.27; P=.28; η^2=0.03) revealed no significant difference between participants.

A paired sample *t* test was conducted to demonstrate that the experimental manipulation was successful (cognitive tasks and anger-eliciting film clips). Subjective tension after the cognitive task and the emotional task were averaged and then compared with subjective tension at baseline. RMSSD values after the cognitive task and emotional task were also averaged. We did this to capture HRV during the total reactivity period (Figure 2). Results showed that subjective levels of tension significantly increased after the cognitive and emotional stressors (mean 4.21, SD 2.18) compared with baseline (mean 1.28, SD 1.66), t_{74}= −10.84, P<.001, d=0.94. Similarly, RMSSD during the stressor period was significantly lower (mean 1.60, SD 0.23) than the baseline RMSSD (mean 1.68, SD 0.27), t_{74}=3.39, P=.001, d=0.48, indicating that the experimental manipulation was successful in physiologically arousing participants.

Poststress Physiological Recovery

To examine the effects of the intervention on cardiac recovery in the poststress period, a between-subjects analysis of covariance for the group (BioBase, mindfulness body scan, and control) was conducted using the HRV at recovery as the outcome and the reactivity change score as a covariate (to statistically remove its influence on recovery). Age was also added as a covariate as responsiveness of autonomic activity

XSL•FO
RenderX

reduces with older age [72]. There was a significant group effect for RMSSD, pNN50, and HF with a medium-sized effect, regardless of participants' individual differences in stress response. In all cases, HRV was higher in the BioBase group than in the mindfulness body scan and the control group. The results are shown in Table 2.

Table 2. Group means, standard deviations, between-group F ratio and effect sizes in the BioBase breathing, mindfulness body scan, and control group during poststress recovery.

Parameters	Groups			Values		
	BioBase breathing, mean (SE)	Mindfulness body scan, mean (SE)	Control, mean (SE)	*F* test (*df*)	*P* value	ηp^2
RMSSD[a,b]	1.82 (0.05)	1.68 (0.05)	1.55 (0.05)	6.801 (2,72)	.002	0.16
pNN50[a,c]	1.37 (0.07)	1.29 (0.07)	1.03 (0.08)	5.261 (2,69)	.008	0.13
HF[a,d]	2.97 (0.10)	2.77 (0.11)	2.50 (0.11)	4.695 (2,72)	.01	0.12

[a]Log transformed.

[b]RMSSD: root mean square of successive interbeat interval differences.

[c]pNN50: percentage of successive interbeat intervals that differ by more than 50 ms.

[d]HF: high frequency.

Discussion

Principal Findings

It is common for employees to encounter stressors at work that can cause physiological activation and, if prolonged, can in the longer term have deleterious effects on cardiovascular health [41]. The aim of this study was to compare the effects of a 5-min guided breathing exercise delivered through the BioBase app with a mindfulness body scan exercise and a control group in facilitating employees' physiological recovery following induced stress exposure.

Our principal finding was that HRV was higher at recovery for participants who performed BioBase breathing after stress induction. Results showed that there was a significant difference between groups, with those performing biofeedback breathing on the BioBase app demonstrating higher HRV compared with mindfulness and the control group during the recovery period. This is of great significance because HRV is an important cardiovascular health index, and higher HRV has been associated with numerous positive health outcomes [73].

Comparison With Previous Work

Recent studies have also investigated the effectiveness of app-based biofeedback breathing interventions. For example, Plans et al [42] assessed the effects of 6-min guided breathing delivered through the BioBase app compared with 6 min of rumination and controlled conditions after stress exposure induced by completing a speech task. They found that individuals who performed guided breathing on the BioBase app had significantly higher HRV than those who ruminated and the control group poststress. Our study findings are in line with their findings, but more importantly, this study was designed to overcome some of their limitations. For example, their study did not account for confounding variables such as fitness and smoking status. Furthermore, they compared the effects of breathing with rumination, which mostly pertains to repetitive thinking of negative nature, whereas we compared breathing with mindfulness, an active intervention historically associated with positive health outcomes. In addition, the breathing rate used in this study was based on past literature suggesting 6 breaths per minute as the average rate to achieve RSA [54]. Our findings also support recent evidence showing that a brief 5-min app-based biofeedback training session was effective in enhancing physiological recovery after laboratory-induced stress, as assessed by levels of salivary alpha-amylase, another commonly used stress biomarker [74].

Previous studies have compared the effects of biofeedback breathing and mindfulness with other self-help methods. For example, a study compared the effects of a 5-week daily biofeedback breathing intervention with mindfulness and physical activity training [75] and reported no overall difference in their effectiveness. Biofeedback training was also found to be equally effective with mindfulness and physical exercise training in reducing worry symptoms and improving attention in young adults [76]. These results are in contrast with our findings, which revealed that biofeedback breathing resulted in greater increases in HRV compared with mindfulness. However, it should be noted that these previous studies [75,76] are limited by the fact that the interventions consisted of numerous sessions, and outcome measures were of a self-report nature while our study looked at physiological changes over a single session. Moreover, these previous studies examined the intervention effects on stress responses, whereas this study examined the effect of the intervention on recovery from physiological stress. Further research is needed to shed light on the differential effects that brief interventions and long-term interventions can have on physiological and self-report outcome measures.

Limitations

Despite our study building upon and extending findings from the previous studies, it comes with its own limitations that warrant discussion. First, HRV during recovery was recorded for only 5 min. Even though this is classified as the minimum accepted short-term recording [65], it could be of great interest to explore how long after the recovery period did the effects

XSL•FO
RenderX

last. Moreover, we induced anger among participants, which is a commonly experienced emotion in the workplace that can pose detrimental effects not only for the employee experiencing it but also on the organizational level [77]. Therefore, our results add value to the notion that BioBase was successful in helping employees physiologically recover from a negative emotion. However, it would be worth exploring the effectiveness of the BioBase breathing app in helping employees overcome negative emotions other than anger. Another limitation is that we did not establish the resonant frequency for each individual but we used the commonly accepted average of 6 breaths per minute. Lehrer et al [54] suggested that practicing breathing at the unique resonant frequency for each individual is a requirement to produce maximal HRV increases. Our study included a single brief breathing session, so establishing each participant's resonant frequency would be time consuming. Greater effect sizes might have emerged if we applied the unique resonance frequency breathing rate for each participant.

Regardless of these limitations, the results derived from this study provide a stimulus for the delivery of future stress recovery interventions. It would be of interest to test the effects of repeated biofeedback breathing exercises on employees' health. For example, future studies could investigate whether the daily practice of short biofeedback sessions could potentially result in HRV increases not only during practice but also in trait physiological changes in HRV. It would also be of interest to test whether biofeedback breathing could increase HRV among employees who ruminate about work during their leisure time [78].

Conclusions

To our knowledge, this study is the first randomized study to investigate the effect of biofeedback and mindfulness on physiological recovery from momentary stress in employees. This study demonstrated that a short period of focused biofeedback breathing can aid physiological recovery in employees following a laboratory-induced stressor. Using a smartphone app to aid physiological relaxation may also help individuals to recover from stressful situations in real life, both inside and outside of work. By taking into consideration the rapid advances in mobile technology, this study showcases the great utility of short smartphone-delivered relaxation methods. With their great practicality and high accessibility, mobile interventions can be a potential solution to the ever-increasing demand for ease of use and parsimonious stress-management interventions.

Conflicts of Interest

DP, SP, and DM are employees of a commercial company, BioBeats Group Ltd, London, United Kingdom. The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Multimedia Appendix 1
Participants' characteristics per group.
[DOCX File , 16 KB - mhealth_v8i10e19412_app1.docx]

References

1. Steiber N, Pichler F. Trends in work stress and exhaustion in advanced economies. Soc Indic Res 2014 Apr 26;121(1):215-239. [doi: 10.1007/s11205-014-0633-7]
2. Kivimäki M, Kawachi I. Work stress as a risk factor for cardiovascular disease. Curr Cardiol Rep 2015 Sep;17(9):630 [FREE Full text] [doi: 10.1007/s11886-015-0630-8] [Medline: 26238744]
3. Chandola T, Britton A, Brunner E, Hemingway H, Malik M, Kumari M, et al. Work stress and coronary heart disease: what are the mechanisms? Eur Heart J 2008 Mar;29(5):640-648. [doi: 10.1093/eurheartj/ehm584] [Medline: 18216031]
4. Johnson NB, Hayes LD, Brown K, Hoo EC, Ethier KA, Centers for Disease Control and Prevention (CDC). CDC national health report: leading causes of morbidity and mortality and associated behavioral risk and protective factors--United States, 2005-2013. MMWR Suppl 2014 Oct 31;63(4):3-27. [Medline: 25356673]
5. Brook RD, Julius S. Autonomic imbalance, hypertension, and cardiovascular risk. Am J Hypertens 2000 Jun;13(6 Pt 2):112S. [doi: 10.1016/s0895-7061(00)00228-4] [Medline: 10921530]
6. Thayer J, Sternberg E. Beyond heart rate variability: vagal regulation of allostatic systems. Ann N Y Acad Sci 2006 Nov;1088:361-372. [doi: 10.1196/annals.1366.014] [Medline: 17192580]
7. McEwen B. Stress, adaptation, and disease. Allostasis and allostatic load. Ann N Y Acad Sci 1998 May 1;840:33-44. [doi: 10.1111/j.1749-6632.1998.tb09546.x] [Medline: 9629234]
8. Brosschot JF, Pieper S, Thayer JF. Expanding stress theory: prolonged activation and perseverative cognition. Psychoneuroendocrinology 2005 Nov;30(10):1043-1049. [doi: 10.1016/j.psyneuen.2005.04.008] [Medline: 15939546]
9. Appelhans BM, Luecken LJ. Heart rate variability as an index of regulated emotional responding. Rev Gen Psychol 2006 Sep;10(3):229-240. [doi: 10.1037/1089-2680.10.3.229]
10. Segerstrom SC, Nes LS. Heart rate variability reflects self-regulatory strength, effort, and fatigue. Psychol Sci 2007 Mar;18(3):275-281. [doi: 10.1111/j.1467-9280.2007.01888.x] [Medline: 17444926]

XSL•FO
RenderX

11. Koval P, Ogrinz B, Kuppens P, van den Bergh O, Tuerlinckx F, Sütterlin S. Affective instability in daily life is predicted by resting heart rate variability. PLoS One 2013;8(11):e81536 [FREE Full text] [doi: 10.1371/journal.pone.0081536] [Medline: 24312315]
12. Thayer J, Siegle G. Neurovisceral integration in cardiac and emotional regulation. IEEE Eng Med Biol Mag 2002;21(4):24-29. [doi: 10.1109/memb.2002.1032635] [Medline: 12222113]
13. Weber F, Exner C. Metacognitive beliefs and rumination: a longitudinal study. Cogn Ther Res 2013 Jun 11;37(6):1257-1261. [doi: 10.1007/s10608-013-9555-y]
14. Tsuji H, Larson MG, Venditti FJ, Manders ES, Evans JC, Feldman CL, et al. Impact of reduced heart rate variability on risk for cardiac events. The Framingham heart study. Circulation 1996 Dec 1;94(11):2850-2855. [doi: 10.1161/01.cir.94.11.2850] [Medline: 8941112]
15. Masi CM, Hawkley LC, Rickett EM, Cacioppo JT. Respiratory sinus arrhythmia and diseases of aging: obesity, diabetes mellitus, and hypertension. Biol Psychol 2007 Feb;74(2):212-223 [FREE Full text] [doi: 10.1016/j.biopsycho.2006.07.006] [Medline: 17034928]
16. Thayer JF, Yamamoto SS, Brosschot JF. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. Int J Cardiol 2010 May 28;141(2):122-131. [doi: 10.1016/j.ijcard.2009.09.543] [Medline: 19910061]
17. Gevirtz R. The promise of heart rate variability biofeedback: evidence-based applications. Biofeedback 2013 Sep;41(3):110-120. [doi: 10.5298/1081-5937-41.3.01]
18. Schwartz M, Collura T, Kamiya J, Schwartz N. Biofeedback: A Practitioner's Guide. New York, USA: Guilford Press; 2014.
19. Lehrer PM, Gevirtz R. Heart rate variability biofeedback: how and why does it work? Front Psychol 2014;5:756 [FREE Full text] [doi: 10.3389/fpsyg.2014.00756] [Medline: 25101026]
20. Yasuma F, Hayano J. Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? Chest 2004 Feb;125(2):683-690. [doi: 10.1378/chest.125.2.683] [Medline: 14769752]
21. Lehrer P, Gevirtz R. Heart rate variability biofeedback: how and why does it work? Front Psychol 2014;5:756 [FREE Full text] [doi: 10.3389/fpsyg.2014.00756] [Medline: 25101026]
22. Vaschillo EG, Vaschillo B, Lehrer PM. Characteristics of resonance in heart rate variability stimulated by biofeedback. Appl Psychophysiol Biofeedback 2006 Jun;31(2):129-142. [doi: 10.1007/s10484-006-9009-3] [Medline: 16838124]
23. Hassett AL, Radvanski DC, Vaschillo EG, Vaschillo B, Sigal LH, Karavidas MK, et al. A pilot study of the efficacy of heart rate variability (HRV) biofeedback in patients with fibromyalgia. Appl Psychophysiol Biofeedback 2007 Mar;32(1):1-10. [doi: 10.1007/s10484-006-9028-0] [Medline: 17219062]
24. Karavidas MK, Lehrer PM, Vaschillo E, Vaschillo B, Marin H, Buyske S, et al. Preliminary results of an open label study of heart rate variability biofeedback for the treatment of major depression. Appl Psychophysiol Biofeedback 2007 Mar;32(1):19-30. [doi: 10.1007/s10484-006-9029-z] [Medline: 17333315]
25. del Pozo JM, Gevirtz RN, Scher B, Guarneri E. Biofeedback treatment increases heart rate variability in patients with known coronary artery disease. Am Heart J 2004 Mar;147(3):E11. [doi: 10.1016/j.ahj.2003.08.013] [Medline: 14999213]
26. Goessl VC, Curtiss JE, Hofmann SG. The effect of heart rate variability biofeedback training on stress and anxiety: a meta-analysis. Psychol Med 2017 Nov;47(15):2578-2586. [doi: 10.1017/S0033291717001003] [Medline: 28478782]
27. Lemaire J, Wallace J, Lewin A, de GJ, Schaefer J. The effect of a biofeedback-based stress management tool on physician stress: a randomized controlled clinical trial. Open Medicine , e154.10.1023/a 2011;5(4):a.
28. Cutshall SM, Wentworth LJ, Wahner-Roedler DL, Vincent A, Schmidt JE, Loehrer LL, et al. Evaluation of a biofeedback-assisted meditation program as a stress management tool for hospital nurses: a pilot study. Explore (NY) 2011;7(2):110-112. [doi: 10.1016/j.explore.2010.12.004] [Medline: 21397872]
29. Wells R, Outhred T, Heathers JA, Quintana DS, Kemp AH. Matter over mind: a randomised-controlled trial of single-session biofeedback training on performance anxiety and heart rate variability in musicians. PLoS One 2012 Oct 4;7(10):e46597 [FREE Full text] [doi: 10.1371/journal.pone.0046597] [Medline: 23056361]
30. Wang S, Li S, Xu X, Lin G, Shao L, Zhao Y, et al. Effect of slow abdominal breathing combined with biofeedback on blood pressure and heart rate variability in prehypertension. J Altern Complement Med 2010 Oct;16(10):1039-1045. [doi: 10.1089/acm.2009.0577] [Medline: 20954960]
31. Steffen PR, Austin T, DeBarros A, Brown T. The impact of resonance frequency breathing on measures of heart rate variability, blood pressure, and mood. Front Public Health 2017;5:222 [FREE Full text] [doi: 10.3389/fpubh.2017.00222] [Medline: 28890890]
32. Kabat-Zinn J. Mindfulness‐based interventions in context: past, present, and future. Clin Psychol 2006 May 11;10(2):144-156. [doi: 10.1093/clipsy.bpg016]
33. Burg JM, Wolf OT. Mindfulness as self-regulated attention. Swiss J Psychol 2012 Jan;71(3):135-139. [doi: 10.1024/1421-0185/a000080]
34. Ditto B, Eclache M, Goldman N. Short-term autonomic and cardiovascular effects of mindfulness body scan meditation. Ann Behav Med 2006 Dec;32(3):227-234. [doi: 10.1207/s15324796abm3203_9] [Medline: 17107296]
35. Shearer A, Hunt M, Chowdhury M, Nicol L. Effects of a brief mindfulness meditation intervention on student stress and heart rate variability. Int J Stress Manag 2016 May;23(2):232-254. [doi: 10.1037/a0039814]

XSL•FO
RenderX

36. Goodman MJ, Schorling JB. A mindfulness course decreases burnout and improves well-being among healthcare providers. Int J Psychiatry Med 2012;43(2):119-128. [doi: 10.2190/PM.43.2.b] [Medline: 22849035]
37. Preziosa A, Grassi A, Gaggioli A, Riva G. Therapeutic applications of the mobile phone. Brit J Guid Couns 2009 Aug;37(3):313-325. [doi: 10.1080/03069880902957031]
38. Parnandi A, Ahmed B, Shipp E, Gutierrez-Osuna R. Chill-Out: Relaxation Training through Respiratory Biofeedback in a Mobile Casual Game. In: International Conference on Mobile Computing, Applications, and Services. 2013 Presented at: MobiCase'13; November 7-8, 2014; Paris, France. [doi: 10.1007/978-3-319-05452-0_18]
39. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app 'calm' to reduce stress among college students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
40. Collins E, Cox A, Wilcock C, Sethu-Jones G. Digital games and mindfulness apps: comparison of effects on post work recovery. JMIR Ment Health 2019 Jul 18;6(7):e12853 [FREE Full text] [doi: 10.2196/12853] [Medline: 31322125]
41. Brosschot JF, Gerin W, Thayer JF. The perseverative cognition hypothesis: a review of worry, prolonged stress-related physiological activation, and health. J Psychosom Res 2006 Feb;60(2):113-124. [doi: 10.1016/j.jpsychores.2005.06.074] [Medline: 16439263]
42. Plans D, Morelli D, Sütterlin S, Ollis L, Derbyshire G, Cropley M. Use of a biofeedback breathing app to augment poststress physiological recovery: randomized pilot study. JMIR Form Res 2019 Jan 11;3(1):e12227 [FREE Full text] [doi: 10.2196/12227] [Medline: 30684437]
43. Dreeben SJ, Mamberg MH, Salmon P. The MBSR body scan in clinical practice. Mindfulness 2013 Apr 7;4(4):394-401. [doi: 10.1007/s12671-013-0212-z]
44. Hayano J, Sakakibara Y, Yamada M, Ohte N, Fujinami T, Yokoyama K, et al. Decreased magnitude of heart rate spectral components in coronary artery disease. Its relation to angiographic severity. Circulation 1990 Apr;81(4):1217-1224. [doi: 10.1161/01.cir.81.4.1217] [Medline: 2317904]
45. Aubert AE, Seps B, Beckers F. Heart rate variability in athletes. Sports Med 2003;33(12):889-919. [doi: 10.2165/00007256-200333120-00003] [Medline: 12974657]
46. Prinsloo GE, Derman WE, Lambert MI, Rauch HG. The effect of a single session of short duration biofeedback-induced deep breathing on measures of heart rate variability during laboratory-induced cognitive stress: a pilot study. Appl Psychophysiol Biofeedback 2013 Jun;38(2):81-90. [doi: 10.1007/s10484-013-9210-0] [Medline: 23435801]
47. Brugnera A, Zarbo C, Tarvainen MP, Marchettini P, Adorni R, Compare A. Heart rate variability during acute psychosocial stress: a randomized cross-over trial of verbal and non-verbal laboratory stressors. Int J Psychophysiol 2018 May;127:17-25. [doi: 10.1016/j.ijpsycho.2018.02.016] [Medline: 29501671]
48. Brinker JK, Campisi M, Gibbs L, Izzard R. Rumination, Mood and Cognitive Performance. Psychology 2013;04(03):224-231. [doi: 10.4236/psych.2013.43a034]
49. Rezvanfard M, Golesorkhi M, Ghassemian E, Safaei H, Eghbali AN, Alizadeh H, et al. Evaluation of inhibition response behavior using the go/no-go paradigm in normal individuals: effects of variations in the task design. Acta Neuropsychologica 2016 Oct 20;14(4):357-366. [doi: 10.5604/17307503.1227530]
50. Kreibig SD. Autonomic nervous system activity in emotion: a review. Biol Psychol 2010 Jul;84(3):394-421. [doi: 10.1016/j.biopsycho.2010.03.010] [Medline: 20371374]
51. Schaefer A, Nils F, Sanchez X, Philippot P. Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers. Cognit Emot 2010 Nov;24(7):1153-1172. [doi: 10.1080/02699930903274322]
52. Fernández C, Pascual JC, Soler J, Elices M, Portella MJ, Fernández-Abascal E. Physiological responses induced by emotion-eliciting films. Appl Psychophysiol Biofeedback 2012 Jun;37(2):73-79. [doi: 10.1007/s10484-012-9180-7] [Medline: 22311202]
53. Berna G, Ott L, Nandrino J. Effects of emotion regulation difficulties on the tonic and phasic cardiac autonomic response. PLoS One 2014;9(7):e102971 [FREE Full text] [doi: 10.1371/journal.pone.0102971] [Medline: 25054913]
54. Lehrer PM, Vaschillo E, Vaschillo B. Resonant frequency biofeedback training to increase cardiac variability: rationale and manual for training. Appl Psychophysiol Biofeedback 2000 Sep;25(3):177-191. [doi: 10.1023/a:1009554825745] [Medline: 10999236]
55. Leslie Rohanczy Band. URL: http://www.leslierohonczy.com/LRBand.htm [accessed 2020-09-29]
56. Bohlmeijer E, ten Klooster PM, Fledderus M, Veehof M, Baer R. Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. Assessment 2011 Sep;18(3):308-320. [doi: 10.1177/1073191111408231] [Medline: 21586480]
57. Faber LG, Maurits NM, Lorist MM. Mental fatigue affects visual selective attention. PLoS One 2012 Oct 31;7(10):e48073 [FREE Full text] [doi: 10.1371/journal.pone.0048073] [Medline: 23118927]
58. Samn S, Perelli L. Estimating aircrew fatigue: a technique with application to airlift operations. School Aerospace Med 1982.
59. Drummond SP, Paulus MP, Tapert SF. Effects of two nights sleep deprivation and two nights recovery sleep on response inhibition. J Sleep Res 2006 Sep;15(3):261-265 [FREE Full text] [doi: 10.1111/j.1365-2869.2006.00535.x] [Medline: 16911028]

XSL•FO
RenderX

60. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. Psychophysiology 1973 Jul;10(4):431-436. [doi: 10.1111/j.1469-8986.1973.tb00801.x] [Medline: 4719486]
61. Stern RA, Arruda JE, Hooper CR, Wolfner GD, Morey CE. Visual analogue mood scales to measure internal mood state in neurologically impaired patients: description and initial validity evidence. Aphasiology 1997 Jan;11(1):59-71. [doi: 10.1080/02687039708248455]
62. Caminal P, Sola F, Gomis P, Guasch E, Perera A, Soriano N, et al. Validity of the Polar V800 monitor for measuring heart rate variability in mountain running route conditions. Eur J Appl Physiol 2018 Mar;118(3):669-677. [doi: 10.1007/s00421-018-3808-0] [Medline: 29356949]
63. Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. Med Biol Eng Comput 2006 Dec;44(12):1031-1051. [doi: 10.1007/s11517-006-0119-0] [Medline: 17111118]
64. Kaufmann T, Sütterlin S, Schulz SM, Vögele C. ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. Behav Res Methods 2011 Dec;43(4):1161-1170. [doi: 10.3758/s13428-011-0107-7] [Medline: 21573720]
65. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use.. Circulation 1996 Mar 1;93(5):1043-1065. [Medline: 8598068]
66. Massaro S, Pecchia L. Heart rate variability (HRV) analysis: a methodology for organizational neuroscience. Organ Res Methods 2016 Dec 13;22(1):354-393. [doi: 10.1177/1094428116681072]
67. Key BL, Campbell TS, Bacon SL, Gerin W. The influence of trait and state rumination on cardiovascular recovery from a negative emotional stressor. J Behav Med 2008 Jun;31(3):237-248. [doi: 10.1007/s10865-008-9152-9] [Medline: 18350377]
68. Laborde S, Mosley E, Thayer JF. Heart rate variability and cardiac vagal tone in psychophysiological research - recommendations for experiment planning, data analysis, and data reporting. Front Psychol 2017;8:213 [FREE Full text] [doi: 10.3389/fpsyg.2017.00213] [Medline: 28265249]
69. Ribeiro GD, Neves VR, Deresz LF, Melo RD, dal Lago P, Karsten M. Can RR intervals editing and selection techniques interfere with the analysis of heart rate variability? Braz J Phys Ther 2018;22(5):383-390 [FREE Full text] [doi: 10.1016/j.bjpt.2018.03.008] [Medline: 29653903]
70. Thomas B, Claassen N, Becker P, Viljoen M. Validity of commonly used heart rate variability markers of autonomic nervous system function. Neuropsychobiology 2019;78(1):14-26. [doi: 10.1159/000495519] [Medline: 30721903]
71. Andy F. Discovering statistics using IBM SPSS statistics. And sex and drugs and rock'n'roll. Pflege 2014 Dec 1;27(6):430. [doi: 10.1024/1012-5302/a000397]
72. Reardon M, Malik M. Changes in heart rate variability with age. Pacing Clin Electrophysiol 1996 Nov;19(11 Pt 2):1863-1866. [doi: 10.1111/j.1540-8159.1996.tb03241.x] [Medline: 8945057]
73. Ernst G. Heart-rate variability-more than heart beats? Front Public Health 2017;5:240 [FREE Full text] [doi: 10.3389/fpubh.2017.00240] [Medline: 28955705]
74. Hunter JF, Olah MS, Williams AL, Parks AC, Pressman SD. Effect of brief biofeedback via a smartphone app on stress recovery: randomized experimental study. JMIR Serious Games 2019 Nov 26;7(4):e15974 [FREE Full text] [doi: 10.2196/15974] [Medline: 31769761]
75. van der Zwan JE, de Vente W, Huizink AC, Bögels SM, de Bruin EI. Physical activity, mindfulness meditation, or heart rate variability biofeedback for stress reduction: a randomized controlled trial. Appl Psychophysiol Biofeedback 2015 Dec;40(4):257-268 [FREE Full text] [doi: 10.1007/s10484-015-9293-x] [Medline: 26111942]
76. de Bruin EI, van der Zwan JE, Bögels SM. A RCT comparing daily mindfulness meditations, biofeedback exercises, and daily physical exercise on attention control, executive functioning, mindful awareness, self-compassion, and worrying in stressed young adults. Mindfulness (N Y) 2016;7(5):1182-1192 [FREE Full text] [doi: 10.1007/s12671-016-0561-5] [Medline: 27642375]
77. Booth J, Mann S. The experience of workplace anger. Leadership Org Development J 2005 Jun;26(4):250-262. [doi: 10.1108/01437730510600634]
78. Cropley M, Plans D, Morelli D, Sütterlin S, Inceoglu I, Thomas G, et al. The association between work-related rumination and heart rate variability: a field study. Front Hum Neurosci 2017;11:27 [FREE Full text] [doi: 10.3389/fnhum.2017.00027] [Medline: 28197087]

Abbreviations

ANOVA: analysis of variance
ANS: autonomic nervous system
BP: blood pressure
FFMQ: Five Factor Mindfulness Questionnaire
HF: high frequency
HR: heart rate
HRV: heart rate variability
pNN50: percentage of successive RR intervals that differ by more than 50 ms

XSL•FO
RenderX

RMSSD: root mean square of successive RR interval differences
RR: interbeat
RSA: respiratory sinus arrhythmia
VAS: Visual Analog Scale

Edited by G Eysenbach; submitted 16.04.20; peer-reviewed by DS Almeida, J Weerdmeester, L Becker; comments to author 12.06.20; revised version received 20.07.20; accepted 26.07.20; published 15.10.20.

Please cite as:
Chelidoni O, Plans D, Ponzo S, Morelli D, Cropley M
Exploring the Effects of a Brief Biofeedback Breathing Session Delivered Through the BioBase App in Facilitating Employee Stress Recovery: Randomized Experimental Study
JMIR Mhealth Uhealth 2020;8(10):e19412
URL: http://mhealth.jmir.org/2020/10/e19412/
doi:10.2196/19412
PMID:33055072

©Olga Chelidoni, David Plans, Sonia Ponzo, Davide Morelli, Mark Cropley. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 15.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Light-Induced Fluorescence-Based Device and Hybrid Mobile App for Oral Hygiene Management at Home: Development and Usability Study

Jun-Min Kim[1,2], PhD; Woo Ram Lee[1,3], PhD; Jun-Ho Kim[4], PhD; Jong-Mo Seo[1], MD, PhD; Changkyun Im[5,6], PhD

[1]Department of Electrical and Computer Engineering, Seoul National University, Seoul, Republic of Korea

[2]Department of Electronics and Information Engineering, Hansung University, Seoul, Republic of Korea

[3]Department of Electronic Communication Engineering, Gyeonggi University of Science Technology, Siheung, Republic of Korea

[4]Department of Electrical Energy Engineering, Keimyung University, Daegu, Republic of Korea

[5]Dental Research Institute, Seoul National University, Seoul, Republic of Korea

[6]Dental Life Science Research Institute, Seoul National University Dental Hospital, Seoul, Republic of Korea

Corresponding Author:
Changkyun Im, PhD
Dental Life Science Research Institute
Seoul National University Dental Hospital
601 Seoul National University Dental Hospital
1 Gwanak-ro, Gwanak-Gu
Seoul,
Republic of Korea
Phone: 82 2 6747 6587
Email: changkyun@snu.ac.kr

Abstract

Background: Dental diseases can be prevented through the management of dental plaques. Dental plaque can be identified using the light-induced fluorescence (LIF) technique that emits light at 405 nm. The LIF technique is more convenient than the commercial technique using a disclosing agent, but the result may vary for each individual as it still requires visual identification.

Objective: The objective of this study is to introduce and validate a deep learning–based oral hygiene monitoring system that makes it easy to identify dental plaques at home.

Methods: We developed a LIF-based system consisting of a device that can visually identify dental plaques and a mobile app that displays the location and area of dental plaques on oral images. The mobile app is programmed to automatically determine the location and distribution of dental plaques using a deep learning–based algorithm and present the results to the user as time series data. The mobile app is also built with convergence of naive and web applications so that the algorithm is executed on a cloud server to efficiently distribute computing resources.

Results: The location and distribution of users' dental plaques could be identified via the hand-held LIF device or mobile app. The color correction filter in the device was developed using a color mixing technique. The mobile app was built as a hybrid app combining the functionalities of a native application and a web application. Through the scrollable WebView on the mobile app, changes in the time series of dental plaque could be confirmed. The algorithm for dental plaque detection was implemented to run on Amazon Web Services for object detection by single shot multibox detector and instance segmentation by Mask region-based convolutional neural network.

Conclusions: This paper shows that the system can be used as a home oral care product for timely identification and management of dental plaques. In the future, it is expected that these products will significantly reduce the social costs associated with dental diseases.

(JMIR Mhealth Uhealth 2020;8(10):e17881) doi:10.2196/17881

KEYWORDS

dental plaque; oral hygiene; red fluorescence; mobile health; deep learning; object detection; instance segmentation

XSL•FO
RenderX

Introduction

Dental plaque is a sticky biofilm associated with oral diseases such as tooth decay and periodontal disease. Management of dental plaque is one of the effective ways to prevent dental diseases, but its transparent and colorless properties make it difficult to visually identify and manage. Therefore, it is important to improve motivation for dental plaque management by making it easy to identify whether the dental plaque adheres to the tooth surface. The most commonly recommended method is to use a disclosing agent that can identify the dental plaque by staining [1,2]. However, the assessment of dental plaque accumulation is subjective and error-prone because it is generally performed through self-monitoring. Therefore, accurate evaluation requires clinical examination by a clinician, which increases cost and time [3,4]. In addition, the mouth and tongue may be stained all day or temporarily by disclosing agents, and some products contain dyes that may cause allergic reactions [5].

It has been demonstrated that dental plaque can also be identified by the light-induced fluorescence (LIF) technique [6,7]. This technique is based on the red fluorescence property of porphyrins, metabolites of heterogeneous bacteria within dental plaque, when irradiated with narrow blue-violet light (centered at 405 nm wavelength) [8,9].

Since it was first reported in the 1920s that dental plaque emits red fluorescence by ultraviolet ray irradiation [10,11], various studies using LIF characteristics have been published, such as fluorescence changes according to oral bacterial species, fluorescence imaging systems, and clinical diagnosis. Studies of fluorescence changes by oral bacterial species reported that red fluorescence was detected in *Prevotella intermedia*, *P. melaninogenica*, *Actinomyces naeslundi*, *A. israelii*, and *Bifidobacterium dentium*, green fluorescence was observed in *Streptococcus oralis*, *S. salivarius*, *S. mutans*, *S. mitis*, *S. sobrinus*, *Fusobacterium nucleatum*, and *Propionibacterium acnes*, and orange fluorescence was found in *Latobacillus fermentans*, *L. rhamnosus*, and *L. casei*, and *Candida albicans* [8,12-14]. Commercial products reported in the scientific literature as LIF imaging systems for dental plaque detection include ACTEON SOPROLIFE (Henry Schein, Inc) [15] and QLF-D Biluminator (Inspektor Research Systems BV) [16]. These products have been used to evaluate plaque levels by comparing clinical and red fluorescent plaques [16-18], and some studies have shown that tooth defects such as caries, calculus, hypomineralization, and discoloration can be observed with red fluorescence [15,19,20]. However, these products are not only classified as medical devices, but also have problems in size and price, making them unsuitable for ordinary users to motivate plaque management at home.

Therefore, we judged that an inexpensive, compact, and intuitively usable system would be suitable for motivating plaque management to ordinary users and developed an oral hygiene monitoring system consisting of a LIF device and a smartphone-based mobile app. The LIF device induces the fluorescence of plaque and the app serves to visualize the plaque area, but in order to greatly motivate, an image processing method that emphasizes the plaque area is required. The traditional image processing method is sufficient to emphasize the plaque area in a limited environment, but an ordinary user needs an image processing method suitable for taking oral images with various cameras such as an Android phone and an iPhone in various environments such as a room, a bathroom, and a living room. Recently, deep learning algorithms have been applied to detection, segmentation, classification, and prediction in various medical fields, including plaque classification, showing surprising results [21,22]. Therefore, it was determined that the deep learning algorithm would be the best way to solve the above-mentioned disadvantages.

In this paper, we introduce a deep learning–based oral hygiene monitoring system that makes it easy to identify dental plaques in our home. The system consists of a device that can visually identify dental plaques and a mobile app that displays the location and area of dental plaques on the oral image. The mobile app was developed based on two deep learning models to sequentially detect tooth areas and highlight plaque areas.

Methods

Overview

In this study, we developed a hand-held LIF device that allows the user to easily monitor dental plaque with the naked eye in a mirror and a hybrid mobile app that provides oral hygiene information. Figure 1 shows the LIF device and the hybrid mobile app.

XSL•FO
RenderX

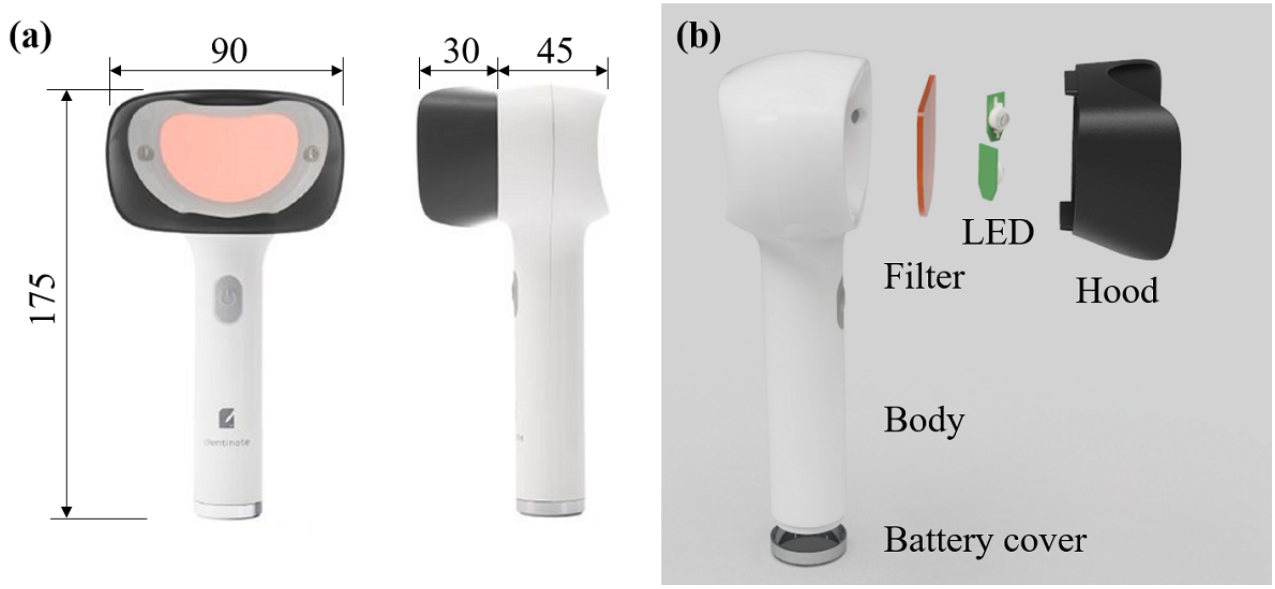
Figure 1. Oral hygiene monitoring system: (a) light-induced fluorescence device and hybrid mobile app; (b) method for monitoring oral hygiene with the naked eye through a mirror; (c) method for monitoring oral hygiene using smartphone; (d) hybrid mobile app that provides time series oral hygiene information.

Development of Light-Induced Fluorescence Hardware Interface

As shown in Figure 2, the LIF device consists of 5 parts: body, two light-emitting diodes (LEDs), filter, hood, and battery cover. The body has a cylindrical handle with 28 mm diameter for easy hand grip, a button manufactured through silicone insert injection molding for better assembly, and a wide viewer suitable for viewing the oral hygiene. The LED (GTPDTV64101, Shenzhen Getian Opto-Electronics Co Ltd) emits 1 W narrow-band spectrum in the wavelength range of 400 to 410 nm with the view angle of 120 degrees. The filter is manufactured through a color-mixing technique in injection molding so that it has a color correction function for identifying clean teeth as white and plaque as red in color. The hood is matte black to block ambient light and minimize light reflection and is interchangeable for each user. The battery cover is designed for high friction through serration technique and it is not easily separated from the body. The device is powered by three AAA batteries and can last for approximately two and a half hours.

XSL•FO
RenderX

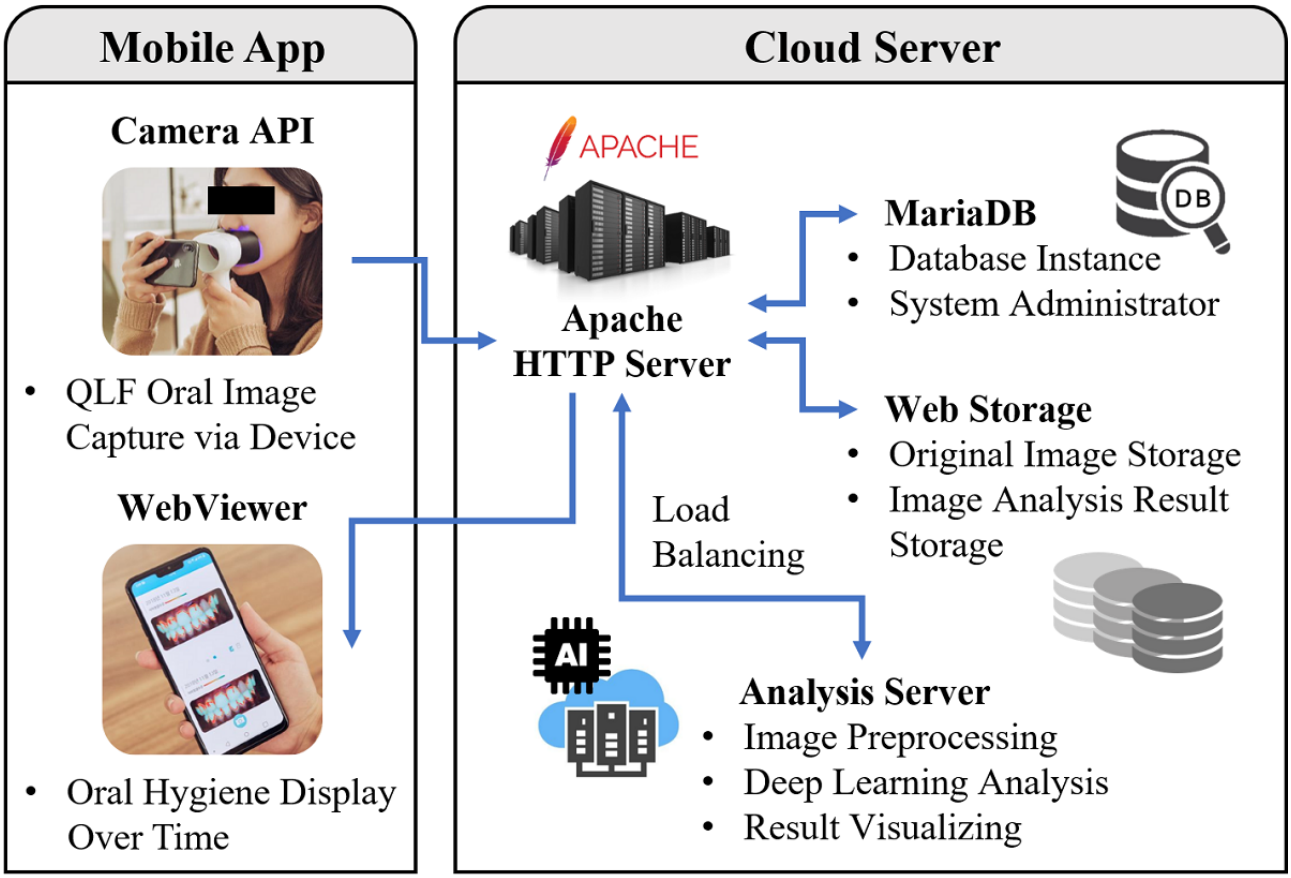
Figure 2. Light-induced fluorescence device for monitoring oral health: (a) dimensions (mm) and (b) components.



Development of Hybrid Mobile App

A hybrid mobile app is one that combines the functionalities of both a native app, which allows access to the smartphone camera, and web application, which can easily be updated without installation by modifying only the code on the server. In this study, we built a hybrid mobile app using JAVA for Android and Swift for iOS and, as shown in Figure 3, its architecture is divided into the mobile app part and the cloud server part. The mobile app consists of a camera application programming interface that captures LIF oral images via the device and a WebViewer that displays oral hygiene over time on the smartphone screen. The cloud server is built on Amazon Web Services, which has abundant computing resources such as CPU and GPU. It contains an Apache http server, which process requests and provides web assets and content over http protocol, a MariaDB, which transforms data into structured information, an Analysis server, which analyzes oral hygiene through deep learning algorithms, and web storage, which stores original images and analyzed results.

Figure 3. Architecture of the hybrid mobile app comprising mobile app and cloud server.



XSL•FO
RenderX

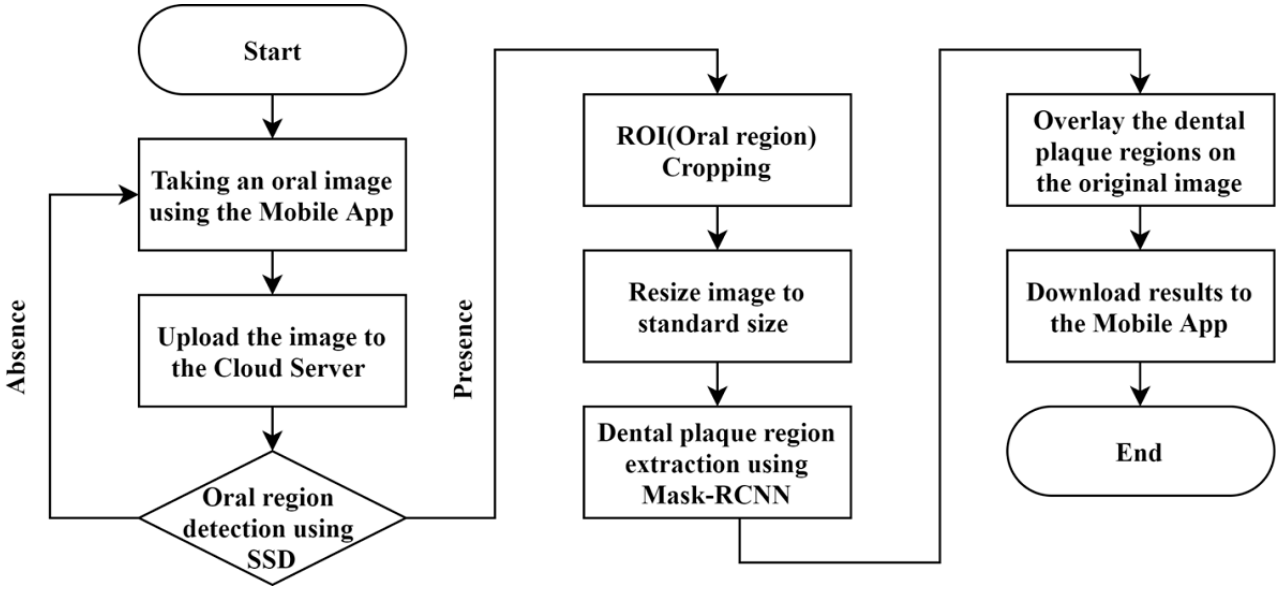
Oral Image Data Flow on Cloud Server

An oral image taken with the user's smartphone camera is uploaded to the http server implemented in Apache using the http protocol. The http servers use load balancing technology to send images to the deep learning–based analysis servers, reducing the load on analyzing large numbers of oral images from multiple users. After analysis is completed, the result image is converted into thumbnails to reduce the amount of transmission and reception and then sent to the user's smartphone screen to check the oral hygiene information. Information such as account, date, plaque location, and cleanliness generated during the oral hygiene analysis are stored and updated in the database on the MariaDB as time series data.

Deep Learning–Based Algorithm for Oral Hygiene Analysis

The oral hygiene analysis is performed through deep learning–based image processing algorithms including object detection, which determines whether the input image is an oral image and localizes the oral region, and instance segmentation, which extracts the dental plaque regions. In this analysis, datasets categorized into oral and nonoral images are used. The 2000 oral images are taken only via the device and stored on the server, and the 2000 nonoral images are Pascal visual object classes images [23] without teeth or gums. Figure 4 illustrates a flowchart of the oral hygiene analysis.

Figure 4. Flowchart of oral hygiene analysis.



For the object detection, the ground-truth bounding box annotations are first performed, which outline the region of interest (RoI) involving only the teeth and gums for each image. Then the single shot multibox detector (SSD) [24], one of the most popular deep learning models for object detection, is applied to detect the RoI within the images (predicted bounding box). The SSD model training is performed with a training dataset after dividing the datasets into a 2000-image training dataset (1000 oral and 1000 nonoral images) and a 2000-image test dataset (1000 oral and 1000 nonoral images). In order to improve the learning performance, the training dataset is augmented with random sample crop, photometric distortion (random transformations in the color HSV domain), rotation, and mirroring. Then the resolutions of the images in the dataset are all converted to 300×300 pixels, and their RGB values are normalized between 0 and 1.

In order to finely extract the red fluorescence–emitted dental plaque region from the oral image, an instance segmentation technique capable of classification and detection of multiple instances in one class is required.

In this study, the Mask region-based convolutional neural network (R-CNN) [25], which is a Faster R-CNN [26] with the addition of a small fully convolutional network that can act as object detection and mask segmentation for each RoI, is used as an instance segmentation technique. The RoI images extracted through the SSD model are used as the input data for Mask R-CNN model training.

Pixel-level annotation for Mask R-CNN training was performed on plaque areas emitting red fluorescence in tooth images according to three criteria:

1. Connected components labeling, which groups its pixels into components based on pixel connectivity, is performed on plaque areas
2. Plaque areas that span several teeth are divided at tooth boundary
3. Plaque thinly distributed between teeth is classified as one instance without following the second criterion

The images are technically prepared in the Common Objects in Context (COCO) data format [27] and augmented with rotation and aspect ratio conversion.

The Mask R-CNN model uses a model pretrained with COCO data as an initial parameter, and the training of the model is performed by the loss function–based stochastic gradient descent method.

This research protocol was approved by the institutional review board (IRB #ERI19046), Seoul National University Dental Hospital. In order to protect users' privacy information, no

XSL•FO
RenderX

personally identifiable information such as name, age, or gender are included in the image, and except for the mouth and the device, the visible parts of the image were mosaicized to make them indistinguishable.

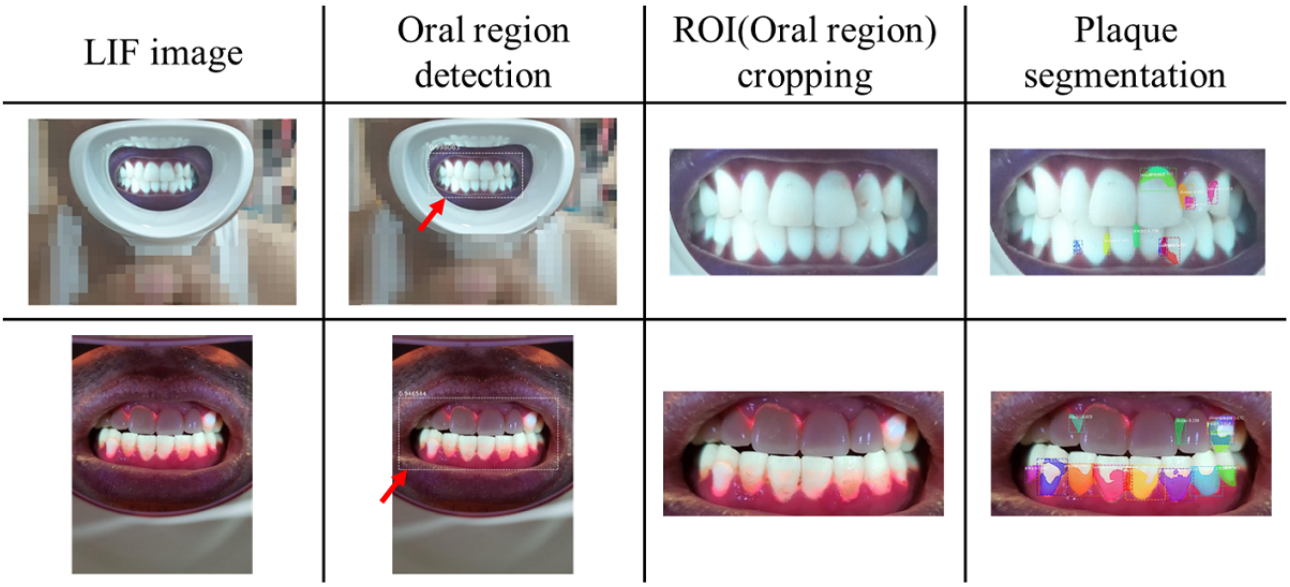
Results

System Characterization

The LIF device consists of a body for easy grip, two LEDs for 400 to 410 nm light emission, filter for color correction, and hood for blocking ambient light, as shown in Figure 2. The electrical features of the device are power consumption of about 2.1 W, current of about 468 mA with 4.5 V, and idle current of about 4.5 mA. When using three AAA batteries with a capacity of 1500 mAh in series, continuous use time is about 150 minutes, and an individual can use it for about 75 days if it is used for 2 minutes per day.

The hybrid mobile app has a scrollable web view, allowing users to easily observe changes in dental hygiene status using time series data as shown in Figure 1(d). Oral hygiene analysis is then performed by taking an oral image through the app and sending it to the cloud server (Figure 3). After analysis, the oral hygiene results are sent back to the user's app. It takes 3.00 (SD 0.020) seconds for the deep learning graphics processing unit and other libraries to load, 1.69 (SD 0.019) seconds for the oral region detection process (SSD), and 4.38 (SD 0.024) seconds for the dental plaque region extraction process (Mask R-CNN). Figure 5 presents representative results obtained during object detection and instance segmentation.

Figure 5. Representative images obtained during deep learning–based oral hygiene analysis. The red arrow points to the predicted bounding box.



Experimental Analysis of Plaque Detection Algorithm

Oral hygiene analysis was performed through object detection and instance segmentation, and each performance was verified using intersection over union (IoU) and average precision (AP). IoU is an evaluation metric used to verify the performance of object detection. Its value is calculated by dividing the area of overlap by the area of union of both the ground-truth bounding box and predicted bounding box. AP, also a metric for evaluating the accuracy of the object detector, is the average of precision values corresponding to recall values between 0 and 1. Here, recall indicates how well all positives are predicted, and precision indicates the accuracy of the predicted result. The object detection performance of SSD is expressed by AP, which is an average value for IoU threshold. The AP values at IoU thresholds of 0.50, 0.60, 0.70, 0.80, and 0.90 are 90.9, 81.8, 63.6, 42.0, and 9.1, respectively. The average over IoU thresholds from 0.50 to 0.95 with a step size of 0.05 is 53.31. The performance of Mask R-CNN is expressed as IoU for the whole segmentation result instead of AP because the dental plaque regions are atypical, and the result value is 0.31.

Discussion

Principal Findings

In this paper, we presented a deep learning–based oral hygiene monitoring system consisting of a LIF device and hybrid mobile app to facilitate oral hygiene at home using a smartphone. The most prominent feature of the LIF device is the filter. Filters of similar devices are manufactured by depositing a thin dielectric layer onto a glass substrate, while the filter of our device is simply manufactured by adopting a method of mixing colors with poly(methyl methacrylate). Thus, the filter of our device is relatively low in manufacturing cost while maintaining the main role of the filter to make the red fluorescence of the plaque stand out compared with the surrounding tooth color. However, performance of the filter is affected by the intensity of ambient light [28], so a hood is required to minimize degradation by ambient light. Although our device also has a black hood, it has been observed that overexposed images are taken because the hood cannot completely block ambient light outdoors or under strong lighting. These results are caused by light leakage due to different oral structures of individuals, and it is expected that these problems will be sufficiently solved when the hood is made of flexible materials or customized. Another source of

XSL•FO
RenderX

ambient light is the large window for visual observation or camera shooting. This is fundamentally unable to block ambient light, but it is believed that providing a light-blocking agent will minimize the influence of ambient light.

The key feature of our hybrid mobile app is that it is programmed with a deep learning algorithm. Since most conventional LIF products are developed for medical purposes, nonmedical users require the help or education of a medical practitioner. On the other hand, our mobile app automatically determines the location and distribution of dental plaques without clinical examination or training. In addition, since the result is stored and displayed as time series data, it is convenient for the user to manage oral hygiene. The deep learning models for oral detection and dental plaque segmentation require high computing resources, but by performing analysis on the server, computing resources can be effectively managed, and results can be quickly generated.

As shown in Figure 6, our segmentation algorithm is somewhat poor in performance. In order to determine the cause, 2000 images were randomly sampled from the images stored on the server and statistically analyzed. Images were categorized as normal, out-focusing, far away, overexposed, too dark, foggy, yawning, and no device. Each category was multiselected, and the results are shown in Table 1. Multimedia Appendix 1 shows representative images categorized in each category.

Figure 6. Representative poor results of each deep learning algorithm.

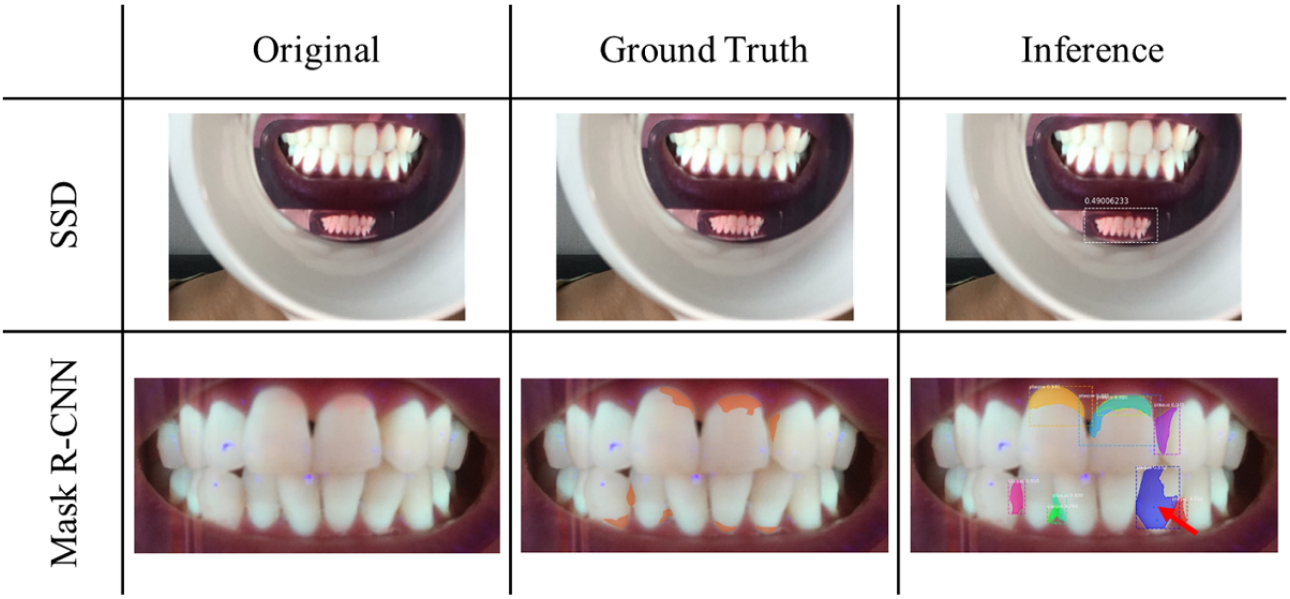


Table 1. Statistical results of the quality of the acquired images.

Category	Value n (%)
Normal	627 (31.4)
Out-focusing	882 (44.1)
Far away	501 (25.1)
Overexposed	489 (24.5)
Too dark	297 (14.9)
Foggy	249 (12.5)
Yawning	223 (11.2)

Of the images from the server, only 31.35% (627/2000) were classified as normal, meaning that most images were described with one or more variables such as out-focusing, far away, overexposed, too dark, foggy, and yawning, which is thought to directly affect the performance of deep learning–based analysis. In addition to these cases, two tooth areas were found by reflecting the tooth area on the smartphone case, and the plaque area was not properly divided when the captured image had a red tone. In the future, if we provide guidelines to users on an appropriate environment for acquiring images and improve the algorithms based on continuously increasing numbers of images as the number of users of the product increases, we expect that plaque identification will improve [29].

Limitations

Due to the different image sensor characteristics of smartphones, a consistent preprocessing method could not be applied. The other limitation was that the size of the training data was smaller than the collected data due to various optical environments.

Conclusion

The primary cause of dental caries and periodontal disease is the failure to remove dental plaque in a timely manner.

XSL•FO
RenderX

Providing preventive care solutions to quickly identify and respond to dental plaques at home can significantly reduce social costs associated with oral disease. The LIF system introduced in this paper consists of a LIF device for visually identifying dental plaques and a mobile app for providing deep learning–based oral hygiene analysis results. The device allows the user to visually check oral hygiene in a mirror and the app motivates the user to perform oral hygiene management by providing the oral hygiene analysis results in time series. In this paper, we introduced a home oral care system, but in the future, we will introduce LIF-based medical devices for marginalized populations, including the elderly, people of lower socioeconomic standing, and those living where the internet is unavailable, by applying edge computing technique and developing low-cost devices in the form of smartphone accessories.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1I1A1A01057598 and NRF-2020R1F1A1070042). LIF device (Dentinote) and mobile app (Dentinote app) were developed with Prizzmable Co Ltd, and this paper is based on the accumulated results from the development process.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Representative images categorized into each category: (a) normal, (b) out-focusing, (c) far away, (d) overexposured, (e) foggy, and (f) yawning.
[PNG File , 737 KB - mhealth_v8i10e17881_app1.png]

References

1. Becker W, Becker BE, Berg LE. Periodontal treatment without maintenance: a retrospective study in 44 patients. J Periodontol 1984 Sep;55(9):505-509. [doi: 10.1902/jop.1984.55.9.505] [Medline: 6592322]
2. Lang NP, Ostergaard E, Löe H. A fluorescent plaque disclosing agent. J Periodontal Res 1972;7(1):59-67. [doi: 10.1111/j.1600-0765.1972.tb00632.x] [Medline: 4272033]
3. Agarwal V, Zhang L, Zhu J, Fang S, Cheng T, Hong C, et al. Impact of predicting health care utilization via web search behavior: a data-driven analysis. J Med Internet Res 2016 Sep 21;18(9):e251 [FREE Full text] [doi: 10.2196/jmir.6240] [Medline: 27655225]
4. Pretty IA, Edgar WM, Smith PW, Higham SM. Quantification of dental plaque in the research environment. J Dent 2005 Mar;33(3):193-207. [doi: 10.1016/j.jdent.2004.10.017] [Medline: 15725520]
5. Datta DD. Disclosing solutions used in dentistry. World J Pharmaceut Res 2017 Jun 1:1648-1656. [doi: 10.20959/wjpr20176-8727]
6. Heinrich-Weltzien R, Kühnisch J, van der Veen M, de Josselin de Jong E, Stösser L. Quantitative light-induced fluorescence (QLF): a potential method for the dental practitioner. Quintessence Int 2003 Mar;34(3):181-188. [Medline: 12731599]
7. Thomas RZ, van der Mei HC, van der Veen MH, de Soet JJ, Huysmans MC. Bacterial composition and red fluorescence of plaque in relation to primary and secondary caries next to composite: an in situ study. Oral Microbiol Immunol 2008 Feb;23(1):7-13. [doi: 10.1111/j.1399-302X.2007.00381.x] [Medline: 18173792]
8. Coulthwaite L, Pretty IA, Smith PW, Higham SM, Verran J. The microbiological origin of fluorescence observed in plaque on dentures during QLF analysis. Caries Res 2006;40(2):112-116. [doi: 10.1159/000091056] [Medline: 16508267]
9. Kim Y, Lee E, Kwon H, Kim B. Monitoring the maturation process of a dental microcosm biofilm using the Quantitative Light-induced Fluorescence-Digital (QLF-D). J Dent 2014 Jun;42(6):691-696. [doi: 10.1016/j.jdent.2014.03.006] [Medline: 24657554]
10. Benedict HC. A note on the fluorescence of teeth in ultra-violet rays. Science 1928 Apr 27;67(1739):442. [doi: 10.1126/science.67.1739.442] [Medline: 17731214]
11. Van Denbergh A. On porphyrin in the mouth. Lancet 1928 Feb;211(5450):281-282. [doi: 10.1016/s0140-6736(00)99843-x]
12. Lennon AM, Buchalla W, Brune L, Zimmermann O, Gross U, Attin T. The ability of selected oral microorganisms to emit red fluorescence. Caries Res 2006;40(1):2-5. [doi: 10.1159/000088898] [Medline: 16352873]
13. van der Veen MH, Thomas RZ, Huysmans MC, de Soet JJ. Red autofluorescence of dental plaque bacteria. Caries Res 2006;40(6):542-545. [doi: 10.1159/000095655] [Medline: 17063027]
14. Volgenant CMC, van der Veen MH, de Soet JJ, ten Cate JM. Effect of metalloporphyrins on red autofluorescence from oral bacteria. Eur J Oral Sci 2013 Jun;121(3 Pt 1):156-161. [doi: 10.1111/eos.12045] [Medline: 23659237]

XSL•FO
RenderX

15. Terrer E, Koubi S, Dionne A, Weisrock G, Sarraquigne C, Mazuir A, et al. A new concept in restorative dentistry: light-induced fluorescence evaluator for diagnosis and treatment. Part 1: Diagnosis and treatment of initial occlusal caries. J Contemp Dent Pract 2009 Nov 01;10(6):E086-E094. [Medline: 20020086]
16. van der Veen MH, Volgenant CMC, Keijser B, Ten Cate JBM, Crielaard W. Dynamics of red fluorescent dental plaque during experimental gingivitis: a cohort study. J Dent 2016 May;48:71-76 [FREE Full text] [doi: 10.1016/j.jdent.2016.02.010] [Medline: 26921667]
17. Rechmann P, Liou SW, Rechmann BMT, Featherstone JDB. Performance of a light fluorescence device for the detection of microbial plaque and gingival inflammation. Clin Oral Investig 2016 Jan;20(1):151-159. [doi: 10.1007/s00784-015-1481-9] [Medline: 25914048]
18. Volgenant CMC, Fernandez Y Mostajo M, Rosema NAM, van der Weijden FA, Ten Cate JM, van der Veen MH. Comparison of red autofluorescing plaque and disclosed plaque: a cross-sectional study. Clin Oral Investig 2016 Dec;20(9):2551-2558 [FREE Full text] [doi: 10.1007/s00784-016-1761-z] [Medline: 26993658]
19. Chen QG, Zhu HH, Xu Y, Lin B, Chen H. Quantitative method to assess caries via fluorescence imaging from the perspective of autofluorescence spectral analysis. Laser Phys 2015 Jul 08;25(8):085601. [doi: 10.1088/1054-660x/25/8/085601]
20. Karlsson L. Caries detection methods based on changes in optical properties between healthy and carious tissue. Int J Dent 2010;2010:270729 [FREE Full text] [doi: 10.1155/2010/270729] [Medline: 20454579]
21. Bakator M, Radosav D. Deep learning and medical diagnosis: a review of literature. Multimodal Technol Interact 2018 Aug 17;2(3):47. [doi: 10.3390/mti2030047]
22. Imangaliyev S, van der Veen MH, Volgenant C, Loos B, Keijser B, Crielaard W. Classification of quantitative light-induced fluorescence images using convolutional neural network. arXiv preprint arXiv. 2017. URL: https://arxiv.org/abs/1705.09193 [accessed 2020-09-24]
23. Everingham M, Van Gool L, Williams CKI, Winn J, Zisserman A. The Pascal visual object classes (VOC) challenge. Int J Comput Vis 2009 Sep 9;88(2):303-338. [doi: 10.1007/s11263-009-0275-4]
24. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY. SSD: single shot multibox detector. Comput Vision 2016:21-37. [doi: 10.1007/978-3-319-46448-0_2]
25. He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. IEEE Trans Pattern Anal Mach Intell 2020 Feb;42(2):386-397. [doi: 10.1109/TPAMI.2018.2844175] [Medline: 29994331]
26. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 2017 Jun;39(6):1137-1149. [doi: 10.1109/TPAMI.2016.2577031] [Medline: 27295650]
27. Lin T, Maire M, Belongie S, Hays J, Perona P, Ramanan D. Microsoft COCO: common objects in context. Comput Vision 2014:740-755. [doi: 10.1007/978-3-319-10602-1_48]
28. Pretty IA, Edgar WM, Higham SM. The effect of ambient light on QLF analyses. J Oral Rehabil 2002 Apr;29(4):369-373. [doi: 10.1046/j.1365-2842.2002.00914.x] [Medline: 11966971]
29. Zappone A, Di Renzo M, Debbah M. Wireless networks design in the era of deep learning: model-based, AI-based, or both? IEEE Trans Commun 2019 Oct;67(10):7331-7376. [doi: 10.1109/tcomm.2019.2924010]

Abbreviations

AP: average precision
COCO: Common Objects in Context
IoU: intersection over union
LED: light-emitting diode
LIF: light-induced fluorescence
R-CNN: region-based convolutional neural network
RoI: region of interest
SSD: single shot multibox detector

Edited by G Eysenbach; submitted 18.01.20; peer-reviewed by E Frontoni, L Zhang, G Lim; comments to author 20.04.20; revised version received 05.06.20; accepted 01.09.20; published 16.10.20.

Please cite as:
Kim JM, Lee WR, Kim JH, Seo JM, Im C
Light-Induced Fluorescence-Based Device and Hybrid Mobile App for Oral Hygiene Management at Home: Development and Usability Study
JMIR Mhealth Uhealth 2020;8(10):e17881
URL: https://mhealth.jmir.org/2020/10/e17881
doi:10.2196/17881
PMID:33064097

XSL•FO
RenderX

©Jun-Min Kim, Woo Ram Lee, Jun-Ho Kim, Jong-Mo Seo, Changkyun Im. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 16.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Feasibility and Acceptability of a Counseling- and mHealth-Based Physical Activity Intervention for Pregnant Women With Diabetes: The Fit for Two Pilot Study

Britta Larsen[1], PhD; Stephanie Micucci[1], BS; Sheri J Hartman[1], PhD; Gladys Ramos[2], MD

[1]Department of Family Medicine & Public Health, University of California, San Diego, San Diego, CA, United States

[2]Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California, San Diego, San Diego, CA, United States

Corresponding Author:
Britta Larsen, PhD
Department of Family Medicine & Public Health
University of California, San Diego
9500 Gilman Dr
San Diego, CA, 92093-0725
United States
Phone: 1 8585348429
Email: blarsen@ucsd.edu

Abstract

Background: Diabetes during pregnancy poses serious health risks to both mother and child. Regular physical activity can reduce these risks, yet few clinic-based interventions of physical activity for pregnant women with diabetes have been attempted.

Objective: The purpose of this single-arm pilot trial is to assess the feasibility and acceptability, and explore the potential efficacy of a counseling- and mobile health–based physical activity intervention for pregnant women with diabetes.

Methods: Participants (N=17) who had type 2 or gestational diabetes, could read and speak in English or Spanish, and were between 10 and 27 weeks of gestation were recruited from the University of California San Diego Diabetes and Pregnancy Program. Participants engaged in a one-on-one counseling and goal-setting session immediately following a clinic visit with their physician. They were given a Fitbit and shown how to use the Fitbit app, including entering personalized step goals, and were encouraged to build up to 10,000 daily steps. Daily steps were recorded for 12 weeks, until they were 36 weeks' gestation, or until 1 week before they gave birth, whichever came first. Feasibility was measured by recruitment, retention, and adherence, and acceptability was measured using consumer satisfaction questionnaires and follow-up interviews. Potential efficacy was explored by examining changes in daily steps over time.

Results: The participants were primarily Hispanic (13/17, 76%), had public insurance (15/17, 88%), and had type 2 diabetes (12/17, 71%). Of the 17 patients who began the intervention, 76% (13/17) completed a follow-up visit, and 71% (12/17) continued wearing the Fitbit regularly after 8 weeks in the intervention. Adherence in wearing the Fitbit was relatively high, with a median wear adherence of 90% of days. The intervention was generally well accepted, with 85% (11/13) indicating that they were motivated to exercise more following the counseling session, 85% (11/13) indicating that the Fitbit helped increase their activity, and 92% (12/13) recommending the program overall. Mean daily steps increased from baseline (mean 6122, SD 2439) to week 3 (mean 6269, SD 2166) and then decreased through week 12 (mean 4191, SD 2228).

Conclusions: High acceptability, retention, and adherence suggest that this may be a promising approach to delivering a simple, low-burden intervention in a clinical setting to a high-risk, underserved population. A randomized controlled trial is needed to determine whether this approach is effective in slowing the reduction in activity typically seen throughout pregnancy.

Trial Registration: ClinicalTrials.gov NCT03302377; https://clinicaltrials.gov/ct2/show/NCT03302377

(JMIR Mhealth Uhealth 2020;8(10):e18915) doi:10.2196/18915

KEYWORDS

exercise; behavioral medicine; mHealth; gestational diabetes; type 2 diabetes; mobile phone

XSL•FO
RenderX

Introduction

Diabetes during pregnancy can have serious health consequences for both mother and child. Women who experience type 2 diabetes (T2D) during pregnancy are less likely to return to prepregnancy weight and are more likely to experience glycemic dysregulation following pregnancy [1]. Infants exposed to T2D in utero have approximately 9 times higher risk of perinatal mortality and are at a higher risk for congenital malformation and developing obesity and T2D later in life [2-4]. Given the rapidly rising rates of diabetes among young adults of childbearing age [5] and estimates that diabetes prevalence will more than double in the United States by 2050 [6] and more than quadruple in youth [7], this constitutes a critical public health issue.

Physical activity has robustly been shown to prevent T2D [8], and limited clinical trial research suggests that it can normalize glucose and prevent insulin use in pregnant women with diabetes [9-11]. The American College of Obstetrics and Gynecologists concordantly now recommends an individualized exercise prescription for all pregnant women with diabetes [12]. Intervening on physical activity during pregnancy is much needed, as activity typically declines during pregnancy [13,14], particularly for women who were initially underactive [15]. Delivery of prenatal exercise counseling, however, is inconsistent and often absent, as most providers lack time or training to provide all the tools necessary to help patients change physical activity [16,17].

Advances in technology, however, have made the dissemination of individualized physical activity interventions more feasible [18,19]. Commercial wearable activity trackers and accompanying smartphone apps use features that have consistently been associated with successful behavioral changes, such as goal setting, self-monitoring, reminders, and social accountability [20]. Accordingly, interventions combining wearable trackers with brief counseling have shown significant improvements in physical activity [21,22]. As wearable devices are not only effective but also low cost relative to other clinical interventions, they could be especially appropriate for incorporation into clinical care, although studies of the efficacy, feasibility, and acceptability of using wearable devices with clinical populations remain sparse. Some small pilot studies have found that using consumer wearable devices may be feasible and effective in increasing physical activity in pregnant women [23]; however, no such studies have been performed to date in pregnant women with diabetes.

Given the growing rates of diabetes in pregnancy and the significant risks associated with this condition, this is a pressing public health issue, and developing effective, low-cost physical activity interventions with potential for implementation in clinical and community settings is essential for promoting healthy pregnancies and infants. Therefore, we designed a low-cost, low-burden physical activity intervention combining an in-person counseling session with the use of a consumer wearable tracker and mobile phone app. The purpose of this study is to evaluate the feasibility and acceptability of this physical activity intervention for pregnant women with diabetes.

Methods

Study Overview

The *Fit for Two* study was a single-arm pilot trial to assess the feasibility and acceptability of a physical activity intervention in a clinical setting for pregnant women with T2D or gestational diabetes (GD). The intervention comprised an in-person counseling session followed by the use of a Fitbit wrist monitor and app for 8-12 weeks to reinforce key behavior change strategies. Physical activity was measured throughout the trial via Fitbit monitors, and feedback on the intervention was provided at the study conclusion via consumer satisfaction questionnaires and individual interviews.

Participants

Participants were pregnant women receiving prenatal care at the University of California San Diego (UCSD) Diabetes and Pregnancy Program (DAPP). Women were eligible to enroll if they (1) were currently pregnant and between 10 and 27 weeks of gestation, (2) were aged 18 years or older, (3) had daily access to a smartphone, (4) were underactive (engaged in less than 100 min per week of at least moderate-intensity activity), (5) could read and speak in English or Spanish, and (6) were diagnosed with either T2D or GD. Patients with T2D were those who had previously been diagnosed with T2D before pregnancy, whereas those with GD were diagnosed with diabetes during pregnancy. Women were ineligible if they had any condition that would make exercise unsafe as determined by their physician or if they were concurrently participating in another behavioral or pharmaceutical trial. The study was approved by the UCSD institutional review board, and all participants provided written informed consent.

Recruitment

Recruitment occurred during DAPP clinic visits. Using electronic medical records, Health Insurance Portability and Accountability Act–certified study staff checked weekly patient visit schedules to identify patients who met the eligibility criteria. The patient's physician was asked to confirm whether she could safely participate in a physical activity study or whether she had any contraindications. At the end of their clinic visits, potentially eligible patients were asked by the attending physician whether they were interested in learning about a study that could help them increase their physical activity. Study staff was on-site to discuss the study with interested participants and complete a screening interview to confirm eligibility.

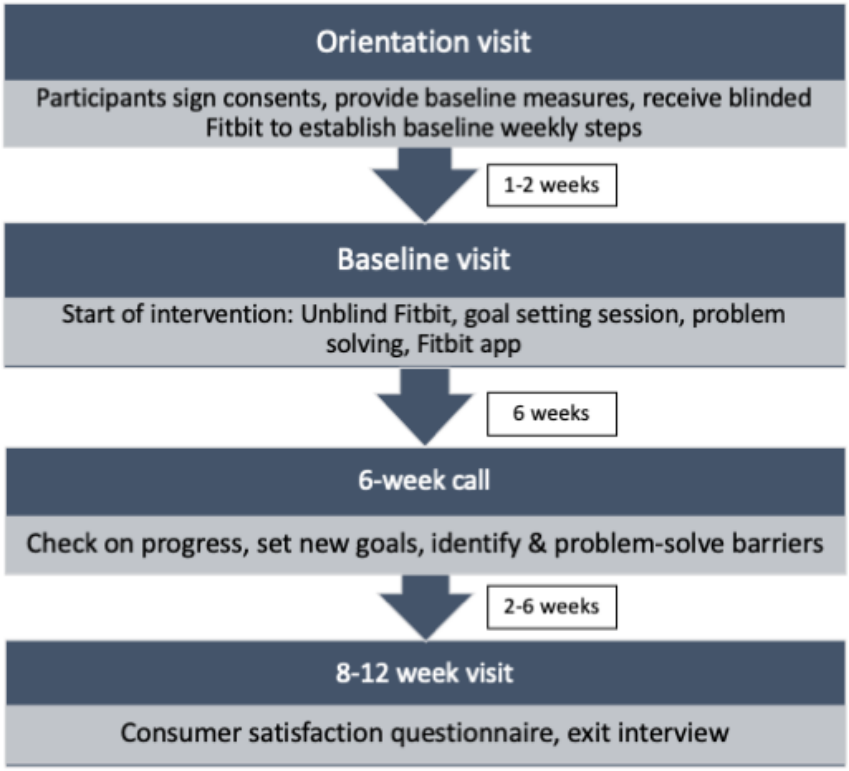
Protocol Overview

As most patients in the DAPP clinic have prenatal care appointments scheduled every 1-3 weeks for glucose review, study visits were designed to coincide with the scheduled clinic visits. Participants attended an orientation visit in an exam room at the clinic to learn more details about the study and give informed consent. Participants who were Spanish-dominant met with a bilingual researcher and completed consent and all forms and study visits in Spanish. They were then given a blinded Fitbit Alta HR wrist-worn activity monitor to wear for 1 week to establish baseline steps and activity (see the *Measures* section). Following their next prenatal care visit 1-3 weeks later,

XSL•FO
RenderX

the participants started the physical activity intervention. Figure 1 shows a timeline of the intervention activities.

Figure 1. Intervention timeline.



Intervention Description

The *Fit for Two* intervention was centered on evidence-based behavior change strategies such as goal setting, self-monitoring, accountability, and social support. The intervention comprised an in-person counseling session to teach key behavior change constructs and initiate behavior change, followed by continued support through the Fitbit app and activity monitor to reinforce behavior change strategies. This allowed for both depth and breadth of the intervention content with relatively little staff time or clinic space required.

During the in-person counseling session at the start of the intervention, participants learned more about the importance of physical activity during pregnancy and the types and amounts of activity they should engage in. The interventionist then unblinded the participant's Fitbit data from the previous week to review their current steps and activity. Using the principles of motivational interviewing, the interventionist helped the participant set her personal goals for daily steps, identify potential barriers to achieving goals, and outline problem-solving strategies. The data were used to set realistic goals for daily steps and discuss potential barriers and problem solving around them using principles of motivational interviewing. The study staff then helped participants learn to use the Fitbit monitor and app to support self-monitoring and get feedback toward goals and reinforcement when goals are met through congratulatory messages, visual displays on the app, and a digital firework display on the wrist monitor. Participants were encouraged to set their own new goals regularly to gradually build up to 10,000 steps at whatever pace was right for them and keep their physician informed of their progress by reporting their average daily steps per week in their weekly email reporting daily glucose levels. For the duration of the intervention (8-12 weeks depending on the timing of pregnancy), participants were instructed to wear the Fitbit wrist monitor daily and synchronize it with the Fitbit app at least weekly. Fitabase (Small Steps Lab, LLC), a third-party software platform that extracts the Fitbit data from authorized users in real time, was used by the study staff to monitor syncing and provided text reminders if participants had not synced their Fitbit monitor with the app for more than 2 weeks.

Participants also received a brief (10-min) phone call after 6 weeks to review progress and report any problems.

Measures

Basic demographics were collected via questionnaires at baseline and included age, race, ethnicity, education, employment, income, marital status, and the number of children. Medical data were collected from participants' electronic medical records both at baseline and follow up and included current weight, weight gain since pregnancy, week of gestation, resting blood pressure (systolic and diastolic), hemoglobin A_{1c} (HbA_{1c}), and nonfasting glucose.

Consumer Satisfaction

We used a questionnaire to assess the feasibility and acceptability of various intervention components, including the goal-setting session, educational material, the Fitbit monitor, and the smartphone app. It was available in English or Spanish; the measure has been translated and back-translated in Spanish

XSL•FO
RenderX

and adapted from previous interventions with Spanish-speaking individuals [24]. Participants were also invited to engage in a one-on-one follow-up interview to give more detailed feedback about their experience in the program.

Physical Activity

Daily steps were measured throughout the study using the Fitbit Alta HR monitor. It is a wrist-worn, combined sensing (ie, accelerometry and photoplethysmography) activity tracker that measures physical activity (ie, intensity, energy expenditure, bouts of exercise, steps, and distance traveled) at varying resolutions ranging from 1 second to 1 min daily. The Fitbit app automatically syncs data from the wrist monitor to a smartphone app via Bluetooth, which was collected and stored through Fitabase. A valid day of Fitbit wear was defined as ≥600 mins (10 hours) of heart rate data, or for participants with inconsistent or missing heart rate data, ≥6000 steps. A partial wear day was defined as any day with ≥1000 steps. The literature has shown that days with <1000 steps are unlikely to show actual wear [25]. To prevent feedback from influencing baseline activity, the displays on the wrist monitor and app were then blinded by removing all information tiles from the app and setting the wrist monitor to only show time, battery life, and distance. The Fitbit has shown excellent reliability and validity for measuring daily steps [26].

Participants were also given an ActiGraph GT3X+ hip-worn accelerometer to wear for 1 week at baseline and follow up to provide a measure of activity at various intensities. However, a high percentage of participants (13/17, 76%) commented that the belt-worn accelerometer was markedly uncomfortable during pregnancy and requested to stop wearing it; thus, this measure was discontinued.

Analysis

Feasibility was defined by 3 factors: (1) recruitment if at least 50% of the women who expressed an interest in the study enrolled, (2) retention if at least 75% of the participants who attended a baseline session completed the follow-up assessment and/or interview, and (3) adherence if the participants wore and synced the Fitbit on at least 75% of days. Acceptability was defined as at least 75% of the respondents indicated that they were *satisfied* or *very satisfied* with the intervention overall and individual intervention components.

As the primary aim was assessing feasibility and acceptability, we did not calculate the power to determine efficacy, but we did evaluate changes in activity from baseline to follow up (week 12) to explore the potential implications for efficacy. Daily steps were converted into average daily steps for the baseline period (the week immediately following the orientation visit) and each week of the trial following the baseline intervention visit. We used 2 separate average daily step variables to plot activity and explore trajectories of behavior change across weeks in the intervention and across gestational weeks: (1) days with any partial Fitbit wear time and (2) days where the Fitbit was worn the entire day. For each of these figures, we present group means plotted with a local polynomial regression fitting or *loess* function to examine nonlinear changes across follow up. Analyses were performed using R Studio version 1.2.5033.

Results

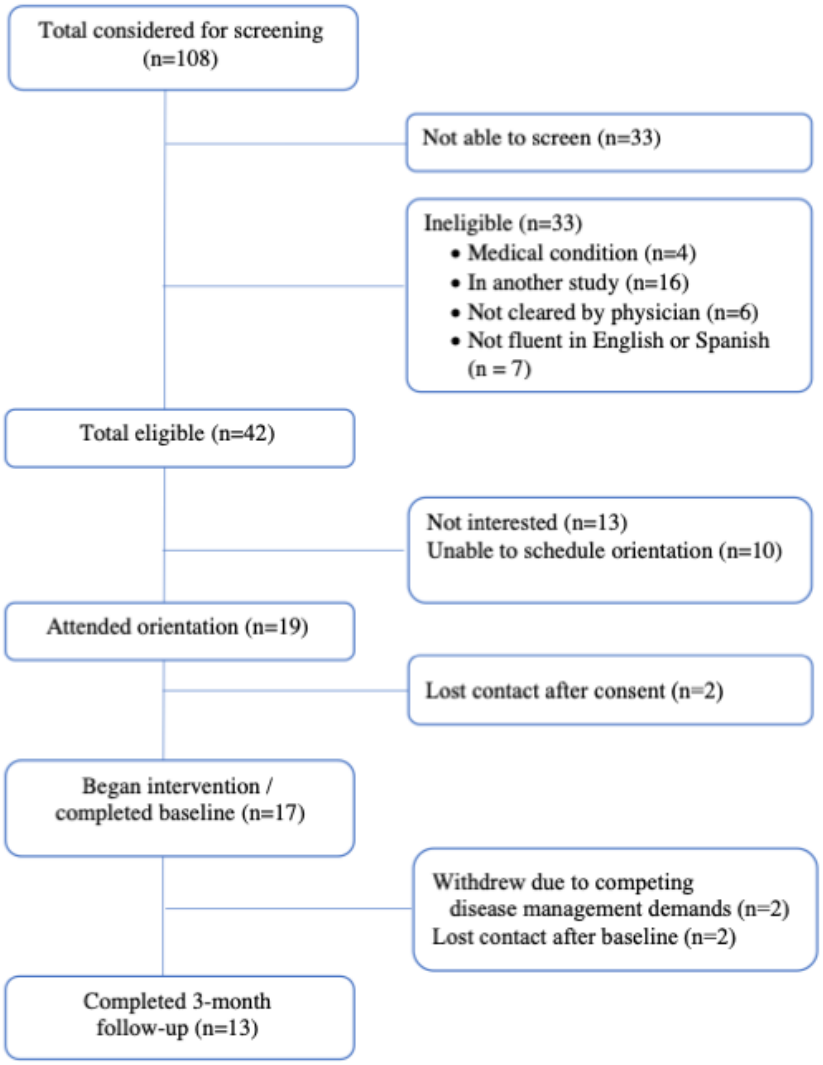
Feasibility

Recruitment

The flow of recruitment is shown in Figure 2. A total of 108 participants had visits scheduled in the clinic during the study period and were considered for participation based on chart review. Approximately one-third of the 108 patients were not screened in person because of missed appointments. Of the 75 patients screened, 42 met the eligibility criteria. The most common reasons for ineligibility were concurrent participation in another study (n=16), not speaking English or Spanish (n=7), and not receiving physician clearance for unsupervised exercise (n=6).

XSL•FO
RenderX

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram.



Of the 42 patients meeting eligibility criteria, 13 declined participation, and 10 expressed interest but were unable to schedule an orientation visit. The most common reason given for declining participation and being unable to schedule a visit was lack of time because of competing demands of diabetes management and prenatal care. Of the 19 participants who attended the orientation session and signed consent forms, 17 completed a baseline intervention visit and began the program. Only 2 participants withdrew following the orientation visit, citing the competing demands of diabetes management.

Baseline

Table 1 shows the baseline characteristics of the 17 participants who began the intervention. Participants were, on average, approximately 30 years old (mean 29.8, SD 3.85; range 23-36), and most were Hispanic (13/17, 76%). Nearly half (8/17, 46%) had no employment, and almost all (15/17, 88%) had Medi-Cal public insurance. Participants were a mix of primigravida and multiparous, with approximately two-thirds indicating that they had other children at home. Most (12/17, 71%) had T2D, and most were in the second trimester (mean 20 weeks and 5 days of gestation, SD 4.6). There was a broad range of metabolic markers, with HbA_{1c} ranging from 4.8 to 10%, and nonfasting glucose ranging from 96 to 179 mg/dL.

XSL•FO
RenderX

Table 1. Baseline characteristics of the study sample (N=17).

Demographics	Value
Age, mean (SD)	29.8 (3.85)
Race and ethnicity, n (%)	
Non-Hispanic White	1 (6)
Non-Hispanic Black	3 (18)
Hispanic	13 (76)
Language, n (%)	
English	13 (76)
Spanish	4 (24)
Employment, n (%)	
None	8 (46)
Part-time	3 (18)
Full-time	3 (18)
No data	3 (18)
Insurance type, n (%)	
Private	2 (12)
Public	15 (88)
Education, n (%)	
Less than high school	4 (24)
High school	6 (36)
Some college	6 (36)
No data	1 (4)
Marital status, n (%)	
Single	5 (29)
Divorced/separated	1 (4)
Married/partnered	10 (59)
No data	1 (4)
Other children at home, n (%)	
Any other children at home	10 (66)
Children younger than 5 years at home	6 (40)
Children aged 6-18 years at home	7 (47)
Medical data	
Diabetes type, n (%)	
Type 2	12 (71)
Gestational	5 (29)
BMI (kg/m^2), mean (SD)	37.0 (4.1)
Weight (lbs), mean (SD)	211.5 (39.2)
Systolic blood pressure (mm Hg), mean (SD)	120.8 (17.2)
Diastolic blood pressure (mm Hg), mean (SD)	70.7 (9.4)
HbA_{1C}[a], mean (SD)	6.1 (1.4)
Glucose (mg/dl) mean (SD)	112.5 (18.2)

[a]HbA_{1c}: hemoglobin A_{1c}.

XSL•FO
RenderX

Follow Up

Of the 17 participants who began the intervention, 13 (76%) completed a follow-up visit, and 12 (71%) continued to wear the Fitbit regularly after 8 weeks. A participant experienced technical difficulties and was not able to connect her Fitbit to the Fitabase system, so her data were not captured. Of the 4 participants lost to follow up, 2 withdrew during the intervention because of medical complications/early labor, whereas 2 chose to discontinue the study because of the competing demands of pregnancy. Participants' time in the intervention ranged from 56 to 84 days, depending on the week of gestation upon enrollment and when they delivered.

Adherence was measured by examining how often participants wore their Fitbit monitors. There was great variability in the wear time among participants. Adherence ranged from 27% of days (23/84) to 100% of days (84/84), with a median wear day of 90% (mean 78%, SD 29%), indicating good adherence. However, full days of wear (≥600 min) were much more infrequent and ranged across participants from 0% to 86% of days, with a median of 50% (mean 48%, SD 32%). The median wear time on days with wear was 706 min (mean 749.1, SD 396.4). Limiting wear time to waking hours, median wear time was 643 min on days with any wear (mean 638.9, SD 282.9).

Acceptability

Of the 13 participants who filled out consumer satisfaction questionnaires, 11 (85%) indicated that they found the counseling session quite/extremely helpful, and another 11 (85%) said that they felt moderately/very motivated to start exercising following the counseling session. The same number (11/13, 85%) indicated that they learned *a great deal* about physical activity in the counseling session, and 13 (13/13, 100%) said that they learned something new. With regard to the Fitbit, 11 (11/13, 85%) found it quite/extremely helpful in helping them increase their physical activity, and all but 1 participant (1/13, 8%) indicated that they wore it regularly. All but 1 participant (1/13, 8%) said that they were quite/extremely satisfied with the program overall, and 12 (12/13, 92%) said that they would probably/definitely recommend it to a friend.

A total of 10 participants completed a semistructured interview to provide additional feedback on the program. A common theme from the interviews was an appreciation for the low-burden/light-touch approach, with participants noting that the study encouraged them without feeling too pushy or intrusive:

> *pushed me harder… but didn't put too much pressure on me… you guys were not invasive at all so I was able to go about my day.* [Participant]

Another noted that the study team "did not hover over me… I had my own space."

Another participant said,

> *I also liked the feeling of having a support team without being pushed. I didn't feel forced to be physically active.* [Participant]

Several participants mentioned that they appreciated not feeling *judged* or *ashamed* when they did not meet their goals.

Consistent with the results of the consumer satisfaction survey, another common theme was that the Fitbit was simple and easy to use. A participant shared:

> *I used the Fitbit every day. It was easy to track how many steps I had. Very user-friendly.* [Participant]

Several participants reported that they explored additional features of the Fitbit app that were not part of the intervention but had been mentioned by their physicians, such as tracking calories, hydration, and heart rate.

Participants also noted a number of features of the Fitbit app and wrist monitor that motivated them. Several participants enjoyed the push notifications from the Fitbit app, reporting daily step progress, or reminding them to move. A participant noted:

> *the updates I received letting me know how close or far I was from my goal also helped motivate me even more to try and achieve my goals.* [Participant]

Another appreciated the reinforcement the app provided when she met her goal:

> *it's cool that the Fitbit shows a little rocket when you meet your goal… (my daughter) always wanted to see the rocket, so we walked around the park just to see the rocket.* [Participant]

Several participants also noted that it became more difficult to stay physically active as their pregnancies progressed. A participant noted,

> *I decreased my step goal toward the end of my pregnancy because it became harder to be physically active and meet my step goal.* [Participant]

Another participant recommended adding components to the intervention at the end to enhance motivation later in pregnancy. Another suggested a web-based support group or in a clinic.

Change in Steps

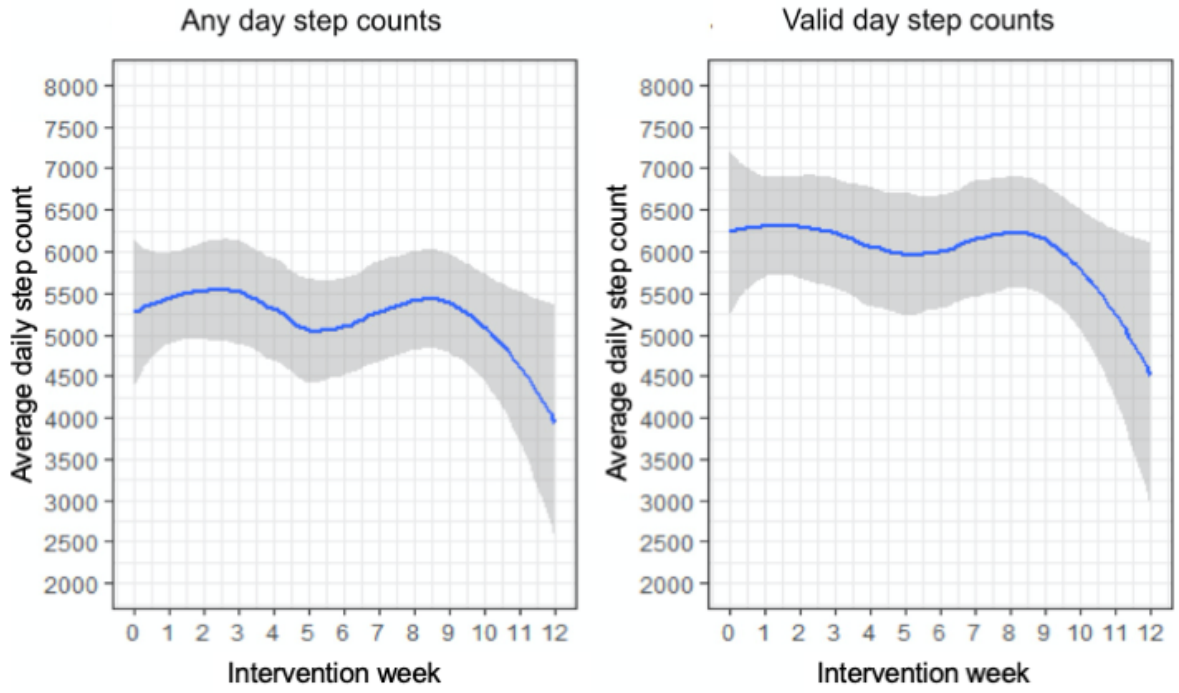
At baseline, the mean number of steps per day was 5281 (SD 1846; Table 2). The average number of steps per day when only considering valid days was higher (mean 6122.1, SD 2356.2 steps per day). The average number of steps slightly increased from baseline to week 3 for both any wear days (mean 5504, SD 1846) and full valid days (mean 6268.6, SD 2081.3; Figure 3). For both any and valid days, the mean daily steps at week 6 were lower (mean 4657, SD 1751.9, and mean 5802.6, SD 2329.1, respectively) before rising again at week 9 (mean 5850.7, SD 2350.6, and mean 6770.4, SD 2563.3, respectively).

XSL•FO
RenderX

Table 2. Average daily steps by week in the study.

Period	Any wear		Valid full day	
	n	Mean (SD)	n	Mean (SD)
Baseline	16	5281 (1846)	15	6122 (2439)
Week 3	14	5504 (1984)	13	6269 (2166)
Week 6	13	4658 (1823)	11	5803 (2443)
Week 12	4	3888 (1551)	4	4191 (2228)

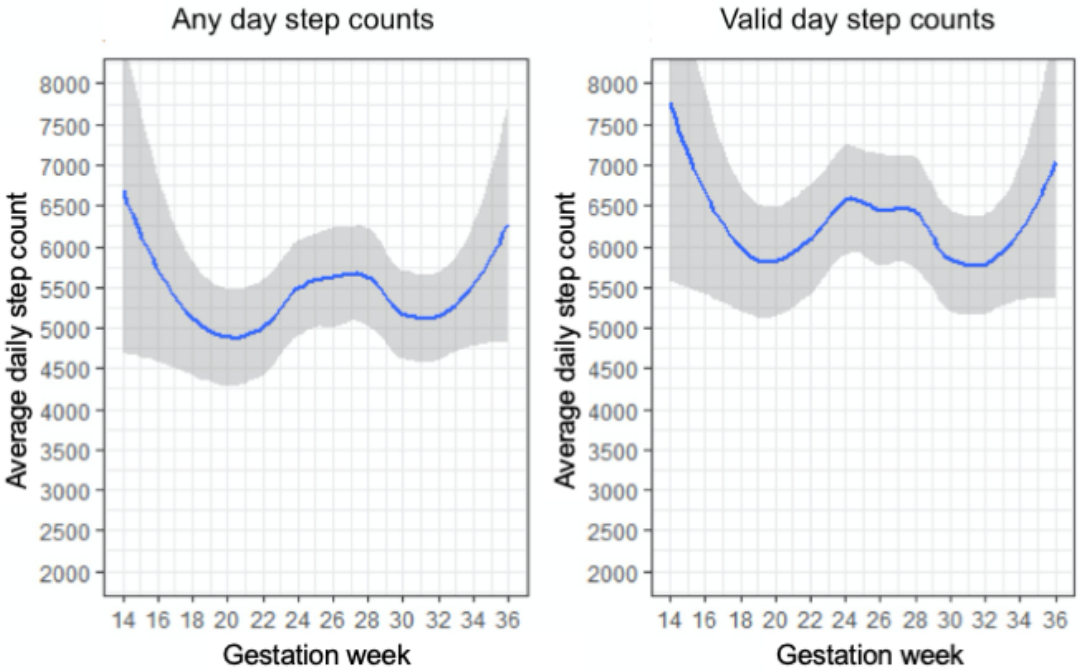
Figure 3. Average daily steps by week in the intervention.



On average, women entered the study at 21 weeks of pregnancy (mean 20.8, median 21, min/max 11/27). Figure 4 displays the average daily step counts by gestational week for any day step counts and full valid day step counts. All wear and valid full-day step counts were highest for women earliest in their pregnancies (14 weeks: mean 6208.1, SD 1324.1 for any wear; mean 7757.5, SD 2009.7 for valid days) and then dropped significantly by 19 weeks (any wear: mean 4973.4, SD 2543.8; valid days: 4334.7, SD 1991.6). Steps increased again toward the beginning of the third trimester (27 weeks, any wear: 6074.6 SD 1968.1; valid days: mean 6629.7, SD 2163.6) and then fell again as the third trimester progressed. Average step counts increased again between 33 and 36 weeks, but notably, only 2 women provided data beyond 33 weeks, contributing to more error in these estimates (Figure 4).

XSL•FO
RenderX

Figure 4. Average daily steps by gestation week.



Discussion

Principal Findings

These results suggest good feasibility for a counseling- and wearable tracker-based physical activity intervention for pregnant women with diabetes. Of all eligible participants, most (29/42, 69%) expressed interest and enthusiasm for the study, and nearly half were enrolled (19/42, 45%). Retention and adherence were high, with a median wear time of 90% of days, and 76% (32/42) of participants completing a follow-up visit. The intervention also appeared highly acceptable to participants, with a large majority (39/42, 93%) saying that they were satisfied; 93% (39/42) saying that they would recommend the program, and 86% (36/42) indicating that they felt motivated to exercise more.

Changes in daily steps were modest and generally increased at the very beginning and then fell again, then rose again after the 6-week call, before falling once more at the end of the intervention; therefore by week 10, the mean daily steps were quite similar to what they were at baseline. Without a control group, it is difficult to know whether the intervention was successful in increasing the number of typical steps for this population. These results are similar to a recent study using Fitbits to increase physical activity in healthy pregnant women, which found small increases in steps in both the intervention and control groups (both received Fitbits) across 12 weeks [23].

Previous research has shown that activity tends to decrease during pregnancy, particularly from the second to the third trimester [13,14]. Schmidt et al [14] showed that although activity remained relatively consistent across the first and second trimesters, there were declines in both moderate activity and total energy expenditure between the second and third trimesters. A study by Pereira et al [15] showed that the decrease in physical activity during pregnancy was especially pronounced in insufficiently active women, as was also with the women in our study. Considering these trends, it is promising that women in this study kept relatively stable steps during the initial 8 to 10 weeks of the trial rather than showing a steady decrease. Data on changes in activity during pregnancy for women with diabetes, however, are scarce, and it is difficult to know how activity in this sample is compared with the larger population.

The greatest barrier to recruitment, apart from an ongoing trial in the same clinic with similar inclusion criteria, was enrollment under 28 weeks of gestation. As DAPP is a tertiary referral clinic, many women were referred late in gestation and did not have a regular primary care physician. Additionally, most women with GD did not receive a diagnosis until at least weeks 24-26 of gestation and thus were ineligible for the study by the time they were screened. Increasing physical activity at any stage of pregnancy is likely to confer benefits to both mother and child; however, earlier intervention may have the greatest clinical impact. This can only be made possible by larger, systems-level changes through greater access to prenatal care and/or universal promotion of physical activity programs to all pregnant women. In addition, despite the low burden of the intervention, of the 42 eligible patients, 23 did not enroll, mostly citing the competing demands of managing diabetes during pregnancy. This highlights the difficulty of recruiting this population and the need to reframe physical activity not as a competing demand but as a key component of diabetes management.

Conversely, retention was relatively high for a high-risk clinical population, particularly as some participated in their third trimester. A meta-analysis of physical activity interventions in pregnancy and risk of GD found that loss to follow up is a major barrier in these populations, reaching up to 33% attrition in some studies [9]. High retention in this intervention was likely

XSL•FO
RenderX

owing to recruiting and delivering the intervention in a clinical setting where patients were expected to see their physician regularly and partnering with the attending physician to recruit and deliver the intervention. That said, following the initial goal-setting session, the entirety of the intervention was delivered remotely. Just 2 women stopped early because of medical complications/early labor, although these patients still provided consumer satisfaction data following delivery. Similarly, adherence was quite high, with participants wearing the Fitbit on approximately 90% of days they were in the intervention, although this decreased across the study. Russo et al [9] also found that adherence to physical activity interventions in pregnancy was highly variable and as low as 16%.

The high retention and adherence may have been owing to the high acceptability and light touch of the intervention. The most common theme in the follow-up interviews was an appreciation for the intervention not being invasive and that participants were allowed to go about their lives without feeling pushed or judged. This was facilitated by utilizing a face-to-face session to teach key behavior change strategies and then using the wearable tracker and app to reinforce these strategies throughout the intervention rather than relying on continued staff contact.

Although this approach was highly acceptable to participants, it is only ideal if it is also efficacious. Despite the fact that some more highly burdensome or *invasive* interventions may be less acceptable, it may also be the burden or the intrusion that leads to behavior change, and although only 1 participant specifically requested more staff contact, several noted a need for additional support during the third trimester. Future iterations of the intervention may need to incorporate greater intervention doses, particularly in the later stages of pregnancy, while still remaining noninvasive, particularly as this patient population is already burdened with tasks of diabetes management and prenatal care. This could be through greater staff contact, or still relying on technology to reinforce behavioral strategies but expanding to other channels (eg, texting) and with greater frequency. The need for this is reinforced by the fact that activity in this trial rose adjacent to contact with the study staff (baseline goal setting and 6-weeks check-in call).

Strengths and Limitations

This study has several strengths, including the patient population recruited. Pregnant women with diabetes are a high-risk clinical population with a great need for lifestyle intervention. A common limitation in physical activity interventions in pregnancy is adherence and retention, but both of these factors remained relatively high through the progression of pregnancy. In addition, participants were racially and ethnically diverse, with Latinas comprising the majority. Several spoke only Spanish, and most (all but 2) relied on public insurance, representing a particularly high-risk, underserved population. In addition, all study visits were completed in exam rooms in the clinic, laying a foundation for an intervention with high potential for clinical implementation.

There are also important limitations to note. As noted previously, this was a single-arm trial, which limited our ability to pinpoint the effects of the intervention or to examine individual intervention components. Although diverse, the study population was difficult to recruit, and the sample size was small. Although parts of the intervention piggy-backed onto existing clinical protocols, such as including their daily steps in weekly glucose reports to their clinicians, clinicians reported a lack of resources for interpreting and responding to these reports, highlighting the need for multilevel interventions to successfully integrate interventions into clinical settings. In addition, because the use of accelerometers was discontinued, more detailed data on time in various activity intensities, including sedentary time, were not available. Despite these limitations, given the lack of research on this population, we feel that these findings lay a valuable foundation for the much-needed interventions for this high-risk group.

Conclusions

Overall, these findings suggest that combined coaching- and mobile health–based physical activity intervention for pregnant women with diabetes is feasible and acceptable. A larger trial is needed to further explore the effect of wearable technology in physical activity interventions targeting pregnant women with diabetes. Future iterations of the intervention should include more staff contact and social support for the participants, especially during the later stages of pregnancy. Given the rapidly growing rates of diabetes in women of reproductive age, further development of such low-cost, low-burden interventions with potential for dissemination is vital to reducing complications and costs of diabetes in this high-risk population.

Acknowledgments

This research was supported by a Health Sciences Research Grant from the UCSD. BL was also supported by grant #K01DK101650. The authors would like to thank Erin Delker for her help with creating the figures and editing the paper. The authors would also like to thank our participants for their time and invaluable feedback.

Conflicts of Interest

None declared.

References

1. Bellamy L, Casas J, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. Lancet 2009 May 23;373(9677):1773-1779. [doi: 10.1016/S0140-6736(09)60731-5] [Medline: 19465232]

XSL•FO
RenderX

2. Dabelea D, Hanson RL, Lindsay RS, Pettitt DJ, Imperatore G, Gabir MM, et al. Intrauterine exposure to diabetes conveys risks for type 2 diabetes and obesity: a study of discordant sibships. Diabetes 2000 Dec;49(12):2208-2211. [doi: 10.2337/diabetes.49.12.2208] [Medline: 11118027]
3. Dabelea D, Mayer-Davis EJ, Lamichhane AP, D'Agostino RB, Liese AD, Vehik KS, et al. Association of intrauterine exposure to maternal diabetes and obesity with type 2 diabetes in youth: the SEARCH case-control study. Diabetes Care 2008 Jul;31(7):1422-1426 [FREE Full text] [doi: 10.2337/dc07-2417] [Medline: 18375420]
4. Clausen TD, Mathiesen E, Ekbom P, Hellmuth E, Mandrup-Poulsen T, Damm P. Poor pregnancy outcome in women with type 2 diabetes. Diabetes Care 2005 Feb;28(2):323-328. [doi: 10.2337/diacare.28.2.323] [Medline: 15677787]
5. Geiss LS, Wang J, Cheng YJ, Thompson TJ, Barker L, Li Y, et al. Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012. J Am Med Assoc 2014 Sep 24;312(12):1218-1226. [doi: 10.1001/jama.2014.11494] [Medline: 25247518]
6. Boyle J, Thompson T, Gregg E, Barker L, Williamson D. Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. Popul Health Metr 2010 Oct 22;8:29 [FREE Full text] [doi: 10.1186/1478-7954-8-29] [Medline: 20969750]
7. Imperatore G, Boyle JP, Thompson TJ, Case D, Dabelea D, Hamman RF, SEARCH for Diabetes in Youth Study Group. Diabetes Care 2012 Dec;35(12):2515-2520 [FREE Full text] [doi: 10.2337/dc12-0669] [Medline: 23173134]
8. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 7;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]
9. Russo LM, Nobles C, Ertel KA, Chasan-Taber L, Whitcomb BW. Physical activity interventions in pregnancy and risk of gestational diabetes mellitus: a systematic review and meta-analysis. Obstet Gynecol 2015 Mar;125(3):576-582. [doi: 10.1097/AOG.0000000000000691] [Medline: 25730218]
10. Bung P, Artal R, Khodiguian N, Kjos S. Exercise in gestational diabetes. An optional therapeutic approach? Diabetes 1991 Dec;40(Suppl 2):182-185. [doi: 10.2337/diab.40.2.s182] [Medline: 1748256]
11. Barakat R, Cordero Y, Coteron J, Luaces M, Montejo R. Exercise during pregnancy improves maternal glucose screen at 24-28 weeks: a randomised controlled trial. Br J Sports Med 2012 Jul;46(9):656-661. [doi: 10.1136/bjsports-2011-090009] [Medline: 21948120]
12. Artal R, O'Toole M. Guidelines of the American college of obstetricians and gynecologists for exercise during pregnancy and the postpartum period. Br J Sports Med 2003 Feb;37(1):6-12; discussion 12. [doi: 10.1136/bjsm.37.1.6] [Medline: 12547738]
13. Rousham EK, Clarke PE, Gross H. Significant changes in physical activity among pregnant women in the UK as assessed by accelerometry and self-reported activity. Eur J Clin Nutr 2006 Mar;60(3):393-400. [doi: 10.1038/sj.ejcn.1602329] [Medline: 16306930]
14. Schmidt MD, Pekow P, Freedson PS, Markenson G, Chasan-Taber L. Physical activity patterns during pregnancy in a diverse population of women. J Womens Health (Larchmt) 2006 Oct;15(8):909-918. [doi: 10.1089/jwh.2006.15.909] [Medline: 17087614]
15. Pereira MA, Rifas-Shiman SL, Kleinman KP, Rich-Edwards JW, Peterson KE, Gillman MW. Predictors of change in physical activity during and after pregnancy: Project Viva. Am J Prev Med 2007 Apr;32(4):312-319 [FREE Full text] [doi: 10.1016/j.amepre.2006.12.017] [Medline: 17383562]
16. Katz S, Feigenbaum A, Pasternak S, Vinker S. An interactive course to enhance self-efficacy of family practitioners to treat obesity. BMC Med Educ 2005 Jan 29;5(1):4 [FREE Full text] [doi: 10.1186/1472-6920-5-4] [Medline: 15679894]
17. Jansink R, Braspenning J, van der Weijden T, Elwyn G, Grol R. Primary care nurses struggle with lifestyle counseling in diabetes care: a qualitative analysis. BMC Fam Pract 2010 May 25;11:41 [FREE Full text] [doi: 10.1186/1471-2296-11-41] [Medline: 20500841]
18. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007 Apr 27;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
19. Prochaska J, Zabinski M, Calfas K, Sallis J, Patrick K. PACE+: interactive communication technology for behavior change in clinical settings. Am J Prev Med 2000 Aug;19(2):127-131. [doi: 10.1016/s0749-3797(00)00187-2] [Medline: 10913904]
20. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. JMIR mHealth uHealth 2016 Apr 27;4(2):e40. [doi: 10.2196/mhealth.4461]
21. Cadmus-Bertram L, Marcus BH, Patterson RE, Parker BA, Morey BL. Use of the FitBit to measure adherence to a physical activity intervention among overweight or obese, postmenopausal women: self-monitoring trajectory during 16 weeks. JMIR Mhealth Uhealth 2015 Nov 19;3(4):e96 [FREE Full text] [doi: 10.2196/mhealth.4229] [Medline: 26586418]
22. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a FitBit-based physical activity intervention for women. Am J Prev Med 2015 Sep;49(3):414-418 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]

XSL•FO
RenderX

23. Choi J, Lee JH, Vittinghoff E, Fukuoka Y. Mhealth physical activity intervention: a randomized pilot study in physically inactive pregnant women. Matern Child Health J 2016 May;20(5):1091-1101 [FREE Full text] [doi: 10.1007/s10995-015-1895-7] [Medline: 26649879]
24. Marcus BH, Dunsiger SI, Pekmezi DW, Larsen BA, Bock BC, Gans KM, et al. The Seamos Saludables study: a randomized controlled physical activity trial of Latinas. Am J Prev Med 2013 Nov;45(5):598-605 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.006] [Medline: 24139773]
25. Sisson SB, Camhi SM, Tudor-Locke C, Johnson WD, Katzmarzyk PT. Characteristics of step-defined physical activity categories in US adults. Am J Health Promot 2012;26(3):152-159. [doi: 10.4278/ajhp.100326-QUAN-95] [Medline: 22208412]
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 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]

Abbreviations

DAPP: Diabetes and Pregnancy Program
GD: gestational diabetes
HbA_{1c}: hemoglobin A_{1c}
T2D: type 2 diabetes
UCSD: University of California, San Diego

Edited by G Eysenbach; submitted 26.03.20; peer-reviewed by J Connelly, E Cotten, R Jiwani, B Balaji; comments to author 12.06.20; revised version received 06.08.20; accepted 13.09.20; published 21.10.20.

Please cite as:
Larsen B, Micucci S, Hartman SJ, Ramos G
Feasibility and Acceptability of a Counseling- and mHealth-Based Physical Activity Intervention for Pregnant Women With Diabetes: The Fit for Two Pilot Study
JMIR Mhealth Uhealth 2020;8(10):e18915
URL: http://mhealth.jmir.org/2020/10/e18915/
doi:10.2196/18915
PMID:33084584

©Britta Larsen, Stephanie Micucci, Sheri J Hartman, Gladys Ramos. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 21.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using mHealth to Provide Mobile App Users With Visualization of Health Checkup Data and Educational Videos on Lifestyle-Related Diseases: Methodological Framework for Content Development

Azusa Aida[1,2], MD; Thomas Svensson[1,3,4], MD, PhD; Akiko Kishi Svensson[1,2,3], MD, MPH, PhD; Hirokazu Urushiyama[5], MD, PhD; Kazuya Okushin[6], MD, PhD; Gaku Oguri[7], MD, PhD; Naoto Kubota[2], MD, PhD; Kazuhiko Koike[6], MD, PhD; Masaomi Nangaku[8], MD, PhD; Takashi Kadowaki[2,9], MD, PhD; Toshimasa Yamauchi[2], MD, PhD; Ung-Il Chung[1,4,10], MD, PhD

[1]Precision Health, Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan
[2]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[3]Department of Clinical Sciences, Lund University, Malmö, Sweden
[4]School of Health Innovation, Kanagawa University of Human Services, Tonomachi, Japan
[5]Department of Respiratory Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[6]Department of Gastroenterology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[7]Department of Cardiology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[8]Department of Nephrology and Endocrinology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[9]Toranomon Hospital, Tokyo, Japan
[10]Clinical Biotechnology, Center for Disease Biology and Integrative Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Akiko Kishi Svensson, MD, MPH, PhD
Precision Health
Department of Bioengineering, Graduate School of Engineering
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo, 113-8656
Japan
Phone: 81 3 5841 4737
Fax: 81 3 5841 4843
Email: akiko-kishi@umin.ac.jp

Abstract

Background: The number of people with lifestyle-related diseases continues to increase worldwide. Improving lifestyle behavior with health literacy may be the key to address lifestyle-related diseases. The delivery of educational videos using mobile health (mHealth) services can replace the conventional way of educating individuals, and visualization can replace the provision of health checkup data.

Objective: This paper aimed to describe the development of educational content for MIRAMED, a mobile app aimed at improving users' lifestyle behaviors and health literacy for lifestyle-related diseases.

Methods: All videos were based on a single unified framework to provide users with a consistent flow of information. The framework was later turned into a storyboard. The final video contents were created based on this storyboard and further discussions with leading experts and specialist physicians on effective communication with app users about lifestyle-related diseases.

Results: The app uses visualization of personal health checkup data and educational videos on lifestyle-related diseases based on the current health guidelines, scientific evidence, and expert opinions of leading specialist physicians in the respective fields. A total of 8 videos were created for specific lifestyle-related diseases affecting 8 organs: (1) brain–cerebrovascular disorder, (2) eyes–diabetic retinopathy, (3) lungs–chronic obstructive pulmonary disease, (4) heart–ischemic heart disease, (5) liver–fatty liver,

XSL•FO
RenderX

(6) kidneys–chronic kidney disease (diabetic kidney disease), (7) blood vessels–peripheral arterial disease, and (8) nerves–diabetic neuropathy.

Conclusions: Providing enhanced mHealth education using novel digital technologies to visualize conventional health checkup data and lifestyle-related diseases is an innovative strategy. Future studies to evaluate the efficacy of the developed content are planned.

(JMIR Mhealth Uhealth 2020;8(10):e20982) doi:10.2196/20982

KEYWORDS

apps; educational videos; health checkup; lifestyle-related disease; mHealth, prevention; telehealth; visualization

Introduction

Measures against noncommunicable diseases, such as cancer, heart disease, and stroke have been strengthened globally [1]. Above all, diabetes, evolving from metabolic syndrome, causes complications, which adversely impact a person's health and overall quality of life and add to the medical expenditures on both individual and societal levels [1]. In Japan, the number of individuals with lifestyle-related diseases (eg, cancer, heart disease, and cerebrovascular disease) has been increasing. Lifestyle-related diseases account for 60% of deaths in Japan, and national medical expenses continue to increase at the rate of 1 trillion yen per year [2-4].

In Japan, annual health checkups are stipulated by various laws of the Ministry of Health, Labour and Welfare, such as the Industrial Safety and Health Act and the Health Promotion Act [5]. Specific health checkups and specific health guidance started in 2008. The targets of these health checkups and health guidance are people aged 40-74 years. As per the Elderly Medical Care Act, insurers must provide annual health checkups to those who are insured [5]. The aim is to identify people with metabolic syndrome and a subsequent risk of developing lifestyle-related diseases, such as diabetes, hypertension, and hyperlipidemia due to visceral fat accumulation [6,7]. Unlike diseases treated only with medication, the key to managing metabolic syndrome and lifestyle-related diseases is to modify lifestyle behavior. Achieving such behavioral change is a challenge, as behavior in any target population differs by age, sex, occupation, lifestyle factors (eg, smoking, alcohol consumption, exercise, and diet), individual background, knowledge, and understanding of health issues that have accumulated over many years [7,8].

Although the impact of risky lifestyle behaviors on health has long been established, the importance of health literacy is now increasingly recognized [8,9]; and the association between health literacy and lifestyle behaviors has been widely confirmed [8,10]. Health literacy can be understood by using lay terms related to the ability of individuals to address health issues in a complex society. However, the rapid growth in its recognition has led to multiple interpretations of the concept, which may cause confusion [8]. Sørensen et al [11] defined health literacy as "people's knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course." Moreover, Berkman et al [12] defined health literacy by emphasizing similar abilities or "know-how" that could be used to "communicate about" issues related to health.

Improved health literacy about lifestyle-related diseases may thus lead to improved health behaviors, which in turn are important for sustainable prevention of lifestyle-related diseases. One route that could prove beneficial for improving health literacy is the use of mobile health (mHealth) apps.

In recent years, information and communications technologies (ICTs) have advanced rapidly, and the number of mHealth apps has been increasing [13]. mHealth is an expanding area within eHealth, which includes medical and public health information services provided via the internet and related technologies [14]. mHealth allows the general public to gain access to health advice or behavioral interventions. In Japan, the provision to use ICT for specific health guidance was initially announced in 2013 [15] and revised in 2018 [16]. Specific health guidance utilizing ICTs has the potential to increase participation rates due to its convenience to remote users as well as busy working professionals, allowing effective health guidance without in-person meetings.

Indeed, the delivery of educational content through videos can replace the current use of in-person communication in providing health checkup data and information about lifestyle-related diseases that may be difficult to understand for those who are not health care professionals [13,17]. With the expansion of mHealth services, the range of methods available to improve users' health literacy is increasing (eg, use of apps with videos, photos, and SMS text messages) [13,18]. To our knowledge, however, no app has yet been developed that encourages healthy behavior among persons with metabolic syndrome and high risk of lifestyle-related diseases, which (1) uses visualization of health checkup data to describe possible future lifestyle-related disease and (2) provides evidence-based educational content to improve health literacy about relevant lifestyle-related diseases.

MIRAMED, the lifestyle intervention app described in this paper, aims to be compatible with the Japanese Ministry of Health, Labour and Welfare's summary [19] of the present health promotion guidelines. In this paper, we describe the development of the visual content of the app that aims to improve user health literacy and lifestyle-related behaviors by converting the individual's numerical health checkup data into easy-to-understand visual information. We focus on the process of creating educational videos for this app based on the current guidelines and expert opinions from specialist physicians.

XSL•FO
RenderX

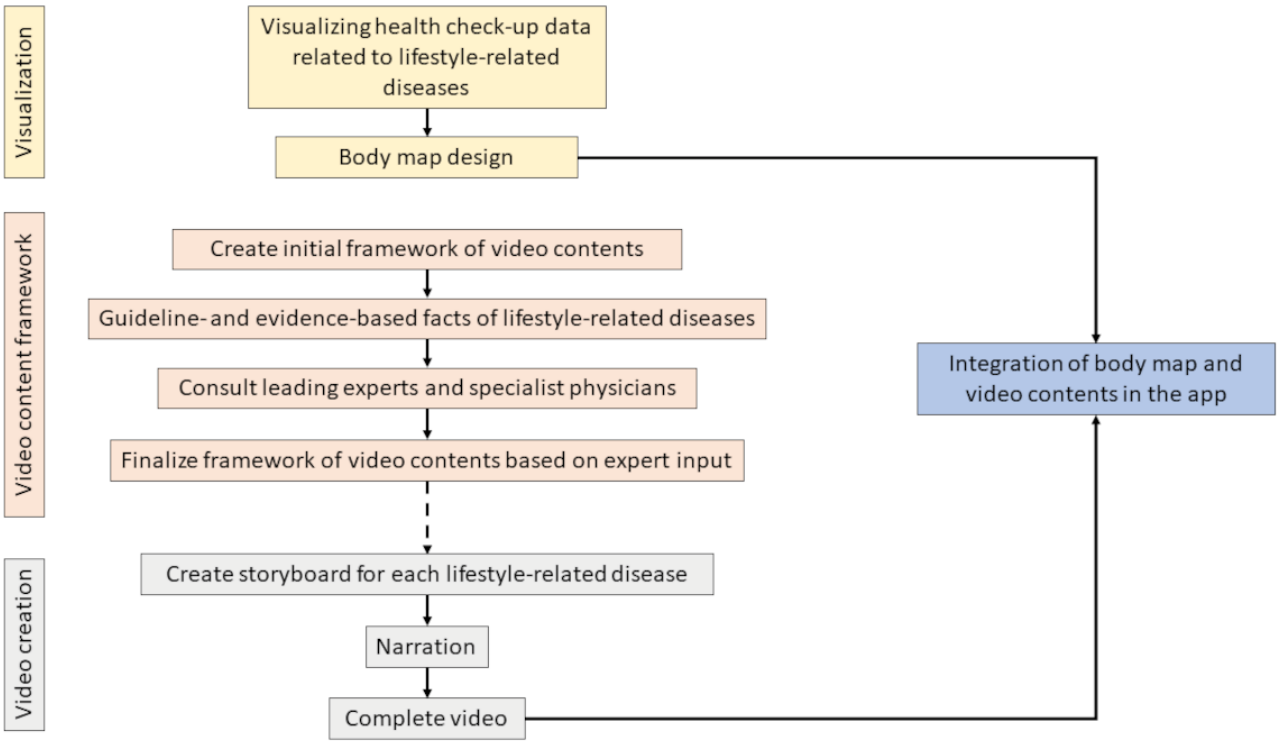
Methods

Development of Educational Contents

The target audiences of the app were people with metabolic syndrome and high risk of metabolic syndrome. The first step in the app's lifestyle intervention process was raising users' awareness of the current status of their health and potential risks of lifestyle-related diseases by visualizing their annual health checkup data. This step was facilitated by using icons displaying key organs in combination with educational videos that describe typical lifestyle-related diseases of those organs. Subsequently, users were prompted to set personal goals for lifestyle change in 5 key lifestyle categories: nutrition, smoking and alcohol, exercise, sleep, and stress. An intervention period of 3 months was set based on the user's understanding of their baseline health status. The user then received daily evaluations related to the lifestyle categories and weekly feedback on their progress. This paper focuses on the process of creating the visualization of health checkup data and the creation of the educational videos used in the MIRAMED app. The overview of this process is illustrated in Figure 1.

The MIRAMED app and its contents were developed by the Precision Health group, Center of Innovation at the University of Tokyo in Japan; and the study was approved by the ethics committee of the Department of Bioengineering at the University of Tokyo (approval number: KE18-44). This research was supported by the Center of Innovation Program from Japan Science and Technology Agency. The funding agency had no role in the design of the study; collection, analysis, and interpretation of data; writing of the report; and decision to submit the paper for publication.

Figure 1. Flowchart of the developmental process of the body map (visualization of health checkup data) and educational videos used in the lifestyle intervention app MIRAMED.

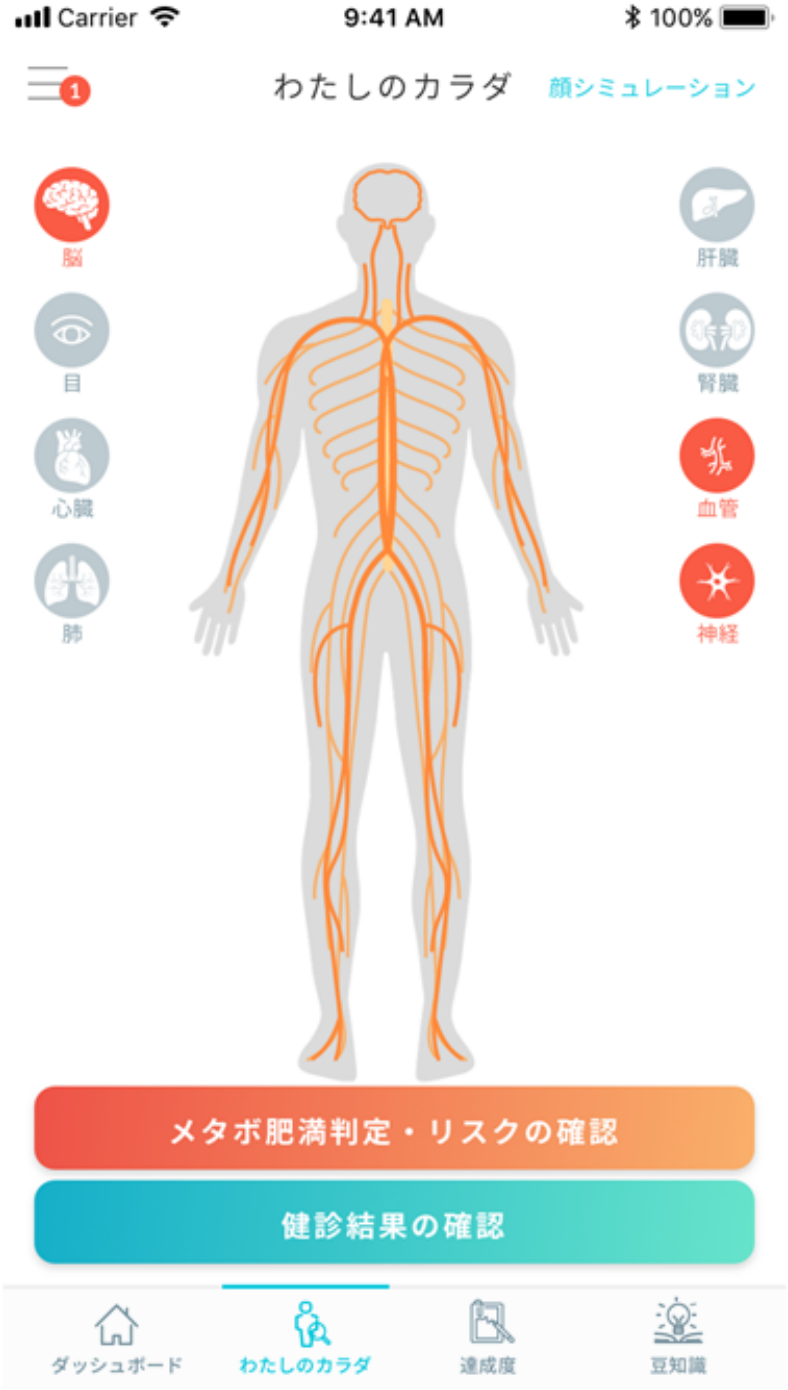


Visualization of Health Checkup Data

Annual health checkup data were visualized in a body map using icons of 8 organs related to lifestyle-related diseases: brain, eyes, lungs, heart, liver, kidneys, blood vessels, and nerves. When a user's health checkup data indicated possible risks associated with a potentially affected organ, the icons were rendered in orange coloring (Figure 2).

XSL•FO
RenderX

Figure 2. User interface of the app showing a body map with the visualization of a user's risks of lifestyle-related diseases.



Creation of Educational Videos About Lifestyle-Related Diseases

The purpose of educational videos in the app was to improve users' health literacy on potential health risks indicated by their health checkup data. In accordance with the Japanese Ministry of Health, Labour and Welfare's summary [19,20] of the present health promotion guidelines in Japan, 8 videos were created. Each video matched with 1 of the 8 selected organs and provided content on a lifestyle-related disease relevant to that organ. First, a unified framework was developed on which all the videos were based. The aim of this framework was to make users familiar with the flow of information in each video and to create cohesiveness across videos covering different lifestyle-related diseases. A framework was then developed into health condition–specific storyboards, which were in turn made into educational videos. Discussions were held with the leading experts and specialist physicians in the fields of neurology, diabetes and metabolic diseases, respiratory medicine, cardiology, gastroenterology, nephrology, and endocrinology. These experts and physicians consulted on how to convey evidence-based and health guidelines–compliant message for each disease to the app users. It was agreed that the unified framework for the videos would contain the following 5 sections (Table 1): disease name, explanation of the disease, symptoms, important facts, and improvement and prevention of the disease. It was required that the contents of each section were grounded in the scientific literature. Owing to the limitations on playback

XSL•FO
RenderX

time of each section, word count for the explanation of each section was also restricted. The framework for all the videos was shared with the production company, and visual designers were provided with instructions on the appropriate animated contents for each section. Extensive discussions were held regarding the comprehensibility of each section from a user's point of view. This feedback allowed for the creation of storyboards, which were discussed further before developing the final product.

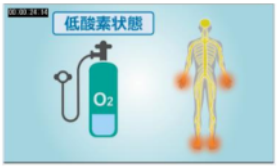
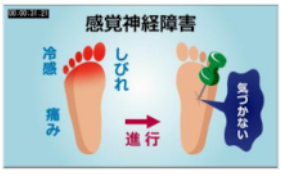
Table 1. The framework for "Diabetic neuropathy."

Sections	Explanation
Disease name	• Diabetic neuropathy
Explanation of the disease	• The systemic peripheral nervous system is damaged due to chronic hyperglycemia. Development of the condition involves peripheral nerve metabolism abnormality, decreased chronic blood flow in nerve tissue due to microangiopathy, and hypoxia. The pathogenic mechanism emphasizes polyol pathway enhancement, free radicals, abnormal lipid metabolism, protein glycation, and inflammatory factors [21]
Symptoms	• In a typical example, signs appear bilaterally in the toes and sole [21,22] • The signs expand to more proximal parts, such as ankles and lower legs. As the condition progresses, hands start showing signs in a glove-sock–like pattern [21,22] • Sensory nerves, motor nerves, and autonomic nerves are damaged, in that order [21-23] • Autonomic neuropathy can involve all systems in the body. It causes a substantial increase in morbidity and mortality, especially in the presence of cardiovascular autonomic neuropathy [21-24]
Important facts	• Diabetes causes various metabolic disorders centered on persistent hyperglycemia due to insufficient action of insulin. This impairment occurs in the order of neuropathy, retinopathy, and nephropathy [21,25]
Improvement and prevention	• Blood glucose control from an early stage [21,25-28]

Figure 3 gives an example of the storyboard, which was later turned into the educational video for "Diabetic neuropathy," a lifestyle-related nerves disease (as indicated by the corresponding icon). The storyboard additionally contained a short description of each section and the maximum length in minutes for the corresponding narration.

XSL•FO
RenderX

Figure 3. The storyboard of diabetic neuropathy.

Section (playback time in parentheses)	Animated content	Short description
a. Title (9 sec)		*"Common consequences likely to happen with your nerves"*, with the image of a nerve.
b. Cause of the disease (4 sec)		"*Diabetes mellitus* was illustrated by an image of a pinprick test to show the process of blood glucose measurement.
c. Explanation and pathological characteristic (13 sec)		Nerve hypoxia was expressed with the image of an oxygen tank in which the oxygen level decreased.
d. Symptoms (12 sec)		The left side of the image shows the starting location of the peripheral symptoms as illustrated with red toes. *"If you have abnormal sensation in the feet and the condition worsens, you may not even notice that you have stepped on a nail"*, the narration was emphasized by an image of a thumbtack stuck in the sole of the foot. (right side)
e. Important fact (5 sec)		Three major complications of diabetes were illustrated with the image of a podium which showed the order of the most common conditions, starting with neuropathy, followed by retinopathy and nephropathy.
f. Improvement and prevention (16 sec)		Healthy lifestyles were illustrated with avatars to encourage as well as visualize healthy lifestyle behaviors.

Results

A total of 8 videos were created by matching an organ with its associated lifestyle-related disease: (1) brain–cerebrovascular disorder, (2) eyes–diabetic retinopathy, (3) lungs–chronic obstructive pulmonary disease, (4) heart–ischemic heart disease, (5) liver–fatty liver, (6) kidneys–chronic kidney disease (diabetic kidney disease), (7) blood vessels–peripheral arterial disease, and (8) nerves–diabetic neuropathy. Each of the videos had a playback time of around 1 minute to retain the user's attention on the contents of the video.

In the example shown in Figure 2, the app indicates that the user had 3 potentially affected organs (brain, blood vessels, and nerves), which are rendered in orange coloring. Upon touching the icon of the glowing organ, the user would be taken to the educational video about the most common lifestyle-related diseases associated with that organ. A detailed description is provided for diabetic neuropathy.

XSL•FO
RenderX

The total playback time of the educational video for diabetic neuropathy was 59 seconds. The title was purposefully chosen as a statement, “Common consequences likely to happen with your nerves,” illustrated with the image of a nerve (Figure 3A). The title was seamlessly connected with a narrated explanation of the underlying cause of the condition: “The main cause of diabetic neuropathy is diabetes mellitus” [21,23]. This was shown along with an image of a pinprick test, meant to illustrate the process of blood glucose measurement (Figure 3B). As a further explanation of disease pathology, the narration continued, “If your blood glucose level remains high, the blood vessels that nourish the nerves will have an accumulation of waste products. This results in damage to the blood vessels, which leads to hypoxia and impaired nerve function.” The nerve hypoxia was expressed with the image of an oxygen tank in which the oxygen level decreased (Figure 3C). This was complemented with the figure of a human with highlighted hands and feet to suggest that symptoms usually start in the peripheral extremities. Typical symptoms were described in the order in which they usually appear, starting with sensory nerves, followed by motor nerves, and finally autonomic nerves [21-23]. The early-stage sensory disorder was explained as starting with spontaneous pain, numbness, and abnormal sensation in the lower limbs at an early stage of the onset [21,22,24,29]. The starting location of the peripheral symptoms was illustrated with red toes (left side of Figure 3D). Narration explained late-stage sensory disorder to the user. “If you have abnormal sensation in the feet and the condition worsens, you may not even notice that you have stepped on a nail.” The narration was combined with an image of a thumbtack stuck in the sole of foot (right side of Figure 3D).

The important fact section highlighted the following 3 major complications of diabetes: (1) neurological disorder–neuropathy, (2) eye disorder–retinopathy, and (3) kidney disorder–nephropathy [21,25]. The image of a podium (Figure 3E) illustrated the order of the most common conditions, starting with neuropathy, followed by retinopathy and nephropathy [21,25]. Finally, for the prevention or improvement of the condition, several factors were considered, such as proper diet (managing calory intake [21,25,28], starting each meal with vegetables [25], not eating sweets [25,28], eating breakfast every morning [25,28]); increased physical activity (combining aerobic and anaerobic exercises) [25,28]; and reducing stress in daily life [28]. In the video, healthy lifestyles were illustrated with avatars to encourage and visualize healthy lifestyle behaviors, such as waking up early in the morning, commuting on foot, bicycling, and allowing oneself time to relax (eg, by reading books) (Figure 3F).

Discussion

The purpose of the MIRAMED app is to improve users’ health literacy and encourage lifestyle-related behaviors through the personalization of health and lifestyle information. This occurs through the following means: (1) visualization of users’ annual health checkup data using icons that feature potential affected organs along with the possible lifestyle-related diseases of the featured organs, and (2) provide educational content on lifestyle-related diseases through videos. Given that informed, active, motivated, knowledgeable, and confident users can improve their health outcomes [30] and reduce medical costs [31], the app’s content may support a change of health behavior in users. To our knowledge, this is the first study to describe potential lifestyle-related diseases by visualizing users’ personal annual health checkup data in combination with educational videos to describe typical lifestyle-related diseases based on the scientific literature and opinions of leading experts and specialist physicians.

Improvement in self-efficacy significantly increases the likelihood that a health intervention (eg, weight loss [32,33] or smoking cessation [34]) will be successfully maintained. Learning and using medical terminology resembles learning a new language, and this is one reason why non–health care professionals face difficulty in understanding health checkup data and lifestyle-related diseases [35]. Visual and auditory augmentation of printed medical terms promotes the understanding of individual health status and provides an opportunity for repeated engagement in learning sessions [35].

App users can access educational videos about typical lifestyle-related diseases by touching icons of the respective organs. The purpose of these videos was to contribute to the improvement of the user’s health literacy. All videos were developed according to a consistent and unified framework. The framework was created in order to provide the user with a familiar structure and consistency across all videos, thus making the information from a wide range of medical specialties easily accessible. The playback time of each video was limited to around 1 minute, and the narration was conducted by the same narrator across all the videos and information. Moreover, both developers and specialists agreed that the contents of the app could be easily understood, even by those without a medical background.

The global penetration of mobile phones is growing rapidly, with up to 90% penetration in countries with high-income economies [36]. For those with a lifestyle-related disease or other chronic health problems, especially individuals living in rural areas who do not have access to medical services by other means, knowledge applicable to daily life is essential for effective self-management of lifestyle-related diseases [35].

This development had a few limitations worth noting. First, the content of the app was developed for the Japanese public. Thus, all the content was written in Japanese. However, it is possible to translate the app and its content into other languages. Second, information on how many times a user played each video and the number of videos watched by a user was not collected. Third, in general, there is a lack of standardized measurement tools to assess health literacy and conceptual framework for health literacy [8,37,38]. However, the future versions of the app can include measures of health literacy, for example, the Rapid Estimate of Adult Literacy in Medicine, Test of Functional Health Literacy in Adults (TOFHLA), or Newest Vital Sign [37-39]. We are presently investigating user experience with the usability of the app; accessibility of the contents; and changes in user lifestyle behaviors, weight, and waist circumference.

XSL•FO
RenderX

In conclusion, this study described the development of visually enhanced health education materials using new digital technologies with an innovative strategy to visualize conventional health checkup data. Although developed as educational videos for lifestyle-related diseases, these contents are expected to be used in various fields. Further evaluation of the effectiveness of the developed contents is needed.

Acknowledgments

We thank all staff members of the Center of Innovation at the University of Tokyo for their extensive efforts and help to conduct the study. We would also like to thank all members of Precision Health at the University of Tokyo for their invaluable assistance. This research was supported by the Center of Innovation Program of the Japan Science and Technology Agency (Grant Number JPMJCE1304).

Conflicts of Interest

AKS and TS have a financial interest in the MIRAMED app through a patent-licensing arrangement.

References

1. Nagai R. What should be health checkups and health guidance? In: Kadowaki T, Tsushita K, editors. The third period, Specific Health Checkups and Specific Health Guidance guide in Japanese. Tokyo: Nanzandou; Sep 20, 2018:4-11.
2. Sone H, Tanaka S, Iimuro S, Oida K, Yamasaki Y, Oikawa S, et al. Components of metabolic syndrome and their combinations as predictors of cardiovascular disease in Japanese patients with type 2 diabetes. Implications for improved definition. Analysis from Japan Diabetes Complications Study (JDCS). J Atheroscler Thromb 2009 Aug;16(4):380-387 [FREE Full text] [doi: 10.5551/jat.no117] [Medline: 19672033]
3. Ministry of Health, Labor and Welfare. Number of deaths, mortality rate (per 100,000 population), composition ratio by cause of death ranking (up to tenth place) by gender. Demographic statistics in Japan. 2018. URL: https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/kakutei18/dl/10_h6.pdf [accessed 2020-04-21]
4. Ministry of Health, Labor and Welfare. Overview of national medical expenses in Japan. 2017. URL: https://www.mhlw.go.jp/toukei/saikin/hw/k-iryohi/17/dl/kekka.pdf [accessed 2020-04-21]
5. Ministry of Health, Labor and Welfare. About medical insurance in Japan. 2019. URL: https://www.mhlw.go.jp/content/12400000/000591715.pdf [accessed 2020-04-21]
6. Hiuge-Shimizu A, Kishida K, Funahashi T, Ishizaka Y, Oka R, Okada M, et al. Absolute value of visceral fat area measured on computed tomography scans and obesity-related cardiovascular risk factors in large-scale Japanese general population (the VACATION-J study). Ann Med 2012 Feb;44(1):82-92. [doi: 10.3109/07853890.2010.526138] [Medline: 20964583]
7. Ministry of Health, Labor and Welfare. Standard health checkups and health guidance program. 2018. URL: https://www.mhlw.go.jp/content/10900000/000496784.pdf [accessed 2020-05-03]
8. Davey J, Holden CA, Smith BJ. The correlates of chronic disease-related health literacy and its components among men: a systematic review. BMC Public Health 2015 Jun 26;15:589 [FREE Full text] [doi: 10.1186/s12889-015-1900-5] [Medline: 26112264]
9. Paasche-Orlow MK, Wilson EAH, McCormack L. The evolving field of health literacy research. J Health Commun 2010;15 Suppl 2:5-8. [doi: 10.1080/10810730.2010.499995] [Medline: 20845188]
10. Australian Institute of Health and Welfare. Risk factors contributing to chronic disease. Australia: Australian Institute of Health and Welfare; 2012.
11. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
12. Berkman ND, Davis TC, McCormack L. Health literacy: what is it? J Health Commun 2010;15 Suppl 2:9-19. [doi: 10.1080/10810730.2010.499985] [Medline: 20845189]
13. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
14. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
15. Ministry of Health, Labour and Welfare. Guidance for implementation of specific health guidance utilizing ICT. 2013. URL: https://www.mhlw.go.jp/bunya/shakaihosho/iryouseido01/dl/info03j-130822_04.pdf [accessed 2020-04-29]
16. Ministry of Health, Labour and Welfare. About the first interview for the specific health guidance utilizing information and communication technology. 2018. URL: https://www.hospital.or.jp/pdf/14_20180209_01.pdf [accessed 2020-04-29]
17. Wanderer JP, Nelson SE, Ehrenfeld JM, Monahan S, Park S. Clinical Data Visualization: The Current State and Future Needs. J Med Syst 2016 Dec;40(12):275. [doi: 10.1007/s10916-016-0643-x] [Medline: 27787779]

XSL•FO
RenderX

18. Stuckey M, Fulkerson R, Read E, Russell-Minda E, Munoz C, Kleinstiver P, et al. Remote monitoring technologies for the prevention of metabolic syndrome: the Diabetes and Technology for Increased Activity (DaTA) study. J Diabetes Sci Technol 2011 Jul 01;5(4):936-944 [FREE Full text] [doi: 10.1177/193229681100500417] [Medline: 21880237]
19. Ministry of Health, Labour and Welfare. About health promotion method. 2003. URL: http://www.kenkounippon21.gr.jp/kenkounippon21/law/index_1.html [accessed 2020-04-24]
20. Kawakubo K. Adult disease called lifestyle-related disease. Kousei 1990;45(1):17-20.
21. The Japanese Diabetes Society. Complications. In: Diabetes specialist training guidebook. Tokyo: Shindan to Chiryo Sha Inc; 2017:279-315.
22. Tesfaye S, Boulton AJM, Dyck PJ, Freeman R, Horowitz M, Kempler P, Toronto Diabetic Neuropathy Expert Group. Diabetic neuropathies: update on definitions, diagnostic criteria, estimation of severity, and treatments. Diabetes Care 2010 Oct;33(10):2285-2293 [FREE Full text] [doi: 10.2337/dc10-1303] [Medline: 20876709]
23. Freeman R. Diabetic autonomic neuropathy. Handb Clin Neurol 2014;126:63-79. [doi: 10.1016/B978-0-444-53480-4.00006-0] [Medline: 25410215]
24. Boulton AJM, Vinik AI, Arezzo JC, Bril V, Feldman EL, Freeman R, American Diabetes Association. Diabetic neuropathies: a statement by the American Diabetes Association. Diabetes Care 2005 Apr;28(4):956-962. [doi: 10.2337/diacare.28.4.956] [Medline: 15793206]
25. The Japanese Diabetes Society. General treatment. In: Diabetes specialist training guidebook. Tokyo: Shindan to Chiryo Sha Inc; 2017:148-164.
26. UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. UK Prospective Diabetes Study Group. BMJ 1998 Sep 12;317(7160):703-713 [FREE Full text] [Medline: 9732337]
27. Ohkubo Y, Kishikawa H, Araki E, Miyata T, Isami S, Motoyoshi S, et al. Intensive insulin therapy prevents the progression of diabetic microvascular complications in Japanese patients with non-insulin-dependent diabetes mellitus: a randomized prospective 6-year study. Diabetes Res Clin Pract 1995 May;28(2):103-117. [doi: 10.1016/0168-8227(95)01064-k] [Medline: 7587918]
28. Belloc NB, Breslow L. Relationship of physical health status and health practices. Prev Med 1972 Aug;1(3):409-421. [doi: 10.1016/0091-7435(72)90014-x] [Medline: 5085007]
29. Himeno T, Kamiya H, Nakamura J. Lumos for the long trail: Strategies for clinical diagnosis and severity staging for diabetic polyneuropathy and future directions. J Diabetes Investig 2020 Jan;11(1):5-16 [FREE Full text] [doi: 10.1111/jdi.13173] [Medline: 31677343]
30. Rusin M, Arsand E, Hartvigsen G. Functionalities and input methods for recording food intake: a systematic review. Int J Med Inform 2013 Aug;82(8):653-664. [doi: 10.1016/j.ijmedinf.2013.01.007] [Medline: 23415822]
31. Ryan DH. The pharmacological and surgical management of adults with obesity. J Fam Pract 2014 Jul;63(7 Suppl):S21-S26. [Medline: 25198216]
32. Kim H, Faw M, Michaelides A. Mobile But Connected: Harnessing the Power of Self-Efficacy and Group Support for Weight Loss Success through mHealth Intervention. J Health Commun 2017 May;22(5):395-402. [doi: 10.1080/10810730.2017.1296510] [Medline: 28339324]
33. Nezami BT, Lang W, Jakicic JM, Davis KK, Polzien K, Rickman AD, et al. The Effect of Self-Efficacy on Behavior and Weight in a Behavioral Weight-Loss Intervention. Health Psychol 2016 May 16 [FREE Full text] [doi: 10.1037/hea0000378] [Medline: 27183306]
34. Blevins CE, Farris SG, Brown RA, Strong DR, Abrantes AM. The Role of Self-Efficacy, Adaptive Coping, and Smoking Urges in Long-Term Cessation Outcomes. Addict Disord Their Treat 2016 Dec;15(4):183-189 [FREE Full text] [doi: 10.1097/ADT.0000000000000087] [Medline: 28133440]
35. Wood FG, Alley E, Baer S, Johnson R. Interactive Multimedia Tailored to Improve Diabetes Self-Management. Nurs Clin North Am 2015 Sep;50(3):565-576. [doi: 10.1016/j.cnur.2015.05.009] [Medline: 26333610]
36. Frøisland DH, Arsand E, Skårderud F. Improving diabetes care for young people with type 1 diabetes through visual learning on mobile phones: mixed-methods study. J Med Internet Res 2012 Aug 06;14(4):e111 [FREE Full text] [doi: 10.2196/jmir.2155] [Medline: 22868871]
37. Ownby RL, Waldrop-Valverde D, Hardigan P, Caballero J, Jacobs R, Acevedo A. Development and validation of a brief computer-administered HIV-Related Health Literacy Scale (HIV-HL). AIDS Behav 2013 Feb;17(2):710-718 [FREE Full text] [doi: 10.1007/s10461-012-0301-3] [Medline: 22961499]
38. Baker DW. The meaning and the measure of health literacy. J Gen Intern Med 2006 Aug;21(8):878-883. [doi: 10.1111/j.1525-1497.2006.00540.x] [Medline: 16881951]
39. Ownby RL, Acevedo A, Waldrop-Valverde D, Jacobs RJ, Caballero J, Davenport R, et al. Development and initial validation of a computer-administered health literacy assessment in Spanish and English: FLIGHT/VIDAS. Patient Relat Outcome Meas 2013;4:21-35 [FREE Full text] [doi: 10.2147/PROM.S48384] [Medline: 23990736]

XSL•FO
RenderX

Abbreviations

ICT: information and communications technology
mHealth: mobile health
TOFHLA: Test of Functional Health Literacy in Adults

Edited by G Eysenbach; submitted 08.06.20; peer-reviewed by B Oh; comments to author 29.06.20; revised version received 22.08.20; accepted 14.09.20; published 21.10.20.

Please cite as:
Aida A, Svensson T, Svensson AK, Urushiyama H, Okushin K, Oguri G, Kubota N, Koike K, Nangaku M, Kadowaki T, Yamauchi T, Chung UI
Using mHealth to Provide Mobile App Users With Visualization of Health Checkup Data and Educational Videos on Lifestyle-Related Diseases: Methodological Framework for Content Development
JMIR Mhealth Uhealth 2020;8(10):e20982
URL: http://mhealth.jmir.org/2020/10/e20982/
doi:10.2196/20982
PMID:33084586

©Azusa Aida, Thomas Svensson, Akiko Kishi Svensson, Hirokazu Urushiyama, Kazuya Okushin, Gaku Oguri, Naoto Kubota, Kazuhiko Koike, Masaomi Nangaku, Takashi Kadowaki, Toshimasa Yamauchi, Ung-Il Chung. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 21.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mobile Social Network–Based Smoking Cessation Intervention for Chinese Male Smokers: Pilot Randomized Controlled Trial

Jinsong Chen[1], MPH; Elsie Ho[2], PhD; Yannan Jiang[1], PhD; Robyn Whittaker[1], PhD; Tingzhong Yang[3], PhD; Christopher Bullen[1], PhD

[1]National Institute for Health Innovation, School of Population Health, University of Auckland, Auckland, New Zealand

[2]School of Population Health, University of Auckland, Auckland, New Zealand

[3]Centre for Tobacco Control Research, Zhejiang University, Hangzhou, China

Corresponding Author:
Jinsong Chen, MPH
National Institute for Health Innovation
School of Population Health
University of Auckland
22-30 Park Ave, Grafton
Auckland, 1023
New Zealand
Phone: 64 0220951065
Email: jinsong.chen@auckland.ac.nz

Abstract

Background: Around 2 million Chinese people, mostly men, die annually from tobacco-related diseases; yet, fewer than 8% of Chinese smokers ever receive any smoking cessation support.

Objective: This study aimed to test the preliminary effectiveness and feasibility for a mobile social network (WeChat)–based smoking cessation intervention (SCAMPI program) among Chinese male smokers.

Methods: Chinese male smokers aged 25-44 years were recruited online from WeChat, the most widely used social media platform in China. Individuals using other smoking cessation interventions or who lacked capacity to provide online informed consent were excluded. Participants were randomly assigned (1:1) to intervention or control groups. Neither participants nor researchers were masked to assignment. The trial was fully online. All data were collected via WeChat. The intervention group received access to the full-version SCAMPI program, a Chinese-language smoking cessation program based on the Behaviour Change Wheel framework and relevant cessation guidelines. Specific intervention functions used in the program include: planning to help users make quitting plans, calculator to record quitting benefits, calendar to record progress, gamification to facilitate quitting, information about smoking harms, motivational messages to help users overcome urges, standardized tests for users to assess their levels of nicotine dependence and lung health, as well as a social platform to encourage social support between users. The control group had access to a static WeChat page of contacts for standard smoking cessation care. Both groups received incentive credit payments for participating. The primary outcome was 30-day biochemically verified smoking abstinence at 6 weeks after randomization, with missing data treated as not quitting. Secondary outcomes were other smoking status measures, reduction of cigarette consumption, study feasibility (recruitment and retention rate), and acceptability of and satisfaction with the program.

Results: The program recorded 5736 visitors over a 13-day recruitment period. We recruited 80 participants who were randomly allocated to two arms (n=40 per arm). At 6 weeks, 36 of 40 (90%) intervention participants and 35 of 40 (88%) control participants provided complete self-reported data on their daily smoking status via WeChat. Biochemically verified smoking abstinence at 6 weeks was determined for 10 of 40 (25%) intervention participants and 2 of 40 (5%) control participants (RR=5, 95% CI 1.2-21.4, *P*=.03). In the intervention group, the calculator function, motivational messages, and health tests were underused (less than once per week per users). Participants rated their satisfaction with the intervention program as 4.56 out of 5.00.

Conclusions: Our program is a novel, accessible, and acceptable smoking cessation intervention for Chinese male smokers. A future trial with a greater sample size and longer follow-up will identify if it is as effective as these preliminary data suggest.

Trial Registration: ANZCTR registry, ACTRN12618001089224; https://tinyurl.com/y536n7sx

International Registered Report Identifier (IRRID): RR2-18071

XSL•FO
RenderX

(JMIR Mhealth Uhealth 2020;8(10):e17522) doi:10.2196/17522

KEYWORDS

mHealth; mobile smoking cessation; social network-based intervention; smoking cessation; public health; gamified health interventions

Introduction

Problems With Smoking in China

Nearly 300 million people in China smoke tobacco daily [1], 95% of whom are men [1]. Smoking causes 25% of Chinese male deaths [2]. However, tobacco control measures are not well adhered to, and fewer than 8% of Chinese smokers have ever received any smoking cessation advice [3]. Interest in smoking cessation is modest (only 38.8% of all current smokers), and success is rare (10.7% among people who had ever tried to quit smoking) among Chinese male smokers [4].

Mobile Phone–Based Smoking Cessation Interventions

Mobile phone–based smoking cessation interventions have the potential to address this problem. A Cochrane review of SMS messaging trials for smoking cessation found them to be effective [5]. Apps may also help; in a trial of a behavioral, decision-aid smartphone app, 23.8% of participants self-reported 3-month continuous smoking abstinence compared with 10.2% of participants who did not receive access to the decision-aid app [6].

Smartphone ownership in China is dramatically increasing, especially in the past 5 years [7], to the point where China now has the largest smartphone market (>700 million people) in the world [8]. The largest user group is composed of men aged 25-44 years [9]. Smartphone-based smoking cessation interventions might be a solution to address the problems of low awareness and usage of smoking cessation services in China.

WeChat Usage

WeChat is the most popular social network platform in China. In 2020, 1.2 billion people were monthly active users of WeChat. WeChat users are highly engaged with the app; nearly 80% of WeChat users use the app for >30 minutes daily [10]. In 2016, WeChat launched a new feature, “mini-program,” to deliver messages, collect user data, and respond to user commands [11]. This makes the WeChat platform more powerful than ever before.

The WeChat super app is becoming a vital component of health care in China [12]. In the past 5 years, a number of studies have attempted to identify the effectiveness and feasibility of delivering health care interventions through WeChat, such as self-management of hypertension or type 2 diabetes [13-16]. Some studies have found improvements in users’ health conditions [15-17]. However, WeChat’s application for smoking cessation remains untapped. Given the large number of WeChat users and the high prevalence of smoking in China, WeChat should be tested to see if it is an appropriate and acceptable platform to deliver smoking cessation interventions.

Aim

We aimed to evaluate the preliminary effectiveness and feasibility of SCAMPI, a standalone online smoking cessation intervention designed, developed, and trialed entirely within the WeChat ecosystem to reach Chinese men aged 25-44 years who smoke tobacco.

Methods

Study Design

We used a two-arm, parallel, randomized controlled trial design to evaluate preliminary effectiveness and acceptability of the SCAMPI program as well as recruitment and retention for estimating the sample size of a future definitive trial. All trial procedures were conducted online via WeChat. Ethical approvals were granted by The University of Auckland Human Participants Ethics Committee (Ref. 021649) and Zhejiang University School of Public Health Research Ethics Committee (Ref. ZGL201801-2). The trial was prospectively registered with the ANZCTR registry (ACTRN12618001089224).

Participants

Participants were recruited from China via WeChat advertisements. People interested in participating were invited to complete an electronic questionnaire on WeChat. Inclusion criteria included smokers (daily smokers “smoking any type of tobacco products on a daily basis” or occasional smokers “smoking any type of tobacco products occasionally”) aged 25-44 years; access to a smartphone; knowledge of the Chinese language and WeChat platform; and willingness to participate and provide follow-up information at scheduled points. Participants were excluded if they used other types of smoking cessation interventions, refused to provide follow-up information, or had a medical condition (eg, mental health conditions or cognitive issues that prohibit understanding of the information provided by the program) that could limit their ability to participate. The first 80 eligible individuals who completed questionnaires and subscribed to the SCAMPI Official Account (OA), a social network account used for messaging and interacting with other subscribers, were consented, enrolled in the trial, and randomized. We collected consent via WeChat; then, we asked participants to provide data on demographics, smoking status, acceptability, and satisfaction (intervention group only) with the program via online questionnaires. Participants’ use patterns and engagement with the program were collected via system records of interaction.

Randomization and Masking

Participants were randomly allocated into intervention and control groups at a 1:1 ratio using a computer-generated randomization sequence with variable block sizes of 2 or 4, provided by the trial statistician (YJ). Participants enrolled themselves to the trial by providing informed consent,

XSL•FO
RenderX

completing the registration questionnaire, and subscribing to the SCAMPI OA. The investigator (JC) assigned participants to the intervention based on the randomization sequence.

Due to the nature of the intervention, participants were not masked to the intervention, as they were provided information about the interventions they would have access to. Participants were not informed which version of the program they used —whether it was the intervention of interest or the comparator. Versions 1 and 2 were used to represent the full and restricted versions of the SCAMPI program, respectively.

Researchers were not masked to the treatment allocation, and statistical analyses were not blinded to the treatment allocation. There are multiple sources used for this trial, and 2 of the sources are only used to collect data (SCAMPI mini-program usage data collected via the mini-program server and users' satisfaction data collected via online questionnaires) from participants in the intervention group. It is impossible to blind researchers to the treatment allocation in order to export data from these sources.

Recruitment

The advertisement and recruitment of the study were completely online through the WeChat platform. WeChat users who were interested in quitting smoking and typed keywords like "smoking," "stop smoking," and "smoking cessation support" would be able to find the SCAMPI OA from WeChat. WeChat users who subscribed to the SCAMPI OA were asked if they would like to take part in the trial through automated WeChat messages. The message was sent in Chinese, translated as "Would you like to be part of a trial for evaluating the preliminary efficacy and feasibility of a WeChat-based smoking cessation program? Use the following link to register as a participant of the trial!" The Tencent Games Dreaming Plan (a department of the Interactive Entertainment Group) also posted the QR code of the SCAMPI OA in their platform to promote the trial and program. After tapping the link, potential participants were directed to the registration system and asked to provide informed consent (see Multimedia Appendix 1). Participants were expected to provide their consent electronically by tapping the "agree" button on the web page. Participation was completely anonymous. Any data related to the participant's personal identity were not collected. A simple IP address check was used to ensure people did not use multiple identities (one participant participated the study with different WeChat accounts).

Procedures

After providing informed consent, participants were randomly allocated to different groups (intervention and control group, n=40 in each group). Participants randomized to the intervention group had access for 6 weeks to the full version of the SCAMPI program, and participants randomized to the control group had access for 6 weeks to a restricted version of the SCAMPI program. On completion of the 6-week follow-up assessment, participants in the control group were offered full access to the SCAMPI program. All data on program use were recorded until the end of the trial. All participation and procedures of the trial were completed online through the WeChat platform, except participants who reported 30-day continuous smoking abstinence at the 6-week follow up received a cotinine test kit via courier to test nicotine levels in their saliva. A link to a courier WeChat mini-program was sent to this group of participants. Participants who received the link were asked to provide their address and recipient details to the courier through WeChat (none of these data were accessible nor recorded by the research team). Based on the details, the kits were sent to the participants. Participants were requested to use the kit as instructed, take a photo or shoot a short video about how they used the kits and the relevant results (personal images were not requested), and send the test results to the SCAMPI OA. All data were collected online via the WeChat platform. Participants in both groups recorded their daily cigarette consumption for the week throughout the trial period. Usage data on the corresponding versions of the SCAMPI program were recorded on the server of the program. Participants' baseline data and satisfaction with the intervention were collected through electronic questionnaires (registration questionnaire and end-of-trial questionnaire). Compensation for participation was delivered in a form called WeChat red packet, a widely used online currency in China and transferrable through the WeChat platform. Recipients of WeChat red packet can use it to purchase goods or use services in China. WeChat red packet was attached to the registration questionnaire, 6-weekly smoking status check-in sessions, and end-of-trial questionnaire (for the intervention group, this is delivered along with the last weekly smoking status check-in session) in the trial. Once participants complete the questionnaires, the WeChat red packet appears in their account (participants in both groups receive same value red packets). Participants' answers to questionnaires had no impact on the WeChat red packet they received.

Interventions

Details of the SCAMPI program and its development have been published elsewhere [18]. In brief, the program was developed through a 1-month collaborative product development process involving 20 Chinese male smokers aged 25-44 years recruited through WeChat. We analyzed their smoking behavior using the "Capability," "Opportunity," "Motivation," "Behaviour" (COM-B) model [19], along with relevant reports of the 2015 Chinese Adult Tobacco Use Research [20] and the national online survey about smoking behaviors of Chinese smokers [21]. The collaborative product development process identified a number of behavioral factors affecting the smoking behavior of the target population, as shown in Textbox 1.

XSL•FO
RenderX

Textbox 1. Summary of behavioral factors for smoking.

Psychological capability

1. Lack of methods to cope with unwanted emotion
2. Low health literacy of smoking harms and second-hand smoking harms
3. Low health literacy of cessation benefits to self and family
4. Lack of methods to quit smoking
5. Lack of awareness of existing smoking cessation services

Reflective motivation

1. Problematic impression of smoking (eg, smoking is cool)

Automatic motivation

1. Perception of smoking as an emotional coping tool (ie, unwanted emotion triggers smoking behavior)
2. Perception of smoking as an entertainment tool (ie, entertainment, such as playing cards, triggers smoking behavior)

Physical opportunity

1. High probability of being in a triggering environment (eg, gifted with cigarettes, peer smoking)
2. Low cost of tobacco

Social opportunity

1. Using smoking as a social communication tool (eg, peer smoking and sharing cigarettes)

Using the Behaviour Change Wheel framework, interventions were designed to address the behavioral factors identified. These interventions adhered to the China Clinical Smoking Cessation Guideline, were coded as functions in the SCAMPI program [19,22], and were deliverable by the WeChat platform [23,24] (see Multimedia Appendix 2). Participants indicated their preferences for the functions in a WeChat-based smoking cessation intervention via an online questionnaire. Prototype iterations and questions were posted on SCAMPI WeChat OA to solicit participants' preferences on the user experience and user interfaces. Participants' comments on the prototype helped developers to iteratively update the SCAMPI program before its launch. At the end of the development phase, the 20 participants were invited to rate usability and acceptability by completing an online questionnaire based on the Mobile App Rating Scale questionnaire [25]. The online questionnaire was completed by 16 of the 20 participants, and the average rating for the SCAMPI program was 4.4 out of 5.0.

The program has the following functions: (1) smoking cessation planning, (2) calculator to record quitting benefits (eg, money saved by not smoking), (3) calendar to record progress of smoking cessation (eg, on which dates did the users not smoke), (4) gamification to facilitate quitting (eg, ranking board for users to compete for the longest continuous smoking abstinence), (5) information about smoking harms and health benefits from smoking cessation, (6) motivational messages (based on requests from users) to help users overcome urges, (7) standardized tests to help users assess their levels of nicotine dependence and lung health, and (8) social support platform to deliver peer support between users. Screenshots and QR codes of the SCAMPI program are provided in Multimedia Appendix 3. At the end of the trial, all WeChat users had free access to the SCAMPI program by either using the keywords "smoking/smoking cessation/SCAMPI" on the WeChat platform or scanning the SCAMPI program QR codes. The full version of the SCAMPI program has the aforementioned functions and content. The restricted version of the program provided the users with contact information for standard smoking cessation care (eg, Quitline in China, local smoking cessation clinics). The minimum frequency of program use is once a week, to provide participants' daily cigarette smoking status throughout the week. Prompts and messages (eg, motivational messages to support users to avoid smoking urges or prompt users to read program posts about smoking harms) were triggered by requests from users. The only messages users received were reminders to provide their smoking status, triggered in the 48 hours after the date users were requested to provide their daily smoking status for the week. The reminders were delivered <3 times a day (once every 8 hours) for 2 days via WeChat. Participants who did not respond to any of these messages were counted as dropouts from the trial. There was no specific training session provided to users. After successfully registering as participants to the study, participants received a message to introduce the versions of program they would use in the trial. Since the program was built natively on the WeChat platform using the WeChat mini-program design guideline, users were expected to have no challenges with using the functions of the program.

Outcomes

The primary outcome was biochemically verified past 30-day smoking abstinence at 6 weeks [25]. Quit failure was defined as any number of cigarettes smoked in the past 30 consecutive days. This measure is known as the 10th level of smoking abstinence in the Chinese Standard for Smoking Cessation [24]. Participants who reported themselves as 30-day smoking

XSL•FO
RenderX

abstainers at the 6-week follow-up received cotinine test strip kits by post. These participants were requested to complete the test at home (using their saliva, with clear instructions provided) and upload a photo or a short video of the strip test results to the investigators via WeChat.

Secondary outcomes were cigarette consumption, 7-day self-reported smoking abstinence at the 4-week and 6-week follow-ups, self-reported 30-day smoking abstinence at the 6-week follow-up (all measured by online, self-reported data), retention in the trial, number of times the participants interacted with their program throughout the trial, and satisfaction ratings for using SCAMPI as a smoking cessation tool (intervention group only).

We used 2 electronic questionnaires. The registration questionnaire asked participants about their demographics, smoking status, and willingness to quit smoking (see Multimedia Appendix 4). The registration questionnaire was developed based on the national online survey about smoking behaviors of Chinese smokers [20,21]. The end-of-trial questionnaire (see Multimedia Appendix 5) was developed based on a standard questionnaire, the Mobile App Rating Scale questionnaire [26]. It focuses on assessing users' perception of and satisfaction with the app, software, or mobile program they used. Both questionnaires were reviewed and commended by experts from WeChat to ensure its user experience met the Chinese users' preferences. Usage was measured by the number of times users visited the program. Participants were encouraged to make comments by directly sending text or audio messages to the program if they wished.

Guidelines for Participant Withdrawal

Participants were informed in the participant information sheet that they were able to withdraw from the trial at any time with no reasons needed. If a participant reported severe physical or mental discomfort by using the SCAMPI program, the investigator (JC) terminated his participation immediately and referred him to health care professionals. Events were recorded and reported.

Statistical Analysis

In this pilot trial, we recruited around 15% of the total sample required for a future adequately powered trial (ie, 80 participants, 40 per group). We estimate a full trial would need a total of 530 participants (n=265 per group) for 90% power at 5% significance (2-sided) to detect an absolute difference of 10% in the primary outcome between 2 groups, assuming a control quit rate of 6.7% [4]. The feasibility outcomes (eg, recruitment, retention, participants' demographics, acceptability) are reported descriptively. All randomized participants were included in the final analysis following the intent-to-treat principle. Missing smoking outcomes were considered as treatment failure and imputed as not quitting. Baseline demographics were summarized by randomized group using descriptive statistics. Primary and secondary outcomes are presented as mean (SD) for continuous variables and n (%) for categorical variables. To test the hypothesis of a difference between the 2 randomized groups after intervention, smoking outcomes were analyzed using a Chi-square test or the Fisher exact test when the counts were <5. Both relative risk (RR) and odds ratio (OR) were estimated with 95% CIs. For continuous outcomes, we used a 2-sample *t* test or Wilcoxon non-parametric test for skewed data. A generalized linear mixed model was used to analyze daily cigarette consumption reported by the participants over the 6-week trial period, using Poisson distribution with a log link. Statistical analysis was performed using SAS version 9.4 (SAS Institute Inc, Cary, NC). All statistical tests were 2-sided at a 5% significance level.

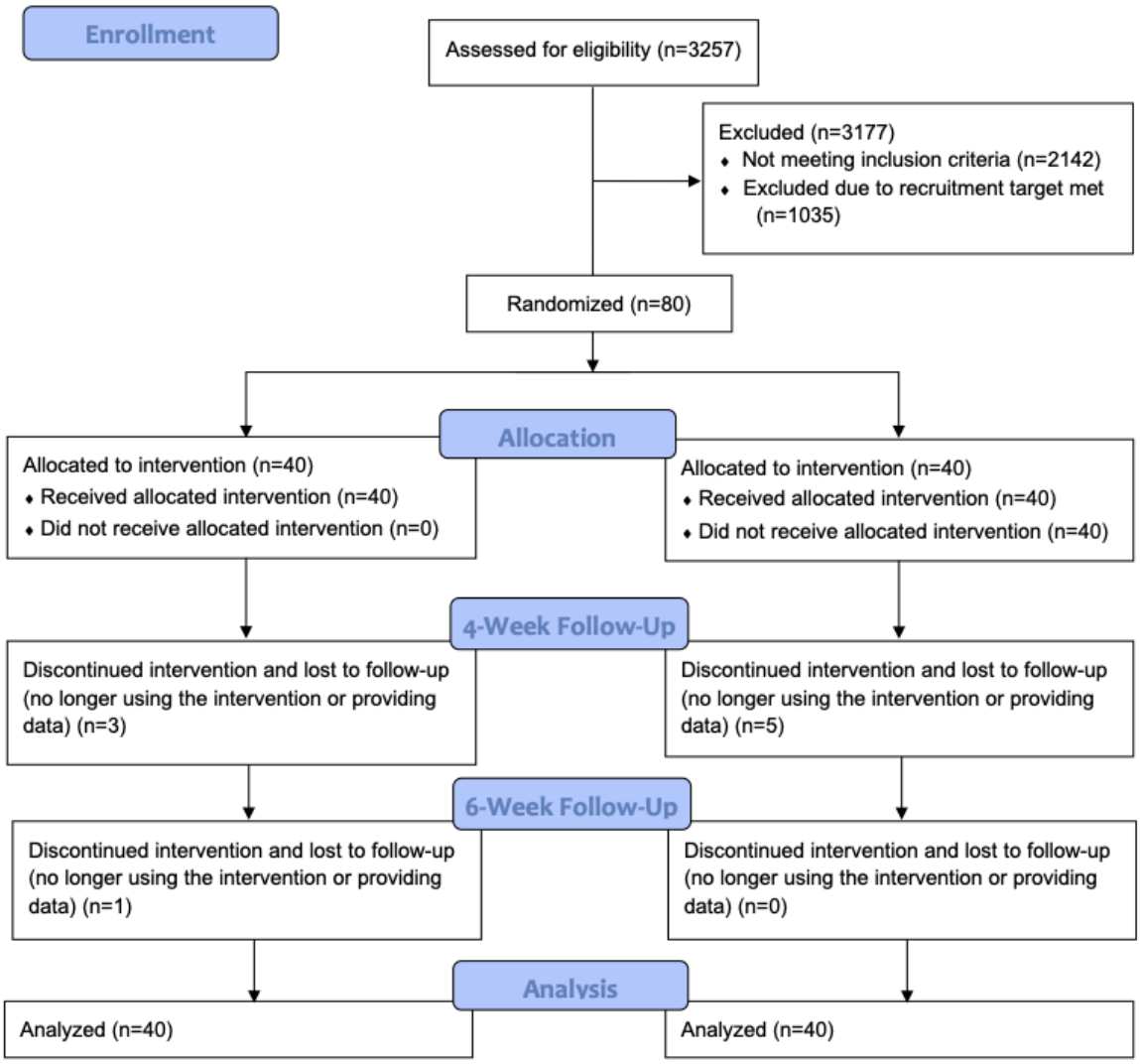
Results

Between January 18, 2019, and January 31, 2019, 80 eligible participants were randomly assigned to the intervention group (n=40) or control group (n=40), as shown in Figure 1.

The intervention and follow-up period lasted for 42 days, from January 31, 2019 to March 13, 2019. In the 13-day recruitment period, 5736 visitors viewed the registration system, 3257 of whom tried to complete the registration form; 1115 eligible individuals provided informed consent and completed the registration forms, among whom the first 80 who completed the registration questionnaire and subscribed to the SCAMPI OA were registered in the study. Table 1 shows the baseline characteristics, smoking status, and quitting reasons of these participants.

XSL•FO
RenderX

Figure 1. Flowchart of trial recruitment, enrollment, and randomization.



XSL•FO
RenderX

Table 1. Baseline characteristics, smoking status, and quitting reasons.

Characteristics	Intervention group (n=40)	Control group (n=40)
Age (years), mean (SD)	32.4 (6.0)	31.4 (5.1)
Smoking status, n (%)		
Daily smoker	34 (85)	38 (95)
Non-daily smoker	6 (15)	2 (5)
Past smoking cessation attempt, n (5)		
Yes	33 (83)	31 (78)
No	7 (18)	9 (23)
Use of smoking cessation services, n (%)		
Yes	24 (72)	28 (90)
No	9 (27)	3 (10)
Factors that trigger smoking, n (%)		
Socializing	17 (43)	19 (48)
At meals	10 (25)	19 (48)
When feeling down	26 (65)	32 (80)
When feeling excited	23 (58)	24 (60)
When nervous and anxious	17 (43)	26 (65)
When feeling tired	26 (65)	29 (73)
When involved with entertainment	13 (33)	16 (40)
When reading or writing	0 (0)	8 (20)
When working	9 (23)	12 (30)
When alone	11 (28)	14 (35)
Drinking alcohol	10 (25)	17 (43)
Reasons to quit smoking, n (%)		
Personal health	20 (50)	21 (53)
Family health	32 (80)	28 (70)
Doctor’s advice	16 (40)	18 (45)
Smoke-free policy in public area	15 (38)	17 (43)
Social stigma	4 (10)	11 (28)
Concerned about environment for next generation	9 (23)	17 (43)
Marital status, n (%)		
Single	9 (23)	8 (20)
In a relationship	9 (23)	6 (15)
Married	22 (55)	26 (65)
Parenting status, n (%)		
Has a child (or children)	31 (78)	23 (58)
No child	9 (23)	17 (43)

Demographics

Of the participants, 90% (72/80) were daily smokers. On average, participants had a strong desire to quit smoking (median quit smoking willingness score=4.3/5.0, interquartile range 4.1-4.5). In addition, 64 of 80 (80%) participants had tried to quit smoking, and 52 of 64 participants (81%) reported having previously used smoking cessation services. Major triggers to smoking were unwanted emotions like feeling depressed and feeling tired as well as concerns for personal or family health.

Retention

At 6 weeks, 36 of 40 (90%) intervention participants and 35 of 40 (88%) control participants provided complete self-reported

XSL•FO
RenderX

data on their daily smoking status via WeChat. For the remaining 9 participants, 7 did not respond, and 2 provided partial smoking data (both were smokers).

Self-reported 30-day smoking abstinence was then collected from a total of 73 participants. Of these participants, 20 (intervention n=15, control n=5) self-reported quitting at 6 weeks and were asked to provide their addresses to receive cotinine test kits to verify their self-reported data. Of these 20 participants, 17 (85%) provided their addresses, of whom 16 (16/20, 80%, intervention n=13, control n=3) completed the tests.

Data on biochemically verified 30-day smoking abstinence at 6 weeks (primary outcome) was available for 69 participants (intervention n= 35, control n=34). Primary analysis included all 80 participants, with missing outcomes treated as not quitting (ie, treatment failure).

Primary Outcome

Of the 40 participants in the intervention group, 10 (25%) were abstinent (biochemically verified) at the 6-week follow-up, compared with only 2 (5%) of the 40 participants in the control group (RR=5.0, 95% CI 1.2-21.4; *P*=.03). The unadjusted OR was 6.3 (95% CI 1.3-31.1).

Secondary Outcomes

Secondary outcomes are summarized in Tables 2 and 3.

Table 2. Summary of participation and smoking data.

Outcomes	Intervention group (n=40)	Control group (n=40)
Retention rate, n (%)		
Week 4	37 (93)	35 (88)
Week 6	36 (90)	35 (88)
Number of times interacting with the SCAMPI program during week 6, mean (SD)	81.88 (37.23)	61.40 (26.40)
Cigarettes smoked per day, mean (SD)		
Week 4	3.43 (6.39)	7.60 (8.79)
Week 6	3.00 (6.02)	5.81 (7.72)

Table 3. Odds ratios for secondary outcomes.

Outcomes	Intervention group, n (%)	Control group, n (%)	Odds ratio (95% CI)	*P* value
Cigarette consumption reduction at week 6	25 (63)	26 (65)	0.89 (0.36-2.23)	.82
7-day self-reported smoking abstinence				
Week 4	25 (63)	13 (33)	3.46 (1.38-8.69)	.007
Week 6	23 (58)	15 (38)	2.25 (0.92-5.52)	.07
30-day self-reported smoking abstinence at week 6	15 (38)	5 (13)	4.2 (1.35-13.06)	.02

Engagement

On average, participants in the intervention group had 82 interactions with their assigned program, while participants in the control group had an average of 61 interactions with their program over the 6-week trial period.

Smoking Behavior Change

Both groups reduced cigarette consumption over the 6 weeks, although the average daily consumption was lower in the intervention group, but the difference was not statistically significant (RR=0.47, 95% CI 0.16-1.35; *P*=.16). Compared to baseline, 25 and 26 participants achieved cigarette consumption reduction at week 6 in the intervention and control groups, respectively (unadjusted OR 0.89, 95% CI 0.36-2.23; *P*=.82). At the 4-week follow-up, 25 of 40 participants (63%) in the intervention group and 23 of 40 participants (33%) in the control group reported 7-day smoking abstinence. At week 6, 23 of 40 participants (58%) in the intervention group and 15 of 40 participants (38%) in the control group achieved 7-day smoking abstinence (unadjusted OR at 4 weeks=3.46, 95% CI 1.38-8.69; *P*=.007; unadjusted OR at 6 weeks=2.25, 95% CI 0.92-5.52; *P*=.07). At 6 weeks, 15 of 40 participants (38%) in the intervention group and 5 of 40 participants (13%) in the control group reported 30-day smoking abstinence (RR=3.0, 95% CI 1.2-7.5; unadjusted OR=4.2, 95% CI 1.35-13.06; *P*=.02).

Satisfaction

Overall, 36 of 40 participants (90%) from the intervention group responded to the end-of-trial questionnaire, rating the SCAMPI program 4.6 out of 5 (95% CI 4.3-4.8). Almost all of the responding participants (35/36, 97%) were willing to introduce SCAMPI to others, 23 of 36 participants (64%) were willing to pay to use the program, 35 of 36 participants (97%) indicated they would use the program at least once in the next 12 months, and 27 of 36 participants (75%) reported that they would use the program more than 10 times in the next 12 months.

XSL•FO
RenderX

Discussion

Principal Findings

In this pilot smoking cessation program underpinned by a theoretical model and empirical evidence [27], Chinese male smokers aged 25-44 years who used the program achieved significantly higher 30-day (both self-reported and biochemically verified) smoking abstinence at the 6-week follow-up as well as 7-day self-reported smoking abstinence (at the 4-week and 6-week follow-ups) than those randomized to the control program. This is the first trial done entirely within the WeChat ecosystem [28-31]. The potential for efficient recruitment (3257 visitors within 13 days) supports the use of this popular social network platform to engage smokers, deliver interventions, and conduct trials via the same platform. Significant numbers of participants (n=52) had ever used smoking cessation services, which may be why the participants engaged with the registration system after reading the advertisement on WeChat. The high retention rate in both arms suggests our strategy of using frequent micro-incentive credits was effective. The program also achieved remarkably high satisfaction ratings from its users.

Limitations

There were a number of limitations. As a pilot trial, it was not powered to detect a significant effect, and the study period (6 weeks) was insufficient to assess long-term (6-month or 12-month) smoking abstinence. However, this pilot trial has shown that the SCAMPI program is feasible and acceptable as a smoking cessation tool on WeChat, for the largest social network user group and smoking population group in China. The recruitment strategy will be able to recruit a bigger sample size within a reasonable length of time (approximately 1 month). The trial design and mode of intervention delivery have shown the ability to maintain a high retention rate of study participants. A future randomized control trial is needed with a larger sample size and 6-month follow-up to evaluate the program's effectiveness in supporting long-term smoking cessation. Our findings may not be generalizable to other people, in China or elsewhere. Nevertheless, WeChat is widely used by people of different age groups and in a range of countries [32,33]. However, SCAMPI was developed using a systematic, replicable method and could be used to tailor mobile social network–based smoking cessation interventions to different population groups.

Another limitation is that participants were not blinded. This may lead to type 1 error as participants in the intervention group may tend to report themselves as smoking- abstinent. However, unlike most online trials relying on self-reported data, biochemically verified data was captured. With more resources, a future trial could use more frequent cotinine testing.

There are a number of concerns with using a social network platform to deliver health interventions and implement clinical trials, including reduced experimental control and the risk of reduced reliability [34-38]. However, these concerns may be addressed by improving the quality of the digital trial and health product application. For example, in future trials, more precise technology could be applied to increase study data validity and reliability (eg, using the mobile or computer camera to verify user identities) [39]. On the other hand, there are advantages; digital trials provide a level of anonymity for people unwilling to participate in face-to-face research, typically because of embarrassment and stigma [40-42].

They also overcome logistical barriers for some participants, such as people with physical disability, transport issues, or geographic remoteness [43]. Applying social network platforms to recruit clinical trial participants and deliver trial interventions can facilitate the inclusion of low-incidence or hidden populations, such as young adult smokers [44]. Therefore, the application of a social network platform may reduce the chances of under-reporting and increase the generalizability of the data [40-45].

User experience with the SCAMPI program may have been impacted by the fact that it was embedded within a social network platform. Users may have been disturbed by WeChat messages while using the program. However, a new WeChat feature allows users to temporarily exit from a mini-program to answer WeChat messages. The SCAMPI program was designed to be simple and concise; no function was designed to be used for longer than 1 minute. No session should take >5 minutes of a user's time. Currently, the program is only able to collect data on the number of times participants interacted with the program. It was difficult to measure the duration of each interaction. Future studies could monitor the duration of user-program interactions to understand user interaction with the program to improve user experience.

Implications

The greatest strength of our program and pilot trial was the use of an existing, widely used social network platform, ensuring wide reach, efficient recruitment, and frequent engagement. The preliminary effectiveness and acceptability of the SCAMPI program is a useful reference point for a larger definitive trial and to guide the development and deployment of other social network–based health behavior change interventions. If applied as a common service on the WeChat platform, it may be necessary to find nonmonetary incentives to maintain the high engagement, such as enhanced gamification.

Acknowledgments

The trial was completely funded by the University of Auckland. Promotion of the program on the WeChat platform was provided by the Tencent company.

XSL•FO
RenderX

Authors' Contributions

CB contributed to the concept development, study design, data interpretation, article writing, and article reviewing. EH provided cultural insights for intervention development, data interpretation, article writing, and article reviewing. YJ developed the figures and contributed to the study design, statistical analysis, data interpretation, article writing, and article review. RW consulted on the product concept development and development process guidance. TY consulted on the product's cultural adaptation and the trial's China ethic applications. JC performed the literature search and contributed to the technical and conceptual development of the SCAMPI program, study design, data collection, data interpretation, and article writing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Informed consent form.
[DOCX File , 20 KB - mhealth_v8i10e17522_app1.docx]

Multimedia Appendix 2
Conversion of behaviour factors to functions and content of the SCAMPI programme.
[DOCX File , 21 KB - mhealth_v8i10e17522_app2.docx]

Multimedia Appendix 3
Screenshots and QR-codes of the SCAMPI programme.
[DOCX File , 1446 KB - mhealth_v8i10e17522_app3.docx]

Multimedia Appendix 4
Registration questionnaire.
[DOCX File , 22 KB - mhealth_v8i10e17522_app4.docx]

Multimedia Appendix 5
End-of-trial questionnaire.
[DOCX File , 17 KB - mhealth_v8i10e17522_app5.docx]

Multimedia Appendix 6
CONSORT-EHEALTH (V 1.6.1) checklist.
[PDF File (Adobe PDF File), 2148 KB - mhealth_v8i10e17522_app6.pdf]

References

1. China Fact Sheet. The Tobacco Atlas. 2018. URL: https://tobaccoatlas.org/wp-content/uploads/pdf/china-country-facts.pdf [accessed 2020-10-03]
2. Countries with the Most Smokers. The Tobacco Atlas. 2015. URL: https://tobaccoatlas.org/data-visualization/countries-with-the-most-smokers/ [accessed 2020-10-03]
3. Global Adult Tobacco Survey (GATS) China 2010 Country Report. World Health Organization. 2010 Aug 06. URL: https://www.who.int/tobacco/surveillance/survey/gats/en_gats_china_report.pdf?ua=1 [accessed 2020-10-03]
4. Zhang M, Wang L, Li Y, Li X, Jiang Y, Hu N, et al. [Cross-sectional survey on smoking and smoking cessation behaviors among Chinese adults in 2010]. Zhonghua Yu Fang Yi Xue Za Zhi 2012 May;46(5):404-408. [Medline: 22883725]
5. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database Syst Rev 2019 Oct 22;10:CD006611. [doi: 10.1002/14651858.CD006611.pub5] [Medline: 31638271]
6. BinDhim NF, McGeechan K, Trevena L. Smartphone Smoking Cessation Application (SSC App) trial: a multicountry double-blind automated randomised controlled trial of a smoking cessation decision-aid 'app'. BMJ Open 2018 Jan 21;8(1):e017105. [doi: 10.1136/bmjopen-2017-017105] [Medline: 29358418]
7. Poushter J. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. Pew Research Center. 2016 Feb 22. URL: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ [accessed 2020-10-03]
8. Top 50 Countries by Smartphone Users and Penetration. Newzoo. 2017. URL: https://newzoo.com/insights/rankings/top-50-countries-by-smartphone-penetration-and-users/ [accessed 2020-10-03]
9. 2012 iiMedia Report of Chinese Smartphone Users' Demographics and Distribution. iiMedia Research. 2013 Mar 06. URL: https://www.iimedia.cn/c400/36505.html [accessed 2020-10-03]

XSL•FO
RenderX

10. WeChat Chinese and International versions combined monthly active users over 1.2 billion, mini-programme daily users over 400 million. TechWeb. 2020. URL: https://tech.sina.cn/2020-05-13/detail-iircuyvi2911382.d.html [accessed 2020-10-03]
11. 2017-2018 WeChat Mini-programme Market Development Research Report. Instant Application. 2018. URL: http://co-image.qichacha.com/upload/chacha/att/20180130/1517306145864730.pdf [accessed 2020-10-03]
12. 9 Ways The WeChat Super App Is Becoming The Front Door To Chinese Healthcare. CBINSIGHTS. 2020. URL: https://www.cbinsights.com/research/wechat-digital-health-china/ [accessed 2020-10-03]
13. Li T, Ding W, Li X, Lin A. Mobile health technology (WeChat) for the hierarchical management of community hypertension: protocol for a cluster randomized controlled trial. Patient Prefer Adherence 2019;13:1339-1352 [FREE Full text] [doi: 10.2147/PPA.S215719] [Medline: 31496664]
14. Dorje T, Zhao G, Scheer A, Tsokey L, Wang J, Chen Y, et al. SMARTphone and social media-based Cardiac Rehabilitation and Secondary Prevention (SMART-CR/SP) for patients with coronary heart disease in China: a randomised controlled trial protocol. BMJ Open 2018 Jun 30;8(6):e021908. [doi: 10.1136/bmjopen-2018-021908] [Medline: 29961032]
15. Dong Y, Wang P, Dai Z, Liu K, Jin Y, Li A, et al. Increased self-care activities and glycemic control rate in relation to health education via Wechat among diabetes patients: A randomized clinical trial. Medicine 2018 Dec;97(50):e13632 [FREE Full text] [doi: 10.1097/MD.0000000000013632] [Medline: 30558051]
16. Mao L, Lu J, Zhang Q, Zhao Y, Chen G, Sun M, et al. Family-based intervention for patients with type 2 diabetes via WeChat in China: protocol for a randomized controlled trial. BMC Public Health 2019 Apr 05;19(1):381 [FREE Full text] [doi: 10.1186/s12889-019-6702-8] [Medline: 30953483]
17. Jiang Y, Liu F, Guo J, Sun P, Chen Z, Li J, et al. Evaluating an Intervention Program Using WeChat for Patients With Chronic Obstructive Pulmonary Disease: Randomized Controlled Trial. J Med Internet Res 2020 Apr 21;22(4):e17089 [FREE Full text] [doi: 10.2196/17089] [Medline: 32314971]
18. Chen J, Ho E, Jiang Y, Whittaker R, Yang T, Bullen C. A Mobile Social Network-Based Smoking Cessation Intervention for Chinese Male Smokers: Protocol for a Pilot Randomized Controlled Trial. JMIR Res Protoc 2020 Sep 18;9(9):e18071 [FREE Full text] [doi: 10.2196/18071] [Medline: 32945261]
19. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
20. 2015 Chinese Adult Tobacco Use Research Report. Chinese Center for Disease Control and Prevention. 2015. URL: http://www.tcrc.org.cn/UploadFiles/2016-03/318/201603231215175500.pdf [accessed 2020-10-03]
21. Big Data about Smoking: Survey of Current Chinese Smokers. Sohu Inc. 2018 Mar 29. URL: http://www.sohu.com/a/226632732_634722 [accessed 2020-10-03]
22. Michie S, Hyder N, Walia A, West R. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. Addict Behav 2011 Apr;36(4):315-319. [doi: 10.1016/j.addbeh.2010.11.016] [Medline: 21215528]
23. WeChat Mini-programme Design Guideline. Tencent Inc. 2019. URL: https://developers.weixin.qq.com/miniprogram/design/ [accessed 2020-10-03]
24. Clinical Guideline for Smoking Cessation in China (2015). National Health and Family Planning Commission of China. 2015. URL: https://untobaccocontrol.org/impldb/wp-content/uploads/reports/china_2016_annex8_clinical_smoking_cessation_guidelines_2015.pdf [accessed 2020-10-03]
25. Piper ME, Bullen C, Krishnan-Sarin S, Rigotti NA, Steinberg ML, Streck JM, et al. Defining and Measuring Abstinence in Clinical Trials of Smoking Cessation Interventions: An Updated Review. Nicotine Tob Res 2020 Jun 12;22(7):1098-1106. [doi: 10.1093/ntr/ntz110] [Medline: 31271211]
26. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
27. Cheng F, Xu J, Su C, Fu X, Bricker J. Content Analysis of Smartphone Apps for Smoking Cessation in China: Empirical Study. JMIR Mhealth Uhealth 2017 Jul 11;5(7):e93 [FREE Full text] [doi: 10.2196/mhealth.7462] [Medline: 28698170]
28. Fu X, Xiang Y, Zhang R, Bricker J, Su C, Cheng F. A content analysis and development strategy exploration for iPhone apps for smoking cessation. Chinese Preventive Medicine 2016;17(11):829-833. [Medline: 28698170]
29. Thornton L, Quinn C, Birrell L, Guillaumier A, Shaw B, Forbes E, et al. Free smoking cessation mobile apps available in Australia: a quality review and content analysis. Aust N Z J Public Health 2017 Dec;41(6):625-630. [doi: 10.1111/1753-6405.12688] [Medline: 28749591]
30. Robinson CD, Seaman EL, Grenen E, Montgomery L, Yockey RA, Coa K, et al. A content analysis of smartphone apps for adolescent smoking cessation. Transl Behav Med 2020 Feb 03;10(1):302-309 [FREE Full text] [doi: 10.1093/tbm/iby113] [Medline: 30476293]
31. Choi J, Noh G, Park D. Smoking cessation apps for smartphones: content analysis with the self-determination theory. J Med Internet Res 2014 Feb 12;16(2):e44 [FREE Full text] [doi: 10.2196/jmir.3061] [Medline: 24521881]
32. Aurora Big Data: Social Network Industry Research Report 2019. Aurora Mobile. 2019 Apr 11. URL: http://www.199it.com/archives/860472.html [accessed 2020-10-03]

XSL•FO
RenderX

33. WeChat applet "fire" overseas? You have to pay attention to this market. PINGWEST. 2019 Dec 21. URL: https://www.pingwest.com/a/200912 [accessed 2020-10-03]
34. Krombholz K, Merkl D, Weippl E. Fake identities in social media: A case study on the sustainability of the Facebook business model. J Serv Sci Res 2012 Dec 31;4(2):175-212. [doi: 10.1007/s12927-012-0008-z]
35. Shahane P, Gore D. A Survey on Classification Techniques to Determine Fake vs. Real Identities on Social Media Platforms. 2018;10(2):175-179.
36. Litman L, Robinson J, Abberbock T. TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. Behav Res Methods 2017 Apr;49(2):433-442 [FREE Full text] [doi: 10.3758/s13428-016-0727-z] [Medline: 27071389]
37. Dandurand F, Shultz TR, Onishi KH. Comparing online and lab methods in a problem-solving experiment. Behav Res Methods 2008 May;40(2):428-434. [doi: 10.3758/brm.40.2.428] [Medline: 18522052]
38. Weigold A, Weigold IK, Drakeford NM, Dykema SA, Smith CA. Equivalence of paper-and-pencil and computerized self-report surveys in older adults. Computers in Human Behavior 2016 Jan;54:407-413. [doi: 10.1016/j.chb.2015.08.033]
39. Stoops WW, Dallery J, Fields NM, Nuzzo PA, Schoenberg NE, Martin CA, et al. An internet-based abstinence reinforcement smoking cessation intervention in rural smokers. Drug Alcohol Depend 2009 Nov 01;105(1-2):56-62 [FREE Full text] [doi: 10.1016/j.drugalcdep.2009.06.010] [Medline: 19615830]
40. Chebli J, Blaszczynski A, Gainsbury SM. Internet-Based Interventions for Addictive Behaviours: A Systematic Review. J Gambl Stud 2016 Dec;32(4):1279-1304. [doi: 10.1007/s10899-016-9599-5] [Medline: 27002522]
41. Whelan R, Cao Z, O'Halloran L, Pennie B. Chapter 27 - Genetics, imaging, cognition: big data approaches to addiction research. In: Verdejo-Garcia A, editor. Cognition and Addiction: A Researcher's Guide from Mechanisms Towards Interventions. San Diego, CA: Elsevier; 2020:365-377.
42. Gainsbury S, Hing N, Suhonen N. Professional help-seeking for gambling problems: awareness, barriers and motivators for treatment. J Gambl Stud 2014 Jun;30(2):503-519. [doi: 10.1007/s10899-013-9373-x] [Medline: 23494244]
43. Proudfoot J, Klein B, Barak A, Carlbring P, Cuijpers P, Lange A, et al. Establishing guidelines for executing and reporting Internet intervention research. Cogn Behav Ther 2011;40(2):82-97. [doi: 10.1080/16506073.2011.573807] [Medline: 25155812]
44. Ramo DE, Rodriguez TMS, Chavez K, Sommer MJ, Prochaska JJ. Facebook Recruitment of Young Adult Smokers for a Cessation Trial: Methods, Metrics, and Lessons Learned. Internet Interv 2014 Apr;1(2):58-64 [FREE Full text] [doi: 10.1016/j.invent.2014.05.001] [Medline: 25045624]
45. Thornton LK, Harris K, Baker A, Johnson M, Kay-Lambkin FJ. Recruiting for addiction research via Facebook. Drug Alcohol Rev 2016 Jul 14;35(4):494-502. [doi: 10.1111/dar.12305] [Medline: 26179224]

Abbreviations

COM-B: “Capability,” “Opportunity,” “Motivation,” “Behaviour”
OA: official account
OR: odds ratio
RR: relative risk

Edited by G Eysenbach; submitted 18.12.19; peer-reviewed by J Bricker, G Klein, N Minian; comments to author 10.03.20; revised version received 04.05.20; accepted 07.07.20; published 23.10.20.

Please cite as:
Chen J, Ho E, Jiang Y, Whittaker R, Yang T, Bullen C
Mobile Social Network–Based Smoking Cessation Intervention for Chinese Male Smokers: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(10):e17522
URL: http://mhealth.jmir.org/2020/10/e17522/
doi:10.2196/17522
PMID:33095184

©Jinsong Chen, Elsie Ho, Yannan Jiang, Robyn Whittaker, Tingzhong Yang, Christopher Bullen. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 23.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Relationship Between Patient Engagement and Depressive Symptoms Among People Living With HIV in a Mobile Health Intervention: Secondary Analysis of a Randomized Controlled Trial

Yu Zeng[1*], MB; Yan Guo[1,2*], PhD; Linghua Li[3*], MD; Y Alicia Hong[4], PhD; Yiran Li[1], MB; Mengting Zhu[1], BS; Chengbo Zeng[5,6], MS; Hanxi Zhang[7], MS; Weiping Cai[3], MD; Cong Liu[3], MSN; Shaomin Wu[1], MS; Peilian Chi[8], PhD; Aliza Monroe-Wise[9], MD; Yuantao Hao[1,2], PhD; Rainbow Tin Hung Ho[10,11], PhD

[1]Department of Medical Statistics, School of Public Health, Sun Yat-sen University, Guangzhou, China

[2]Sun Yat-sen Center for Global Health, Guangzhou, China

[3]Department of Infectious Diseases, Guangzhou Number Eight People's Hospital, Guangzhou, China

[4]Department of Health Administration and Policy, College of Health and Human Services, George Mason University, Fairfax, VA, United States

[5]South Carolina SmartState Center of Healthcare Quality, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[6]Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[7]National Center of AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

[8]Department of Psychology, Faculty of Social Sciences, University of Macau, Macau, China

[9]Department of Global Health, University of Washington, Seattle, WA, United States

[10]Department of Social Work & Social Administration, The University of Hong Kong, Hong Kong, China

[11]Centre on Behavioral Health, The University of Hong Kong, Hong Kong, China

[*]these authors contributed equally

Corresponding Author:
Yan Guo, PhD
Department of Medical Statistics
School of Public Health
Sun Yat-sen University
No.74, 2nd Zhongshan Road
Guangzhou, 510000
China
Phone: 86 13501502582
Email: guoy8@mail.sysu.edu.cn

Abstract

Background: Associations between higher levels of patient engagement and better health outcomes have been found in face-to-face interventions; studies on such associations with mobile health (mHealth) interventions have been limited and the results are inconclusive.

Objective: The objective of this study is to investigate the relationship between patient engagement in an mHealth intervention and depressive symptoms using repeated measures of both patient engagement and patient outcomes at 4 time points.

Methods: Data were drawn from a randomized controlled trial (RCT) of an mHealth intervention aimed at reducing depressive symptoms among people living with HIV and elevated depressive symptoms. We examined the association between patient engagement and depressive symptoms in the intervention group (n=150) where participants received an adapted cognitive-behavioral stress management (CBSM) course and physical activity promotion on their WeChat social media app. Depressive symptoms were repeatedly measured using the Patient Health Questionnaire (PHQ-9) at baseline and 1 month, 2 months, and 3 months. Patient engagement was correspondingly measured by the completion rate, frequency of items completed, and time spent on the program at 1 month, 2 months, and 3 months. Latent growth curve models (LGCMs) were used to explore the relationship between patient engagement and depressive symptoms at multiple time points in the intervention.

XSL•FO
RenderX

Results: The mean PHQ-9 scores were 10.2 (SD 4.5), 7.7 (SD 4.8), 6.5 (SD 4.7), and 6.7 (SD 4.1) at baseline, 1 month, 2 months, and 3 months, respectively. The mean completion rates were 50.6% (SD 31.8%), 51.5% (SD 32.2%), and 50.8% (SD 33.7%) at 1, 2, and 3 months, respectively; the average frequencies of items completed were 18.0 (SD 14.6), 32.6 (SD 24.8), and 47.5 (SD 37.2) at 1, 2, and 3 months, respectively, and the mean times spent on the program were 32.7 (SD 66.7), 65.4 (SD 120.8), and 96.4 (SD 180.4) minutes at 1, 2, and 3 months, respectively. LGCMs showed good model fit and indicated that a higher completion rate (β at 3 months=–2.184, *P*=.048) and a greater frequency of items completed (β at 3 months=–0.018, *P*=.04) were associated with fewer depressive symptoms at 3 months. Although not significant, similar trends were found in the abovementioned relationships at 1 and 2 months. There was no significant relationship between time spent on the program and depressive symptoms.

Conclusions: This study revealed a positive association between patient engagement and health outcomes at 3 months of an mHealth intervention using LGCMs and repeated measures data. The results underscore the importance of improving patient engagement in mHealth interventions to improve patient-centered health outcomes.

Trial Registration: Chinese Clinical Trial Registry ChiCTR-IPR-17012606; https://tinyurl.com/yxb64mef

International Registered Report Identifier (IRRID): RR2-10.1186/s12889-018-5693-1

(JMIR Mhealth Uhealth 2020;8(10):e20847) doi:10.2196/20847

KEYWORDS

mHealth; patient engagement; latent growth curve model; depressive symptoms; HIV

Introduction

Mobile health (mHealth) has increasingly become a promising tool to deliver treatments for a range of psychosocial disorders [1-4]. Studies have shown that mHealth interventions are effective in reducing depressive symptoms among people living with HIV and elevated depressive symptoms [5-8]. In comparison with traditional face-to-face interventions, mHealth interventions have the potential to implement personalized and cost-effective interventions, provide real-time feedback, and overcome geographical and logistical barriers [9].

Despite the growing success of mHealth interventions, the relationship between patient engagement and intervention outcomes is still unclear. Patient engagement is defined as the extent to which participants actively interact with intervention content [10]. In face-to-face randomized controlled trials (RCTs), patient engagement has been found to be an essential predictor of intervention outcomes, with higher levels of patient engagement predicting better health outcomes of participants [11]. However, such a relationship has not yet been confirmed in mHealth interventions [12]. Findings regarding patient engagement and intervention outcomes in mHealth interventions are mixed. Some studies found a relationship similar to that observed in face-to-face interventions [13-18], while others found no significant relationship [19-21].

Understanding the relationship between patient engagement and health outcomes in mHealth interventions is critical for developing and implementing effective mHealth programs [12]. In most mHealth studies, patient engagement was only assessed at a single point in time and changes in patient outcomes were typically evaluated between pre- and postintervention [21]. There is a lack of examination of dose-response relationships between patient engagement and treatment outcomes during the entire course of the interventions. It is unclear whether such a dose-response relationship exists in mHealth interventions, and if yes, whether such a relationship occurs early in the intervention as demonstrated in some face-to-face interventions [22] or appears later in the intervention. Existing mHealth studies have varying intervention durations, which may lead to differential results of the dose-response relationship. For example, for the same purpose of reducing depressive symptoms among the general population, one mHealth intervention lasted 4 weeks whereas another was 24 weeks in duration [19,23]. With different lengths of treatment, it might be difficult to compare the dose-response relationship in these mHealth interventions, especially when the relationship was only evaluated pre- and postintervention. More studies are needed to explore the potentially changing relationship between patient engagement and health outcomes during the course of mHealth interventions. Longitudinal analyses with repeated measures of both patient engagement and treatment outcomes are likely to improve our understanding of the dose-response relationship throughout the course of an intervention.

To fill the gaps in the existing literature, this study aimed to explore the relationship between patient engagement and depressive symptoms during the 3-month course of an RCT of an mHealth program for people living with HIV and elevated depressive symptoms, the Run4Love intervention [24]. We used 3 measures of patient engagement to capture multiple aspects of this construct in the study, as suggested by a recent systematic review on adherence reporting in online RCTs [12]. Using 4 time points of outcome measures and the corresponding patient engagement measures, we examined the levels of and changes in patient engagement and depressive symptoms at 1, 2, and 3 months and the relationship between patient engagement and depressive symptoms at 1, 2, and 3 months.

Methods

Study Setting

This study is a secondary analysis of data from the intervention group of a 3-month RCT (Chinese Clinical Trial Registry [ChiCTR-IPR-17012606]) that was conducted to examine the efficacy of a WeChat-based intervention on reducing depressive symptoms among people living with HIV and elevated depressive symptoms [8]. WeChat, similar to WhatsApp, is a widely used social media platform for instant communication

XSL•FO
RenderX

in China with 1.1 billion active users [25]. Participants were recruited from the outpatient department of a large hospital designated for HIV treatment in 2017 in Guangzhou, the third largest city in China, with a population of more than 13 million [26]. A total of 300 people living with HIV who had elevated depressive symptoms (Center for Epidemiological Studies Depression Scale [CES-D] score ≥16) and used WeChat were randomized into the intervention or wait-list control group in a 1:1 ratio. The institutional review board of Sun Yat-sen University approved the intervention protocol, and the detailed recruitment and randomization procedure were described elsewhere [24].

Intervention

In addition to routine care, the Run4Love intervention consisted of two major components, an adapted cognitive-behavioral stress management (CBSM) course and physical activities promotion. The adapted CBSM course included content covering coping skills, muscle relaxation, and meditation tutorials. The physical activities promotion consisted of information on the benefits of regular exercise, exercise guidance, and healthy dietary suggestions. During the 3-month intervention, multimedia items were delivered to participants to support the adapted CBSM course and physical activities promotion via WeChat in the form of short articles, audio recordings, and posters almost once a day (5 to 6 items per week). Out of 65 items in total, 29 were short articles with an average of 1300 words, requiring about 5 minutes to read; 24 were posters with motivational messages, requiring about 30 seconds to read; and 12 were audio recordings in Chinese requiring 5 to 10 minutes to listen to. The WeChat-based Run4Love platform with extended functions was used to automatically deliver intervention items and collect real-time data on patient engagement (eg, time spent on an article, audio recording, or poster). The intervention design has been detailed elsewhere [24].

During the intervention, we used several measures to promote engagement: (1) automatically sending feedback on each participant's compliance via WeChat on a weekly basis (eg, number of items completed and not completed); (2) giving verbal encouragement and financial incentives (up to US $2) via WeChat on a weekly basis based on the level of compliance; and (3) conducting motivational interviews by phone to sustain participation and help overcome barriers of compliance at 1 week, 1 month, and 2 months after baseline.

Measurement

Depressive Symptoms

Depressive symptoms were assessed by the Chinese version of the Patient Health Questionnaire (PHQ-9) at baseline, 1 month, 2 months, and 3 months [27]. Compared with the 20-item measurement of CES-D, PHQ-9 is shorter but has demonstrated both good reliability and validity in the Chinese populations [28,29]. The scale consists of 9 items such as "Little interest or pleasure in doing things" and "Feeling down, depressed, or hopeless" to measure the depressive status of the participants over the previous 2 weeks. Each item was rated on a 4-point Likert scale ranging from 0 (not at all) to 3 (nearly every day), providing a 0 to 27 total severity score, with higher scores indicating increased depressive symptoms [30]. A good reliability of PHQ-9 was shown in the study, and the Cronbach alphas were .814, .861, .872, and .852 at baseline, 1 month, 2 months, and 3 months, respectively.

Patient Engagement

Three cumulative measures, including completion rate, frequency of items completed, and time spent on the program, were used to assess patient engagement in this study. Completion rate was the ratio of the number of items a patient completed to the number of items assigned, where the items that had been clicked were considered completed. Frequency of items completed was assessed not only by the number of items read or listened to by the participants, but also by the number of times the items were repeatedly read or listened to. Compared with the completion rate, frequency of items completed captured the repeat aspect of patient engagement. For example, if the same item was read or listened to twice by a participant, frequency was counted as 2 and the frequency of items completed for the participant was the accumulative frequency for all the items. The time spent on the program was measured by the overall time a participant spent on the Run4Love program, including reading or listening to the delivered items during the course of the intervention. We repeatedly assessed these 3 measures of patient engagement in conjunction with measurements of depressive symptoms that were evaluated at 1, 2, and 3 months after baseline, as has been recommended in previous studies [31]. Specifically, all 3 measures at 1, 2, and 3 months were used to assess the average patient engagement from baseline until 1, 2, and 3 months after baseline.

Baseline Characteristics

Sociodemographic characteristics and HIV-related variables were collected at baseline including age, gender, sexual orientation, education, marital status, employment status, income, and duration of HIV infection (years). Income was measured by asking the participants whether they had adequate income to cover their daily expenses (adequate or inadequate).

Statistical Analysis

Descriptive statistics on depressive symptoms, 3 measures of patient engagement, and baseline characteristics were provided. Mean and standard deviation were used to describe continuous variables with normal distribution, median and interquartile range (IQR) for continuous variables that did not follow a normal distribution, and frequency and percentage for categorical variables. Latent growth curve models (LGCMs) were constructed to examine the time-varying associations between patient engagement and depressive symptoms using repeated measure data at 4 time points [32]. Maximum likelihood estimation with robust standard errors (MLR) was used to handle missing data and obtain parameter estimates [32]. Descriptive analyses were conducted in SPSS Statistics version 20.0 (IBM Corporation), and LGCMs were conducted in Mplus version 7 (Muthen & Muthen).

The analyses of LGCMs were conducted in two steps [32]. First, we developed an unconditional LGCM to estimate the growth

XSL•FO
RenderX

trajectory of depressive symptoms across time. The initial level of depressive symptoms was represented as the intercept, and the average rate of change was represented as the slope in the model. For the intercept factor, all loadings were fixed to 1. For the slope factor, the first and second loadings were fixed to 0 and 1, and the remaining loadings were freely estimated from the data to test whether there was a linear or nonlinear growth trajectory (eg, linear growth pattern: 0, 1, 2, 3; nonlinear growth pattern: the third and fourth loadings significantly differ from 2 and 3, respectively).

Second, we developed 3 conditional LGCMs to examine the effects of patient engagement on depressive symptoms based on the 3 measures of patient engagement. The conditional model was extended from the unconditional model to incorporate one measure of patient engagement and related baseline characteristics as covariates. Since no baseline characteristics were found to significantly associate with the 3 measures of patient engagement or the outcome of depressive symptoms in this study, we included variables such as education, income, and duration of HIV infection that were associated with depressive symptoms among people living with HIV in the literature as control variables in the conditional models [33-35]. Associations between patient engagement and depressive symptoms were treated as time-varying effects, which were specified by regressing depressive symptoms (eg, PHQ-9 at 1 month) on the corresponding patient engagement variable (eg, completion rate at 1 month). Patient engagement, therefore, could be considered as a time-specific predictor of depressive symptoms after controlling for the influences of the underlying growth trajectory of depressive symptoms and baseline characteristics [32].

Model fit of LGCM was evaluated using chi-square tests and indices including the standardized root mean of residual (SRMR), confirmatory fit index (CFI), Tucker-Lewis index (TLI), and root mean-squared error of approximation (RMSEA). The criteria used to determine adequate model fit included: CFI≥0.95, TLI≥0.95, SRMR≤0.08, and RMSEA≤0.08 [36-38].

Results

Baseline Characteristics

As shown in Table 1, the mean age of participants was 28.0 (SD 5.8) years and the average duration of HIV infection was 2.4 (SD 2.3) years. The participants were predominately male (142/150, 94.7%), nonheterosexual (130/150, 86.7%), well-educated (98/150, 65.3%, with at least some college education), single (132/150, 88.0%), employed (123/150, 82.0%), and had adequate income to cover daily expenses (129/150, 86.0%).

Table 1. Baseline characteristics of people living with HIV and elevated depressive symptoms in the intervention group (n=150).

Variables	Value
Age in years, mean (SD)	28.0 (5.8)
Gender, n (%)	
Male	142 (94.7)
Female	8 (5.3)
Sexual orientation, n (%)	
Heterosexual	20 (13.3)
Homosexual, bisexual, uncertain	130 (86.7)
Education, n (%)	
Below or at high school level	52 (34.7)
Above high school level	98 (65.3)
Marital status, n (%)	
Single, divorced, widowed	132 (88.0)
Married	18 (12.0)
Employment status, n (%)	
Unemployed	27 (18.0)
Employed	123 (82.0)
Income, n (%)	
Adequate to cover daily expenses	129 (86.0)
Inadequate to cover daily expenses	21 (14.0)
Duration of HIV infection in years, mean (SD)	2.4 (2.3)

XSL•FO
RenderX

Change in Depressive Symptoms and Patient Engagement Over Time

Of the participants, 87.3% (131/150), 76.0% (114/150), and 92.0% (138/150) completed the online questionnaire for depressive symptoms at 1, 2, and 3 months, respectively, after baseline. Compared with the depressive symptom score at baseline (PHQ-9: mean 10.2 [SD 4.5]), all follow-up assessments reported a decrease in depressive symptoms with mean values of 7.7 (SD 4.8), 6.5 (SD 4.7), and 6.7 (SD 4.1) at 1, 2, and 3 months, respectively (Table 2).

Nearly all participants (149/150) read or listened to at least one item during the 3-month intervention. A moderate level of patient engagement was found during the intervention process. The completion rate remained around 50.0% (33/65) during the course of the intervention, with a mean of 50.6% (SD 31.8%), 51.5% (SD 32.2%), and 50.8% (SD 33.7%) at 1, 2, and 3 months, respectively. The mean frequency of items completed was 18.0 (SD 14.6), 32.6 (SD 24.8), and 47.5 (SD 37.2) at 1, 2, and 3 months, respectively. The mean time spent on the program was 32.7 (SD 66.7), 65.4 (SD 120.8), and 96.4 (SD 180.4) minutes at 1, 2, and 3 months, respectively. When comparing the mean and median scores in Table 2, we found that time spent on the program was positively skewed, with means much larger than the medians at 1, 2, and 3 months.

Table 2. Change in depressive symptoms and patient engagement across the 3-month intervention[a].

Variables	Baseline		1 month		2 months		3 months	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
PHQ-9[b]	10.2 (4.5)	10.0 (7.0-13.0)	7.7 (4.8)	8.0 (4.0-11.0)	6.5 (4.7)	6.0 (3.0-9.0)	6.7 (4.1)	7.0 (3.0-9.0)
Completion rate (%)	—[c]	—	50.6 (31.8)	54.2 (16.7-81.8)	51.5 (32.2)	53.0 (20.6-81.7)	50.8 (33.7)	47.6 (16.9-85.3)
Frequency of items completed	—	—	18.0 (14.6)	17.0 (5.8-26.0)	32.6 (24.8)	29.5 (10.0-51.3)	47.5 (37.2)	41.0 (11.0-78.3)
Time spent on program (min)	—	—	32.7 (66.7)	9.7 (2.5-42.1)	65.4 (120.8)	22.0 (5.3-80.3)	96.4 (180.4)	31.3 (6.7-117.2)

[a]For PHQ-9, the sample sizes were 150, 131, 114, and 138 at baseline, 1 month, 2 months, and 3 months, respectively. For patient engagement, the sample size remained 150 across the 3-month intervention.

[b]PHQ-9: Patient Health Questionnaire.

[c]Not available.

Unconditional Latent Growth Curve Model

Results of the unconditional model indicated that depressive symptoms decreased during the 3-month intervention, and the largest reduction occurred at 1 month after baseline. This model had good model fit (Figure 1). The mean intercept was 10.157 (P<.001), indicating a mean PHQ-9 score of 10.157 at baseline. The mean slope was –2.336 (P<.001), and the factor loadings on the slope were 0, 1, 1.565, and 1.425 at baseline, 1 month, 2 months, and 3 months, respectively. These results indicated a nonlinear pattern of the reduction of depressive symptoms: a rapid decrease of PHQ-9 score (2.336 points) occurred at the first month, and then the magnitude of reduction flattened out. Variances of the intercept and slope were 15.069 (P=.003) and 5.564 (P=.12), indicating significant individual differences in depressive symptoms at baseline, but no individual differences in the rates of change over time. Covariance between the intercept and slope was –4.381 (P=.25), suggesting that the rate of change in depressive symptoms was independent of its initial levels (Table 3).

XSL•FO
RenderX

Figure 1. Unconditional latent growth curve model for depressive symptoms (n=150). Model fit statistic: χ^2=7.0, *P*=.07, degrees of freedom=3; CFI=0.975, TLI=0.949, SRMR=0.039, RMSEA=0.094. Observed variables were indicated by boxes and latent variables were indicated by ovals. Unidirectional arrows indicated the effect of one variable on another and bidirectional arrows indicated correlations. The nonsignificant path was indicated by dotted line. 0: baseline; 1: 1 month after baseline; 2: 2 months after baseline; 3: 3 months after baseline; **P*<.05, ***P*<.001.

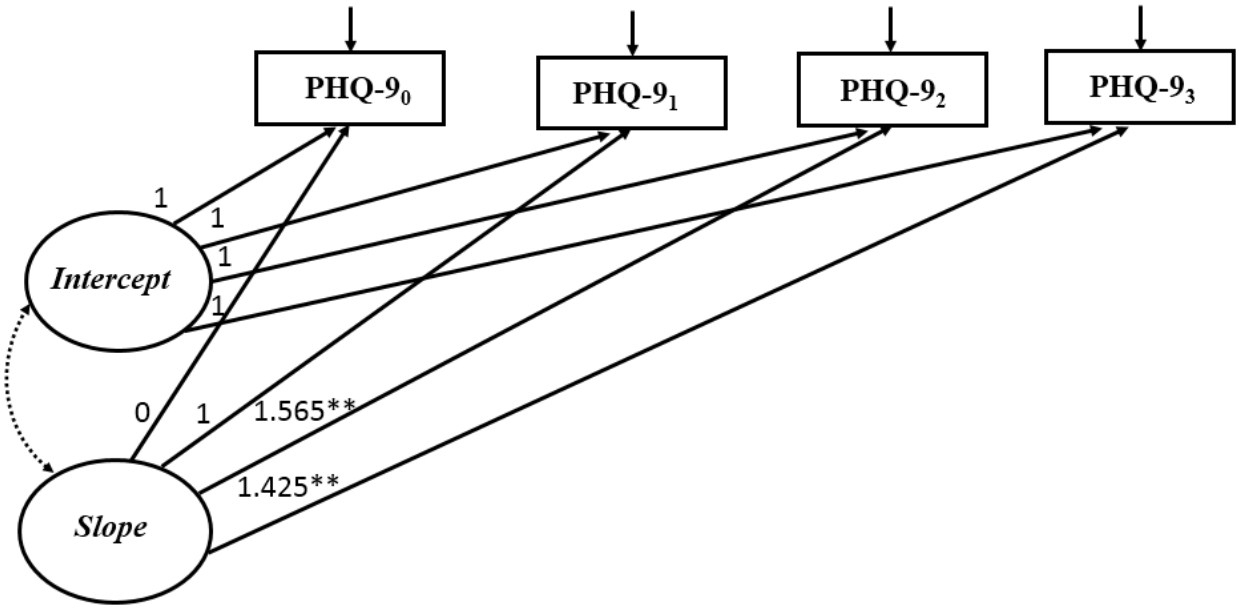


Table 3. Parameter estimates for unconditional latent growth curve model for depressive symptoms (n=150).

Growth parameters	Estimate	SE	*P* value
Slope$_0$→PHQ-9$_0$[a]	0.000	0.000	—[b]
Slope$_1$→PHQ-9$_1$	1.000	0.000	—
Slope$_2$→PHQ-9$_2$	1.565	0.196	<.001
Slope$_3$→PHQ-9$_3$	1.425	0.225	<.001
Mean intercept	10.157	0.370	<.001
Mean slope	–2.336	0.411	<.001
Variance of intercept	15.069	5.100	.003
Variance of slope	5.564	3.614	.12
Covariance of intercept and slope	–4.381	3.799	.25

[a]PHQ-9: Patient Health Questionnaire.

[b]Not available.

Conditional Latent Growth Curve Models

Three conditional LGCMs were constructed, each using one measure of patient engagement as a time-varying variable. Results of the conditional models indicated that higher completion rate and greater frequency of items completed were significantly associated with lower depressive symptoms at 3 months, but there was no significant relationship between time spent on the program and depressive symptoms over the course of the intervention. All conditional models reported good model fit (Figures 2-4).

XSL•FO
RenderX

Figure 2. Conditional latent growth curve model for depressive symptoms with completion rate as the measure of patient engagement (n=150). Model fit statistic: χ^2=13.0, *P*=.79, degrees of freedom=18; CFI=1.000; TLI=1.000; RMSEA<0.001; SRMR=0.030. Baseline characteristics included education, income, and duration of HIV infection. Observed variables were indicated by boxes and latent variables were indicated by ovals. Unidirectional arrows indicated the effect of one variable on another and bidirectional arrows indicated correlations. Nonsignificant paths were indicated by dotted lines. Rate: completion rate; 0: baseline; 1: 1 month after baseline; 2: 2 months after baseline; 3: 3 months after baseline; **P*<.05, ***P*<.001.

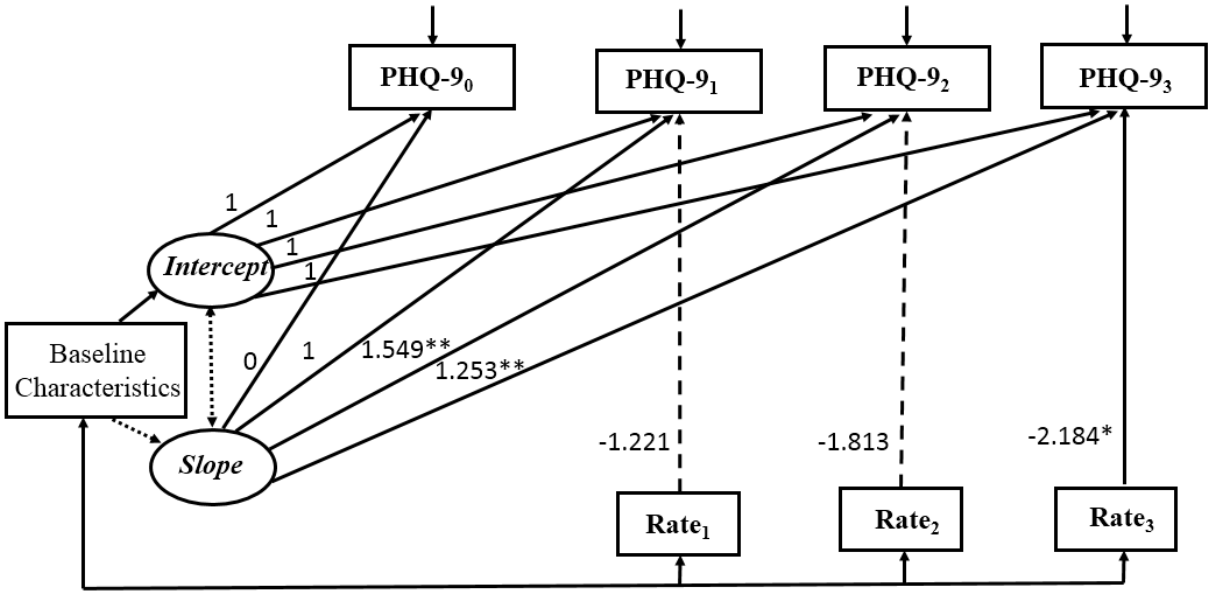
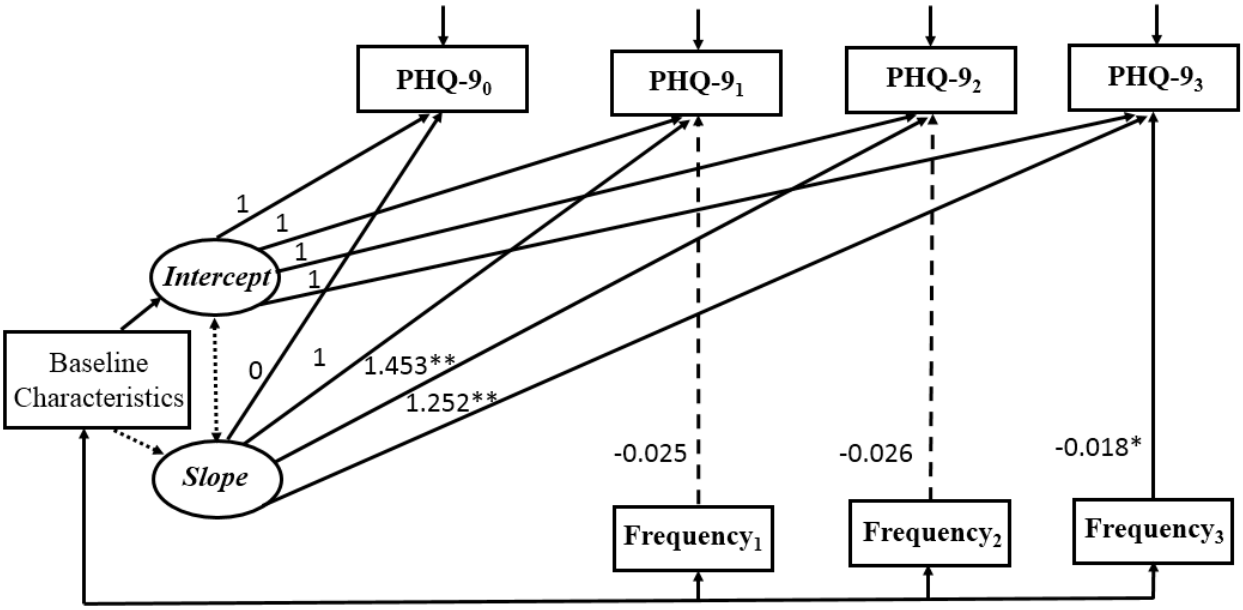
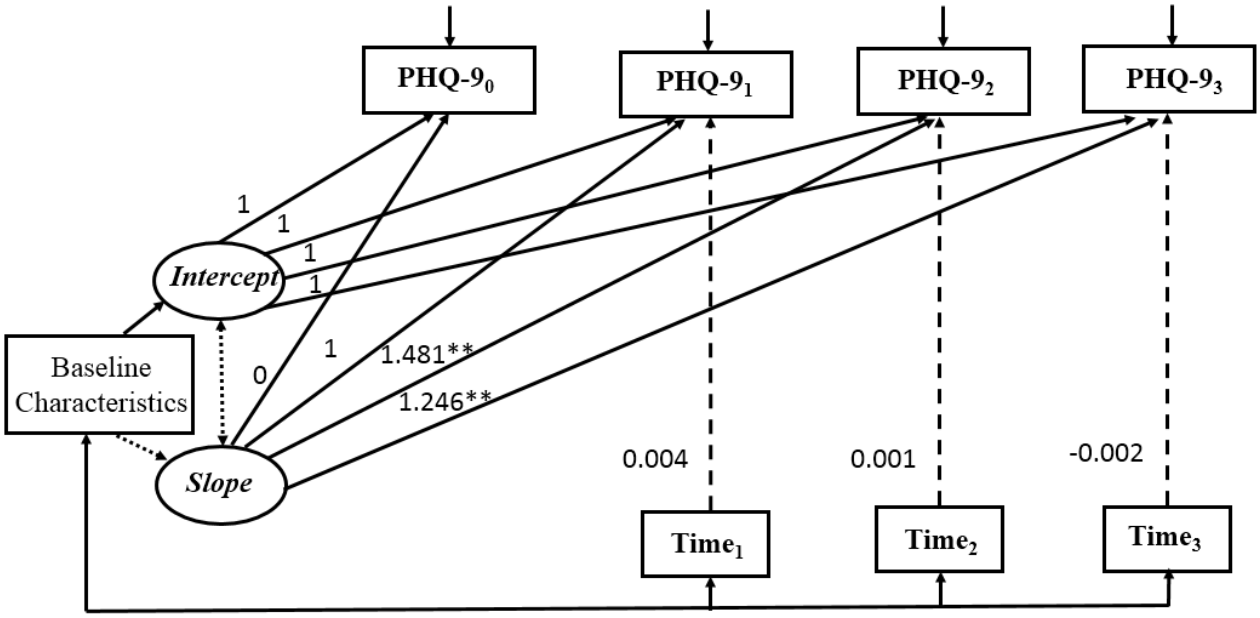


Figure 3. Conditional latent growth curve model for depressive symptoms with frequency of items completed as the measure of patient engagement (n=150). Model fit statistic: χ^2=14.0, *P*=.73, degrees of freedom=18; CFI=1.000; TLI=1.000; RMSEA<0.001; SRMR=0.028. Baseline characteristics included education, income, and duration of HIV infection. Observed variables were indicated by boxes and latent variables were indicated by ovals. Unidirectional arrows indicated the effect of one variable on another and bidirectional arrows indicated correlations. Nonsignificant paths were indicated by dotted lines. Frequency: frequency of items completed; 0: baseline; 1: 1 month after baseline; 2: 2 months after baseline; 3: 3 months after baseline; **P*<.05, ***P*<.001.



XSL•FO
RenderX

Figure 4. Conditional latent growth curve model for depressive symptoms with the time spent on the program as the measure of patient engagement (n=150). Model fit statistic: χ^2=19.6, *P*=.36, degrees of freedom=18; CFI=0.993; TLI=0.988; RMSEA=0.024; SRMR=0.025. Baseline characteristics included education, income, and duration of HIV infection. Observed variables were indicated by boxes and latent variables were indicated by ovals. Unidirectional arrows indicated the effect of one variable on another and bidirectional arrows indicated correlations. Nonsignificant paths were indicated by dotted lines. Time: time spent on the program; 0: baseline; 1: 1 month after baseline; 2: 2 months after baseline; 3: 3 months after baseline; **P*<.05, ***P*<.001.



Based on the literature, baseline characteristics of education, income to cover daily expenses, and duration of HIV infection were included as control variables. As shown in Table 4, education and duration of HIV infection did not have significant effect either on the initial levels (intercept) or the rate of change (slope) of depressive symptoms in all conditional models. Income had a significant effect on the intercept in all conditional models, suggesting that participants with adequate income to cover daily expenses had a lower level of depressive symptoms at baseline. The time-varying variable, patient engagement, had a significant effect on depressive symptoms at 3 months after baseline (completion rate: β3=–2.184, *P*=.048; frequency of items completed: β3=–0.018, *P*=.04), whereas the effect of patient engagement, regardless of the measurement, was not significant on depressive symptoms at 1 or 2 months. Specifically, participants who had higher completion rates or greater frequency of items completed were more likely to report fewer depressive symptoms at 3 months after controlling for the underlying growth trajectory of depressive symptoms and baseline characteristics. By contrast, the time spent on the program was not significantly associated with depressive symptoms throughout the intervention process (β1=0.004, *P*=.29; β2=–0.001, *P*=.74; β3=–0.002, *P*=.16), suggesting that the amount of time patients spent on the intervention program was not related to the changes in depressive symptoms.

XSL•FO
RenderX

Table 4. Parameter estimates for the three conditional latent growth curve models (n=150).

Covariate parameters	Completion rate, estimate (SE)	Frequency of items completed, estimate (SE)	Time spent on program, estimate (SE)
Baseline characteristics			
Education→intercept	–0.893 (0.858)	–0.914 (0.849)	–0.909 (0.843)
Income→intercept	–2.329[b] (1.179)	–2.306[b] (1.175)	–2.309[b] (1.169)
Duration of HIV infection→intercept	0.064 (0.165)	0.058 (0.157)	0.061 (0.156)
Education→slope	–0.257 (0.650)	–0.274 (0.632)	–0.239 (0.628)
Income→slope	1.382 (0.860)	1.298 (0.899)	1.399 (0.898)
Duration of HIV infection→slope	0.103 (0.155)	0.116 (0.147)	0.092 (0.139)
Measures of patient engagement			
$Engagement_1$→$PHQ\text{-}9_1$[a]	–1.221 (0.881)	–0.025 (0.019)	0.004 (0.004)
$Engagement_2$→$PHQ\text{-}9_2$	–1.813 (1.158)	–0.026 (0.014)	0.001 (0.002)
$Engagement_3$→$PHQ\text{-}9_3$	–2.184[b] (0.945)	–0.018[b] (0.009)	–0.002 (–0.082)

[a]PHQ-9: Patient Health Questionnaire.

[b]$P<.05$.

Discussion

Principal Findings

This study was among the first to explore dose-response relationship between patient engagement and depressive symptoms in a social media–based mHealth intervention among people living with HIV using longitudinal data. Three measures of patient engagement were used to assess different aspects of the construct. By investigating the time-varying relationship between patient engagement and health outcomes, our study provided a better understanding of the dose-response relationship in mHealth interventions and the time of occurrence of this association. Results of LGCMs revealed that two measures of patient engagement, completion rate and frequency of items completed, were significantly associated with reduced depressive symptoms, and these relationships occurred at 3 months of the intervention but not at 1 or 2 months. However, time spent on the program was not significantly related to depressive symptoms throughout the intervention.

First and foremost, this study provides new evidence on the dose-response relationship in mHealth interventions as the literature in this regard was limited and inconclusive [39]. Potential reasons for the ambivalent relationship in previous mHealth intervention studies could include inadequate research design or implementation, inadequate intervention durations for significant treatment effect, low levels of patient engagement, high dropout rate, and/or inappropriate or inadequate measurements of patient engagement or data analysis methods [12,20,39]. This study, however, shows that an evidence-based mHealth intervention with rigorous design and implementation could produce significant treatment effects, in particular, for participants with better engagement.

Our data suggests that the significant relationship of patient engagement and health outcomes did not occur early (1 or 2 months) in the intervention, but only at 3 months. Few studies have explored the dose-response relationship across the span of an intervention, especially in mHealth programs. To our knowledge, only one mHealth study intended to explore the patient engagement–outcome relationship using multiple measures of patient engagement [13]. However, the web-based iCBT intervention conducted in a community setting in Dublin, Ireland, in 2014 only measured pre- and postintervention depressive symptoms [13]. The methodology of repeated-measures analysis of variance used in the study could only reveal that patients with significant improvement in depressive symptoms had higher levels of patient engagement during the 8-week intervention [13]. Although the study further found that patients tended to engage in the online program more in the first half of the program (1 to 4 weeks) than in the latter half (5 to 8 weeks) [13], the design and findings of the study could not identify when the linkage between patient engagement and outcome occurred during the course of the intervention.

Compared with mHealth interventions, face-to-face studies have focused more on the relationship between patient engagement and health outcomes. One face-to-face intervention investigated the dose-response relationship in the first 5 weeks of the intervention using both weekly engagement measures and depressive symptoms outcomes among people with major depressive disorder [22]. However, the study used repeated measures regression analysis that could only conclude that there was a significant dose-response relationship during the 5-week intervention but not in which week(s) the relationship started to appear. Another face-to-face study conducted among patients with diabetes in Netherlands investigated the dose-response relationship between patient engagement and depressive symptoms during an 8-week intervention [40]. Patient engagement and depressive symptoms were repeatedly measured at 2, 4, and 8 weeks [40]. Although the study used multilevel analysis that was able to identify both the dose-response relationship and timing of the occurrence of any relationship, the study did not find a dose-response relationship in any of the

XSL•FO
RenderX

time points [40]. One potential explanation for the nonsignificant dose-response relationship might be that individuals with chronic diseases and depressive symptoms need a longer time of exposure to an intervention to obtain the beneficial outcomes. Likewise, our study did not find a dose-response relationship at 1 or 2 months of the intervention but only at 3 months of the intervention among people living with HIV and depressive symptoms. Such a relationship among individuals with chronic diseases and psychological disorders needs to be further explored in future studies. Although the nonsignificant dose-response relationship at 1 or 2 months seemed to be contradictory to the finding that most reduction of depressive symptoms occurred at 1 month, reasons for the reduction of depressive symptoms could be many. Besides patient engagement, other factors such as social support might also play a role in the reduction of depressive symptoms. For example, several phone calls made by research staff to motivate participants to continue the intervention at 1 week and 1 and 2 months after baseline may provide social support needed by people living with HIV and elevated depressive symptoms.

Identifying the time point of effective dose-response relationship between patient engagement and treatment outcomes is critical for optimizing the design, pace, and duration of interventions. Unnecessarily lengthy interventions waste resources and lead to high dropout rates [41], whereas interventions with inadequate duration may not generate intended outcomes [42]. Repeatedly examining the relationship between patient engagement and health outcomes allows us to identify the optimal time for intervention. This study demonstrates such measurement in an mHealth intervention.

In addition to identifying the timing of occurrence of the dose-relationship in the intervention, this study also found the differential relationships between different measures of patient engagement and health outcomes. Compared with face-to-face interventions, mHealth interventions have more measurement tools to capture different aspects of patient engagement [12]. These different measurements may not have similar impacts on patient outcomes and may explain the conflicting results regarding the dose-response relationships in mHealth interventions [12].

In this study, participants with higher completion rates reported fewer depressive symptoms at 3 months. On average, participants completed 50.8% (33/65) of the program items (ie, short articles, audio recordings, and posters) at 3 months, which was similar to the completion rates reported in other online interventions for depressive symptoms [10,13]. A recent meta-analysis found a significant association between higher completion rates and reduced depressive symptoms in online programs [43]. However, some more recent internet-based RCTs did not find such a relationship [19,20]. One potential reason for the nonsignificant relationship was low levels of patient engagement in these mHealth interventions [10]. Therefore, adherence to the program needs to be underscored in mHealth interventions, especially considering that most of the existing mHealth interventions have relatively low completion rates [10,13].

Frequency of items completed was another patient engagement measure we used in the study, and it was also positively associated with patients' outcomes. Such a measure of patient engagement has rarely been reported in the existing mHealth interventions. At 3 months of the intervention, participants had read or listened to the items 47.5 times on average. As the average number of items completed by participants was 33, each item was repeatedly read or listened to for approximately 1.5 times (47.5/33). This measure captured an important aspect of patient engagement as it reflected to what extent participants were willing to repeatedly engage with the intervention items. Based on this frequency data, we subsequently chose items that had been most read or listened to and re-sent these items to the patients in a booster session 3 months after the intervention [8]. In a separate study, we found a continued intervention effect with further reductions of patients' depressive symptoms at 3 and 6 months after the intervention [8].

In contrast, the third patient engagement measure, time spent on the program, was found not to relate to the reduced depressive symptoms, which was consistent with previous studies [21,44]. Time spent on the program may not reflect the actual levels of patient engagement as it is likely to be influenced by a variety of factors such as reading speed, cognitive capacity, and familiarity with the program [39]. Participants might keep the program on without actively engaging. For example, patients might fall asleep when listening to a relaxation audio recording; yet time stayed on the page was still counted as time spent on the program. In this study, there was a noticeable difference (31.3 vs 96.7 minutes) between the median and mean time spent on the platform at 3 months of the intervention, indicating a positively skewed distribution of patients' time spent online, suggesting few patients spent extra-long time on the platform. Future interventions should incorporate a time-out function to better capture the actual patient engagement.

Limitations

Several limitations in this study should be noted. First, although the sample was drawn from a large hospital designated for HIV treatment in Guangzhou with more than 10,000 HIV seropositive patients on treatment, participants were mostly from an urban setting and the majority of them were nonheterosexual young men. Thus, the generalizability of the findings to other people living with HIV or other populations may be limited. Second, although depressive symptoms are the major outcome in this study, more types of outcome measures such as other psychological indicators (eg, anxiety) and/or biomarkers (eg, hair or salivary cortisol concentration) can be incorporated in future studies to provide further evidence on dose-response relationship between patient engagement and intervention outcomes. Third, while measures of patient engagement were collected objectively through the Run4Love platform, depressive symptoms were self-reported and might suffer from social desirability and recall bias. Fourth, there might be measurement biases in assessing patient engagement. For example, when clicked, a material was readily considered as completed. Nevertheless, this definition of completion was widely used in mHealth interventions, and the relevant patient engagement measures served as reliable and efficient metrics to assess dose-response relationship [12]. Additionally, time spent on the

XSL•FO
RenderX

program might not capture the real patient engagement as it was not able to differentiate between active and inactive engagement and time on distraction or interruption could not be measured. Finally, we did not differentiate patient engagement in different formats of the intervention items (ie, short articles, audio recordings, and posters) nor did we examine the relationship between the content of the Run4Love intervention and patient engagement in this study, as it goes beyond the purposes of this study. We call for future research to further explore the relationships between patient engagement and the content of the intervention as well as the potential differences in patient engagement by different types of multimedia items in mHealth interventions.

Implications

This study indicated that higher completion rate or greater frequency of accessing items in an mHealth intervention contributed to reduced depressive symptoms among people living with HIV and elevated depressive symptoms at 3 months in the intervention. Given the observed positive impact of patient engagement on intervention outcomes, adherence to mHealth interventions should be evaluated and strengthened as a critical and integral strategy to improve treatment effects. Approaches for promoting patient engagement include enhancing patients' initial motivation in program participation, education on the importance of active participation, frequent and effective reminders, financial and verbal incentives, and engaging patients in goal setting (eg, physical activities) [45-47]. Future studies with repeated assessments of patient engagement and outcomes throughout the intervention process have the potential to analyze the time-varying dose-response relationship. Such studies would optimize the intervention duration and maximize intervention outcomes through improving the design and implementation of mHealth interventions.

Conclusion

In conclusion, this study revealed a positive association between patient engagement and reduction in depressive symptoms at 3 months of an mHealth intervention using latent growth curve models and repeated measure data from 4 time points. The results underscore the importance of enhancing patient engagement in mHealth interventions to improve health outcomes of the participants.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant No 71573290) and China Medical Board open competition funding (Grant No. 17-271). The funding sources had no role in the study design, data collection, statistical analyses, and writing of the manuscript.

Authors' Contributions

YG and YZ had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. YG, YAH, and YZ were responsible for the concept and design of the study. YZ, YL, MZ, HZ, CZ, SW, CP, and YH were responsible for the acquisition, analysis, or interpretation of data. YZ and YG drafted the manuscript. YG, YZ, YAH, AMW, YH, and RTHH completed the critical revision of the manuscript for important intellectual content. LL, WC, CL, and YH were responsible for administrative, technical, and material support, and YG supervised the study.

Conflicts of Interest

None declared.

References

1. McManus RJ, Mant J, Bray EP, Holder R, Jones MI, Greenfield S, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): a randomised controlled trial. Lancet 2010 Jul 17;376(9736):163-172. [doi: 10.1016/S0140-6736(10)60964-6] [Medline: 20619448]
2. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The Impact of Information Technology on Patient Engagement and Health Behavior Change: A Systematic Review of the Literature. JMIR Med Inform 2016;4(1):e1 [FREE Full text] [doi: 10.2196/medinform.4514] [Medline: 26795082]
3. Ben-Zeev D, Scherer EA, Gottlieb JD, Rotondi AJ, Brunette MF, Achtyes ED, et al. mHealth for schizophrenia: patient engagement with a mobile phone intervention following hospital discharge. JMIR Ment Health 2016;3(3):e34 [FREE Full text] [doi: 10.2196/mental.6348] [Medline: 27465803]
4. Cook DJ, Manning DM, Holland DE, Prinsen SK, Rudzik SD, Roger VL, et al. Patient engagement and reported outcomes in surgical recovery: effectiveness of an e-health platform. J Am Coll Surg 2013 Oct;217(4):648-655. [doi: 10.1016/j.jamcollsurg.2013.05.003] [Medline: 23891066]
5. van Luenen S, Garnefski N, Spinhoven P, Kraaij V. Guided internet-based intervention for people with HIV and depressive symptoms: a randomised controlled trial in the Netherlands. Lancet HIV 2018 Dec;5(9):e488-e497. [doi: 10.1016/S2352-3018(18)30133-4] [Medline: 30135045]
6. Cheng LJ, Kumar PA, Wong SN, Lau Y. Technology-delivered psychotherapeutic interventions in improving depressive symptoms among people with HIV/AIDS: a systematic review and meta-analysis of randomised controlled trials. AIDS Behav 2020 Jun;24(6):1663-1675. [doi: 10.1007/s10461-019-02691-6] [Medline: 31587115]

XSL•FO
RenderX

7. Karyotaki E. Internet-based interventions for people with HIV and depression. Lancet HIV 2018 Dec;5(9):e474-e475. [doi: 10.1016/S2352-3018(18)30172-3] [Medline: 30135044]
8. Guo Y, Hong YA, Cai W, Li L, Hao Y, Qiao J, et al. Effect of a WeChat-based intervention (Run4Love) on depressive symptoms among people living with HIV in China: a randomized controlled trial. J Med Internet Res 2020 Feb 09;22(2):e16715 [FREE Full text] [doi: 10.2196/16715] [Medline: 32044751]
9. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
10. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
11. Mausbach BT, Moore R, Roesch S, Cardenas V, Patterson TL. The relationship between homework compliance and therapy outcomes: an updated meta-analysis. Cognit Ther Res 2010 Oct;34(5):429-438 [FREE Full text] [doi: 10.1007/s10608-010-9297-z] [Medline: 20930925]
12. Beintner I, Vollert B, Zarski A, Bolinski F, Musiat P, Görlich D, et al. Adherence reporting in randomized controlled trials examining manualized multisession online interventions: systematic review of practices and proposal for reporting standards. J Med Internet Res 2019 Aug 15;21(8):e14181 [FREE Full text] [doi: 10.2196/14181] [Medline: 31414664]
13. Enrique A, Palacios JE, Ryan H, Richards D. Exploring the relationship between usage and outcomes of an internet-based intervention for individuals with depressive symptoms: secondary analysis of data from a randomized controlled trial. J Med Internet Res 2019 Aug 01;21(8):e12775 [FREE Full text] [doi: 10.2196/12775] [Medline: 31373272]
14. Couper MP, Alexander GL, Zhang N, Little RJA, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res 2010;12(4):e52 [FREE Full text] [doi: 10.2196/jmir.1430] [Medline: 21087922]
15. Mattila E, Lappalainen R, Välkkynen P, Sairanen E, Lappalainen P, Karhunen L, et al. Usage and dose response of a mobile acceptance and commitment therapy app: secondary analysis of the intervention arm of a randomized controlled trial. JMIR Mhealth Uhealth 2016 Jul 28;4(3):e90 [FREE Full text] [doi: 10.2196/mhealth.5241] [Medline: 27468653]
16. Carroll KM, Ball SA, Martino S, Nich C, Babuscio TA, Nuro KF, et al. Computer-assisted delivery of cognitive-behavioral therapy for addiction: a randomized trial of CBT4CBT. Am J Psychiatry 2008 Jul;165(7):881-888 [FREE Full text] [doi: 10.1176/appi.ajp.2008.07111835] [Medline: 18450927]
17. Hadgkiss EJ, Jelinek GA, Taylor KL, Marck CH, van der Meer DM, Pereira NG, et al. Engagement in a program promoting lifestyle modification is associated with better patient-reported outcomes for people with MS. Neurol Sci 2015 Jun;36(6):845-852 [FREE Full text] [doi: 10.1007/s10072-015-2089-1] [Medline: 25638416]
18. Van Gemert-Pijnen JE, Kelders SM, Bohlmeijer ET. Understanding the usage of content in a mental health intervention for depression: an analysis of log data. J Med Internet Res 2014;16(1):e27 [FREE Full text] [doi: 10.2196/jmir.2991] [Medline: 24486914]
19. Geramita EM, Herbeck Belnap B, Abebe KZ, Rothenberger SD, Rotondi AJ, Rollman BL. The association between increased levels of patient engagement with an internet support group and improved mental health outcomes at 6-month follow-up: post-hoc analyses from a randomized controlled trial. J Med Internet Res 2018 Jul 17;20(7):e10402 [FREE Full text] [doi: 10.2196/10402] [Medline: 30021711]
20. Fuhr K, Schröder J, Berger T, Moritz S, Meyer B, Lutz W, et al. The association between adherence and outcome in an Internet intervention for depression. J Affect Disord 2018 Dec 15;229:443-449. [doi: 10.1016/j.jad.2017.12.028] [Medline: 29331706]
21. Donkin L, Hickie IB, Christensen H, Naismith SL, Neal B, Cockayne NL, et al. Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: randomized controlled trial. J Med Internet Res 2013 Oct;15(10):e231 [FREE Full text] [doi: 10.2196/jmir.2771] [Medline: 24135213]
22. Conklin LR, Strunk DR. A session-to-session examination of homework engagement in cognitive therapy for depression: do patients experience immediate benefits? Behav Res Ther 2015 Sep;72:56-62 [FREE Full text] [doi: 10.1016/j.brat.2015.06.011] [Medline: 26183022]
23. van Straten A, Cuijpers P, Smits N. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 2008;10(1):e7 [FREE Full text] [doi: 10.2196/jmir.954] [Medline: 18364344]
24. Guo Y, Hong YA, Qiao J, Xu Z, Zhang H, Zeng C, et al. Run4Love, a mHealth (WeChat-based) intervention to improve mental health of people living with HIV: a randomized controlled trial protocol. BMC Public Health 2018 Dec 26;18(1):793 [FREE Full text] [doi: 10.1186/s12889-018-5693-1] [Medline: 29940921]
25. WeChat. Wikipedia. URL: https://zh.wikipedia.org/wiki/%E5%BE%AE%E4%BF%A1 [accessed 2020-05-30]
26. Guangzhou Population 2020. World Population Review. URL: https://worldpopulationreview.com/world-cities/guangzhou-population/ [accessed 2020-05-30]
27. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry 2014;36(5):539-544. [doi: 10.1016/j.genhosppsych.2014.05.021] [Medline: 25023953]

XSL•FO
RenderX

28. Sun XY, Li YX, Yu CQ, Li LM. [Reliability and validity of depression scales of Chinese version: a systematic review]. Zhonghua Liu Xing Bing Xue Za Zhi 2017 Jan 10;38(1):110-116. [doi: 10.3760/cma.j.issn.0254-6450.2017.01.021] [Medline: 28100388]
29. Chen S, Fang Y, Chiu H, Fan H, Jin T, Conwell Y. Validation of the nine-item Patient Health Questionnaire to screen for major depression in a Chinese primary care population. Asia Pac Psychiatry 2013 Jun;5(2):61-68. [doi: 10.1111/appy.12063] [Medline: 23857806]
30. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
31. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
32. Bollen KA, Curran PJ. Latent Curve Models: A Structural Equation Perspective. Hoboken: John Wiley & Sons; 2006.
33. Chen W, Shiu C, Yang JP, Simoni JM, Fredriksen-Goldsen KI, Lee TS, et al. Antiretroviral therapy (ART) side effect impacted on quality of life, and depressive symptomatology: a mixed-method study. J AIDS Clin Res 2013 Jun 29;4:218 [FREE Full text] [doi: 10.4172/2155-6113.1000218] [Medline: 24083060]
34. Remien RH, Exner T, Kertzner RM, Ehrhardt AA, Rotheram-Borus MJ, Johnson MO, NIMH Healthy Living Project Trial Group. Depressive symptomatology among HIV-positive women in the era of HAART: a stress and coping model. Am J Community Psychol 2006 Dec;38(3-4):275-285. [doi: 10.1007/s10464-006-9083-y] [Medline: 16967343]
35. van Servellen G, Aguirre M, Sarna L, Brecht M. Differential predictors of emotional distress in HIV-infected men and women. West J Nurs Res 2002 Feb;24(1):49-72. [doi: 10.1177/019394590202400105] [Medline: 11829273]
36. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociol Method Res 2016 Jun 29;21(2):230-258. [doi: 10.1177/0049124192021002005]
37. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equation Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
38. Bentler PM. Fit indexes, lagrange multipliers, constraint changes and incomplete data in structural models. Multivariate Behav Res 1990 Apr 01;25(2):163-172. [doi: 10.1207/s15327906mbr2502_3] [Medline: 26794478]
39. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
40. Snippe E, Schroevers MJ, Annika Tovote K, Sanderman R, Emmelkamp PMG, Fleer J. Patients' outcome expectations matter in psychological interventions for patients with diabetes and comorbid depressive symptoms. Cognit Ther Res 2015;39(3):307-317 [FREE Full text] [doi: 10.1007/s10608-014-9667-z] [Medline: 25983355]
41. Christensen H, Griffiths KM, Mackinnon AJ, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. Psychol Med 2006 Dec;36(12):1737-1746. [doi: 10.1017/S0033291706008695] [Medline: 16938144]
42. Chu BC, Kendall PC. Positive association of child involvement and treatment outcome within a manual-based cognitive-behavioral treatment for children with anxiety. J Consult Clin Psychol 2004 Oct;72(5):821-829. [doi: 10.1037/0022-006X.72.5.821] [Medline: 15482040]
43. Kazantzis N, Whittington C, Zelencich L, Kyrios M, Norton PJ, Hofmann SG. Quantity and quality of homework compliance: a meta-analysis of relations with outcome in cognitive behavior therapy. Behav Ther 2016 Sep;47(5):755-772. [doi: 10.1016/j.beth.2016.05.002] [Medline: 27816086]
44. Clarke G, Kelleher C, Hornbrook M, Debar L, Dickerson J, Gullion C. Randomized effectiveness trial of an Internet, pure self-help, cognitive behavioral intervention for depressive symptoms in young adults. Cogn Behav Ther 2009;38(4):222-234 [FREE Full text] [doi: 10.1080/16506070802675353] [Medline: 19440896]
45. Carroll SL, Embuldeniya G, Abelson J, McGillion M, Berkesse A, Healey JS. Questioning patient engagement: research scientists' perceptions of the challenges of patient engagement in a cardiovascular research network. Patient Prefer Adherence 2017;11:1573-1583 [FREE Full text] [doi: 10.2147/PPA.S135457] [Medline: 28979105]
46. Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. J Med Internet Res 2005;7(2):e16 [FREE Full text] [doi: 10.2196/jmir.7.2.e16] [Medline: 15998607]
47. Loh A, Leonhart R, Wills CE, Simon D, Härter M. The impact of patient participation on adherence and clinical outcome in primary care of depression. Patient Educ Couns 2007 Jan;65(1):69-78. [doi: 10.1016/j.pec.2006.05.007] [Medline: 17141112]

Abbreviations

CBSM: cognitive-behavioral stress management course
CES-D: Center for Epidemiological Studies Depression Scale
CFI: confirmatory fit index

XSL•FO
RenderX

IQR: interquartile range
LGCM: latent growth curve model
mHealth: mobile Health
MLR: maximum likelihood estimation with robust standard error
PHQ-9: Patient Health Questionnaire
RCT: randomized controlled trial
RMSEA: root mean-squared error of approximation
SRMR: standardized root mean of residual
TLI: Tucker-Lewis index

Edited by G Eysenbach; submitted 10.07.20; peer-reviewed by E White, S van Luenen; comments to author 19.08.20; revised version received 30.08.20; accepted 22.09.20; published 29.10.20.

Please cite as:
Zeng Y, Guo Y, Li L, Hong YA, Li Y, Zhu M, Zeng C, Zhang H, Cai W, Liu C, Wu S, Chi P, Monroe-Wise A, Hao Y, Ho RTH
Relationship Between Patient Engagement and Depressive Symptoms Among People Living With HIV in a Mobile Health Intervention: Secondary Analysis of a Randomized Controlled Trial
JMIR Mhealth Uhealth 2020;8(10):e20847
URL: http://mhealth.jmir.org/2020/10/e20847/
doi:10.2196/20847
PMID:33118956

©Yu Zeng, Yan Guo, Linghua Li, Y Alicia Hong, Yiran Li, Mengting Zhu, Chengbo Zeng, Hanxi Zhang, Weiping Cai, Cong Liu, Shaomin Wu, Peilian Chi, Aliza Monroe-Wise, Yuantao Hao, Rainbow Tin Hung Ho. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 29.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Mobile App, KhunLook, to Support Thai Parents and Caregivers With Child Health Supervision: Development, Validation, and Acceptability Study

Rosawan Areemit[1], MSc, MD; Pagakrong Lumbiganon[1], MD; Chanyut Suphakunpinyo[1], MD; Arunee Jetsrisuparb[1], MD; Sumitr Sutra[1], MD; Kunwadee Sripanidkulchai[2], PhD

[1]Department of Pediatrics, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

[2]Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand

Corresponding Author:
Kunwadee Sripanidkulchai, PhD
Department of Computer Engineering, Faculty of Engineering
Chulalongkorn University
Phayatai Road, Wang Mai, Pathumwan
Bangkok, 10330
Thailand
Phone: 66 2 218 6998
Email: kunwadee@cp.eng.chula.ac.th

Abstract

Background: In Thailand, children born in government hospitals receive a maternal and child health handbook (MCHH). However, when a new MCHH edition is released, those with the previous editions do not have access to the updated information. A mobile app is an appealing platform to fill this gap. We developed a mobile app called “KhunLook” as an interactive electronic MCHH intended to assist parents in child health supervision.

Objective: This study describes the user requirements and development of the KhunLook mobile app, validity of parents’ growth assessments, and parents’ evaluation of feasibility and acceptability of the app.

Methods: Phase 1 was a qualitative study using individual interviews. The interview data were used to revise the prototype. In phase 2, parents were randomly assigned to assess their children’s growth with the app or the MCHH. The outcomes were compared to those of the physician’s assessment, and congruence was determined. In phase 3, parents evaluated the feasibility and acceptability of the app in comparison to the MCHH through a web-based survey.

Results: Four health care providers and 8 parents participated in phase 1. Two themes were identified: (1) the mobile app potentially counters parents’ infrequent use of the MCHH with accuracy, attractiveness, convenience, and simplicity, and (2) health supervision needs to be standard, up-to-date, and understandable. KhunLook was publicly launched with a family page and 7 key features: growth and nutrition, development, immunizations, oral health, reminders for the next appointment, memories, and health advice. In phase 2, 56 parents participated in the growth parameter assessments; 34 were in the App group and 22 in the MCHH group. The outcomes of the growth parameter assessments between parents and physicians in both the App and MCHH groups were not significantly different. The congruence proportions were higher in the App group for weight and head circumference, but the differences were not statistically significant. In phase 3, 356 parents from all over Thailand participated in a web-based survey. Parents rated the app feasibility as “very easy to easy” to use at higher proportions than the MCHH in all health assessment domains (growth, development, and immunizations) and ease-of-use domains with statistical significance (P<.001). The KhunLook app received a significantly higher mean score (8.59/10) than the MCHH (7.6/10) (P<.001). Most parents (317/356, 89.0%) preferred the app over MCHH. Further, 93.5% (333/356) of the parents stated that they would continue to use the app and 96.9% (345/356) would recommend others to use it.

Conclusions: KhunLook, a Thai mobile app for child health supervision, was developed, validated for growth assessments, and was well accepted for ease-of-use by parents. Further studies should be conducted with a large scale of users, and the impact of this app on health behaviors and health outcomes must be evaluated.

(JMIR Mhealth Uhealth 2020;8(10):e15116) doi:10.2196/15116

XSL•FO
RenderX

KEYWORDS

mobile app; mHealth; KhunLook; child health supervision; maternal and child health handbook; feasibility; acceptability; Thailand; mobile phone

Introduction

Child health supervision is critical to the development of the child, family, community, and the future of populations. Periodic well-child visits foster strong relationships between the health care provider and the child and family, enabling provisioning of preventive measures such as appropriate surveillance, screening, anticipatory guidance, counseling, and immunizations [1,2]. For children, parents and primary caregivers are fundamental for the early detection of abnormalities as well as promotion of life-long practices that support their health. "Parents" in this paper refer to all primary caregivers. In Thailand, the maternal and child health handbook (MCHH) is the medium to facilitate preventive health care communication between parents and the child's health care provider [3-5].

The MCHH was developed and provided to pregnant women at antenatal visits and for children who were born in government hospitals since 1985 [3]. It provides relevant information and serves as the standard hard copy of personal prenatal, natal, postnatal, and child health supervision record for children up to 6 years of age [3]. Only a minority of the children will use a similar book provided by hospitals in private sectors. However, health records from different sectors (ie, government hospitals, private hospitals, or clinics) are not integrated. Often, the MCHH serves as the sole continuous health supervision record.

While the MCHH is widely used and well-accepted, several studies have identified its shortcomings. The MCHH is revised intermittently to include up-to-date growth curve standards, relevant information, and changes in the country's expanded immunization program [3]. However, parents who received previous editions do not have access to updated information, which could potentially cause conflicting messages. The MCHH has growth curves with instructions for parents to keep track of their child's growth. Although this is practical for monitoring the health status, this part is seldomly used [3,6]. Health care providers and parents have voiced the difficulty and confusion in plotting growth curves [7-9]. Studies suggest the MCHH could be improved by including more detailed and graphical information and that it should be made from more durable material [3,10]. Printed books also carry substantial publishing and distribution costs. For the best care, parents must remember to bring the MCHH to well-child visits. Lastly, the book can be lost because of personal chaos or natural disasters. To this end, mobile health (mHealth) technology can be used to overcome the disadvantages of MCHH.

mHealth technology is defined as "medical and public health practice supported by mobile devices," which includes delivering health care services and useful information to patients, family members, and health care providers through the use of mobile devices and communication [11]. Internet and mobile phone usage have proliferated in Thailand. The use of smartphones in Thailand increased from 26.4% in 2014 to 37.9% in 2015 and 72.3% in 2017 [11-13]. Mobile apps are conveniently distributed and downloadable to the smartphone. Parents are active users of apps. Studies have shown that parents find information on apps to support them on child health and parenting issues [14].

A mobile app is an appealing platform for the development of an interactive electronic MCHH because it addresses many of the printed book's disadvantages [15]. Growth parameters can be automatically plotted, and percentiles can be computed and initially interpreted. Developmental milestones and immunization records can be tracked and assessed. Loss of information can be prevented through data backups. Moreover, updates in health supervision can be automatically incorporated through the app.

In this study, we developed and evaluated an interactive electronic MCHH mobile app called "KhunLook," which translates to "my dear child" to assist Thai parents and health care providers with child health supervision. The validity of the growth assessments and user acceptability were also examined.

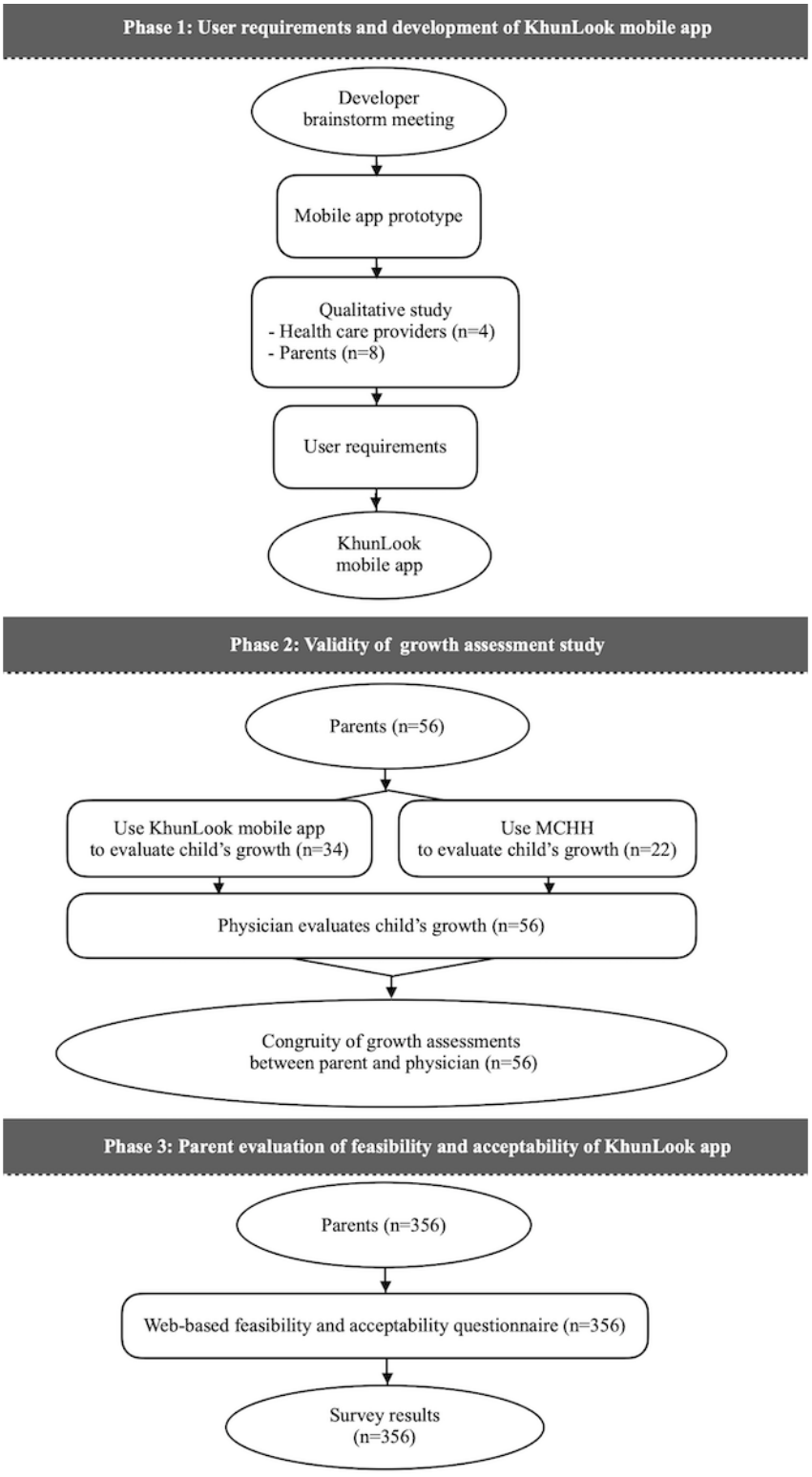
Methods

Overview

This study was approved by the Khon Kaen University Ethics Committee for Human Research. This study consisted of 3 phases: (1) phase 1, understanding user requirements and development of the KhunLook mobile app, conducted during 2013-2015, (2) phase 2, validation of the growth assessment study, conducted in 2015, and (3) phase 3, parent evaluation of the feasibility and acceptability of the KhunLook app, conducted during 2017 (Figure 1). Phases 1 and 2 were conducted at the Faculty of Medicine, Khon Kaen University, Thailand, and phase 3 was conducted through a web-based survey platform.

XSL•FO
RenderX

Figure 1. An overview of the methodology of the study. MCHH: maternal and child health handbook.



Recruitment

Posters were used to publicize the study. Participation in the study was voluntary. Confidentiality was assured and written or electronic consent was obtained.

Phase 1: User Requirements and Development of the KhunLook Mobile App

App Prototype

A mobile app wireframe prototype was developed based on the major requirements from a brainstorm meeting with pediatricians, dentists, and app developers (n=12). The app prototype had sections pertinent to child health supervision, including a family page, birth history, growth, development,

XSL•FO
RenderX

immunizations, oral care, reminders for the next visit appointment, memories, and health advice. A panel of experts in pediatric subspecialties related to each content domain selected evidence-based and culturally appropriate content that is in accordance with the latest standards and written in a brief and comprehensible manner. The content was cross-checked for content appropriateness and understandability by 2 other subspecialists who were nationally recognized in their field, and adjustments were made until consensus was reached. The content was then incorporated into the app.

User Requirements

The objective was to understand user requirements in order to maximize the benefits and usage of the mobile app. The phase 1 study design was as follows. We used a qualitative design to explore participants' perspectives by asking the question, "What are your perspectives of using a mobile app in lieu of the hard copy of MCHH?" The participants were then provided with the prototype on a tablet. They were encouraged to use and explore the prototype by themselves and then provide comments. The interviews were conducted by a trained research assistant who was not involved in the development of the app. An investigator (KS) and 2 trained research assistants who were the mobile app developers observed the process, took field notes, and assisted the participants as requested. The participants were then asked the next question, "What are your comments and additional requirements for the app?"

Sampling

We used purposive sampling—a nonprobability selection of participants—based on our criteria of interest for richness of data. We targeted health care professionals who worked with children and parents with children under the age of 6 years, who use smartphones and an MCHH, to obtain a wide range of perspectives [16].

Data Collection

We used individual in-depth interviews and field notes to collect data. Interviews were audio-recorded and transcribed verbatim, except for personal identifiers, which were removed [16]. The average health care provider interview lasted for 37.5 minutes and the average parent interview lasted for 38.8 minutes. In total, we had 7.2 hours of interview data and field notes that were analyzed.

Data Analysis

Data collection and analysis occurred concurrently. Researchers RA and KS read the initial 3 transcripts to develop the codes and coding categories. Codes and coding categories were discussed, compared, and clarified for the development of the initial coding schemes. After reaching consensus on the coding scheme, further transcripts were analyzed individually, followed by team discussions and consensus. Data collection continued until further themes failed to emerge. The last interview did not introduce new concepts, suggesting that theoretical saturation was achieved. After the transcripts were coded, related codes were grouped into categories and used to generate key themes. Themes and their relationships were discussed and linked to construct an understanding and address the research questions. An updated version of the KhunLook mobile app based on phase 1 input was publicly launched and made available for users to download without charge in January 2015.

Phase 2: Validity of the Growth Assessment Study

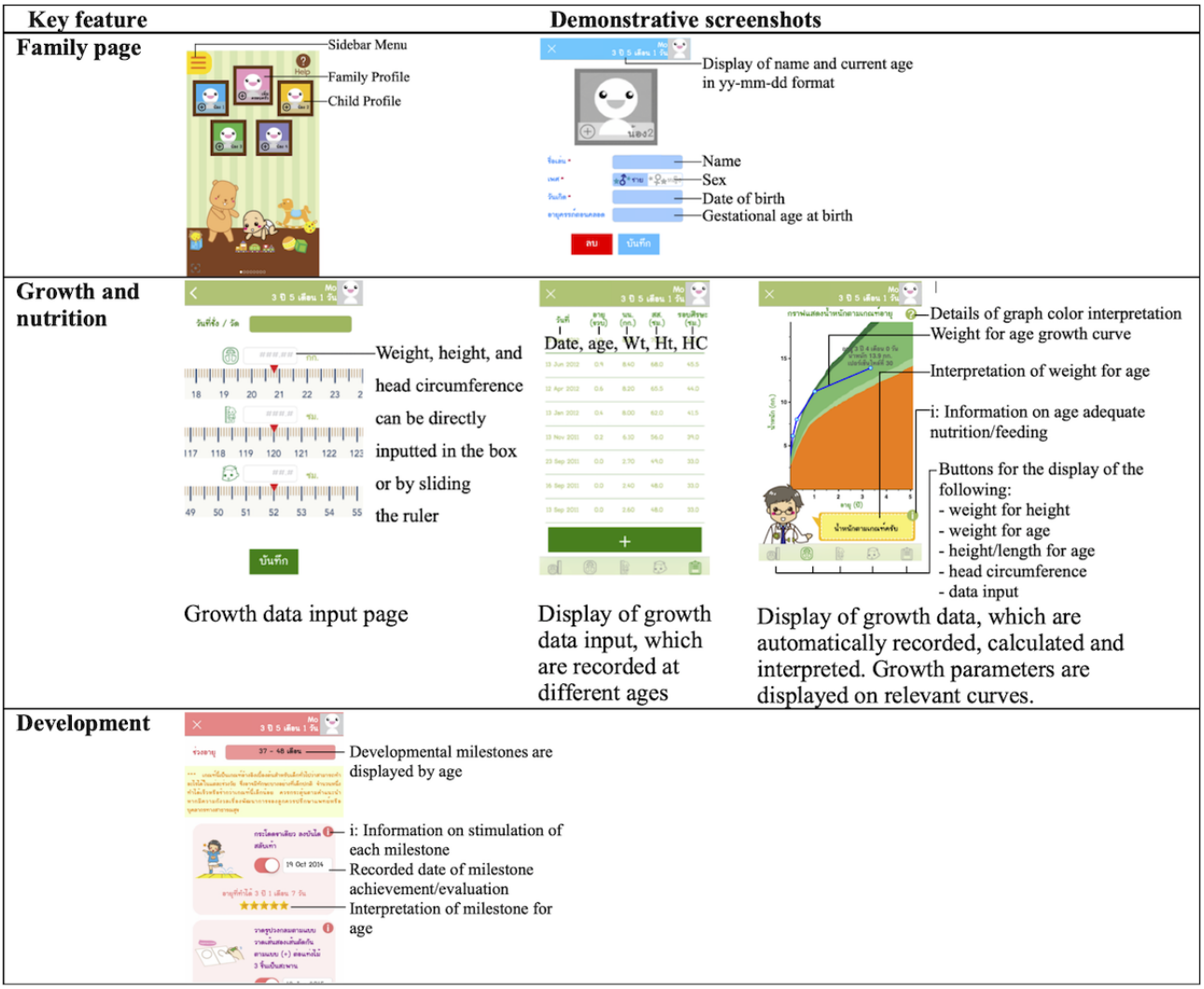
The objectives of the phase 2 of the study were as follows: (1) to evaluate the congruence of the child growth assessments between parents and physicians and (2) to compare the proportion of congruence of parent and physician child growth assessments between the App group and the MCHH group.

Study Design

Although KhunLook has many health assessment domains as shown in Figure 2 and Figure 3, we used only the growth assessment in our phase 2 study because growth assessment is an objective, important, and practical way to monitor a child's health. We selectively focused on this parameter to investigate whether parents could use and understand the outcomes of their child's growth assessment. Moreover, with correct usage of the same growth curves, parent and physician growth assessments should yield the same outcomes, but with typical use, outcomes may differ. We investigated the effectiveness of the app in assisting parents in the early detection of abnormalities in their child's growth parameters. In this quasi-experimental study design, children's growth parameters were measured by health care providers in well-child clinics, and the parameters were provided to parents. A convenient group of parents was recruited and randomly assigned to 2 groups: the "App group" used the KhunLook app and the "MCHH group" used the MCHH to assess their child's growth. The App group had no other guidance other than assistance with loading the app. Parents who own an MCHH are familiar with the handbook, and theoretically, app users should intuitively be able to use an engaging well-designed app without extra training [17]. The KhunLook app version 1.6 for the iPhone operating system or version 1.23 for the Android operating system, which had equivalent functionality, were used. Physicians assessed the child's growth and provided standard child health supervision care by using the same growth chart standards. Physicians were blinded to each group. They did not use the app and they were not part of the app development team. The inclusion criteria of the parents of children under the age of 6 years were as follows: (1) own an MCHH and (2) use a smartphone/tablet with the iPhone or Android operating system.

XSL•FO
RenderX

Figure 2. Demonstrative screenshots of the KhunLook app.



Data Collection

We used a questionnaire to collect the baseline characteristics and outcomes of growth assessment by parents and physicians. Growth parameters that were assessed included weight for age, height for age, and head circumference.

Data Analysis

Growth parameters were categorized into 3 categories: below normal, normal, and above normal. Growth parameters were considered normal when they were between the 3rd and 97th percentile for age and sex, which was the reference provided in the app and MCHH [18]. Parents' assessments that matched the physician's assessment were considered congruent. For demographics, two-sided *t* test and chi-square test were used, when appropriate. The Stuart-Maxwell test for homogeneity was used to determine the agreement of the child growth assessments between parents and physicians. The chi-square test was used to determine the differences between the proportions of parent-physician agreement of child growth assessments between the App group and the MCHH group. The STATA Statistical Software (Release 14. College Station, TX: StataCorp LP) was used for the statistical calculations.

Phase 3: Parents' Evaluation of the Feasibility and Acceptability of the KhunLook App

The phase 3 objective was to assess parents' evaluations of the KhunLook app in terms of feasibility and acceptability.

Study Design

For a wider range of user evaluations, a web-based survey was used to collect the parents' evaluations of the KhunLook app and the MCHH in comparable sections in terms of feasibility and acceptability. The survey was launched in July 2017. The inclusion criteria of the parents of children under the age of 6 years were as follows: (1) own an MCHH and (2) used the KhunLook app within the past month.

Data Collection

We used a web-based questionnaire to collect information on the baseline characteristics, feasibility, and acceptability of the app and MCHH. For feasibility, parents rated the ease of using the app compared to the MCHH on 5 health assessment domains and 6 ease-of-use domains on a 4-rank rating scale (very easy, easy, difficult, and very difficult). For acceptability, parents were asked to provide an overall score from 0 to 10 (0=least acceptable, 10=most acceptable) for the app and MCHH; they were asked to choose the preferred method, if they would

XSL•FO
RenderX

continue to use the app, and if they would advise others to use the app.

Data Analysis

For feasibility, the first 2 ratings were grouped as "very easy to easy" and the last 2 were grouped as "difficult to very difficult," and then, the proportions were calculated. We used McNemar test for change to compare the feasibility ratings between the mobile app and MCHH. For acceptability, a paired two-sided *t* test was used to compare the scores between the KhunLook app and MCHH. The STATA Statistical Software was used for the statistical calculations.

Results

Phase 1: User Requirements and Development of the KhunLook Mobile App

User Requirements

Four health care providers and 8 parents participated in this phase. Half of the health care providers were women with a mean (SD) age of 49 (19) years (range, 36-76 years) and 50% (2/4) had practice experiences of more than 15 years. Three providers were pediatricians and one was a dentist. All of the parents were mothers with mean (SD) age of 32 (2.7) years (range, 30-36 years), and regarding their levels of education, 75% (6/8) had a doctorate degree and 25% (2/8) had a bachelor's degree; 88% (7/8) of the mothers had a monthly income of more than 15,000 Baht per month. Five mothers had 2 children, while 3 had 1 child. The mean (SD) age of their children was 3.6 (2.3) years (range 0.9-6 years). Five mothers used iPhones and 3 used Android smartphones. One mother had previously used mobile apps to track her child's health. From the qualitative data analysis, we identified 2 major themes that were raised by the participants and relevant to our research questions. These themes are discussed in greater detail along with the representative quotes below (Table 1).

Table 1. Themes and the illustrative quotes.

Theme	Illustrative quote
The mobile app potentially counters parents' infrequent use of the MCHH[a] with accuracy, attractiveness, convenience, and simplicity.	*…I don't know how many parents even open the MCHH. Most of them don't. They tell me it's for the doctor to read even when previous versions have included a section for parents to record. The MCHH changes from year to year and parents sometime forget to bring it to health supervision visits. Some lose them in floods. An app could easily and accurately help them check if their numbers (child's growth parameters) are normal, and it should provide brief information for further actions*. [Quote 1, Health care provider 2]
	…When I had my first child, I always read the MCHH and plotted his growth parameters. As for my second child, I just used the MCHH to record vaccines. I rely on the doctors for immunization appointments. I never recorded immunizations either; the doctor does that. I don't really record things in the MCHH book. I bring it along for health supervision visits, not sick visits. If there were an app on my mobile, it would make things more convenient. [Quote 2, Parent 1]
	…This is great! But does it need to be constantly connected to the internet? Because not many people will pay for the charge, parents in remote areas won't. [Quote 3, Parent 2]
Health supervision needs to be standard, up-to-date, and understandable.	*…For child health supervision, the app needs to have sections about growth, vaccination, development, which are standard and up-to-date. Once guidelines or standards change, they should be reflected in the app. The MCHH cannot do this. The parameters should be corrected for premature babies.* [Quote 4, Health care provider 3]
	I don't think the parents understand what the Bacillus Calmette-Guérin vaccine is; it needs to be explained in Thai briefly. I think all vaccinations should be presented by age; that makes it easier for parents. [Quote 5, Health care provider 4]
	…I would use it for growth and development assessment; it should tell the parent what to do if their child has abnormal development, like go see the doctor or the kind of stimulation that is needed, and it should use a corrected age for preterm babies. I've searched for apps that could do this but have not yet found any that are free of charge. [Quote 6, Parent 3]
	...This app is very complete, organized, and easy to use. I can automatically assess my child's growth, development, vaccines. The information comes from a reliable source and it even has a journal where I can add pictures. I would definitely use this app. [Quote 7, Parent 4]

[a]MCHH: maternal and child health handbook.

Theme 1: Health care providers and parents expressed that mobile apps are potentially useful by countering the infrequent usage of the MCHH and by helping caregivers to understand their child's health. The downside would be the need for internet connectivity.

Theme 2: Health care providers and parents expressed the need for standard and up-to-date information that is in Thai language.

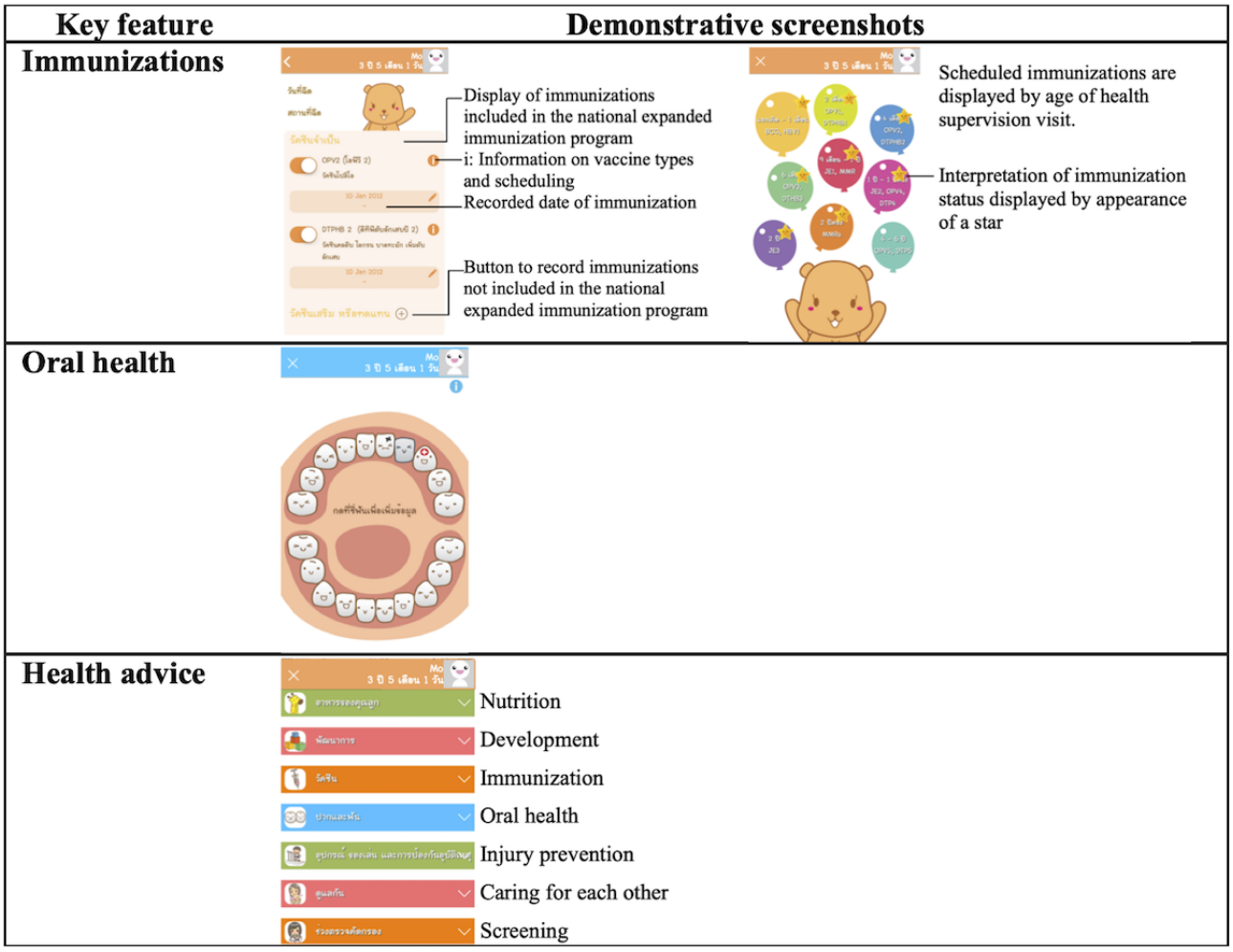
KhunLook Mobile App

The updated version of the KhunLook mobile app based on phase 1 input contains a family page and 7 key features: growth and nutrition, development, immunizations, oral health, reminders for the next visit appointment, memories, and health advice. Demonstrative screenshots are shown in Figure 2 and Figure 3. For assessments of growth, development and

XSL•FO
RenderX

immunizations, parents are required to input data for the app to record and automatically calculate, interpret, and display.

Figure 3. Demonstrative screenshots of the KhunLook app.



Growth standards were based on the Thai national growth reference [18], and growth parameters were assessed in relation to sex and chronological age for term infants (born at gestational age≥37 weeks) and in relation to corrected age (up to gestational age of 40 weeks) until 2 years of age for preterm infants born at gestational age <37 weeks. Assessment of development was based on the Royal College of Pediatricians of Thailand's child developmental milestone reference [19]. Immunization status was assessed based on the Ministry of Public Health and the Pediatric Infectious Disease Society of Thailand's immunization schedule standard [20,21]. Health advice was based on the Ministry of Public Health and the Royal College of Pediatricians of Thailand child health supervision guidelines [22]. After installation, the app fully functions without requiring internet connectivity, as mobile broadband internet use was not prevalent in Thailand when the app was first launched. Data inputted by users were stored locally on the mobile phone to ensure user privacy. The overall design of the app is colorful and playful to appeal to the target user group. Internet connection is only required for app updates and optional data backup.

Phase 2: Validity of the Growth Assessment Study

Of the 56 parents who participated, 34 were in the App group and 22 were in the MCHH group. The baseline characteristics of the parents are shown in Table 2. There were no statistically significant differences between the parental characteristics in both the groups. The differences in the children's mean age was significantly less, and there was a lower proportion of girls in the App group. The distribution of the growth parameters was similar in both the groups. Most of the children received an MCHH version printed within 1 year of birth (36/56, 64%), and some received versions printed within 1-2 years (12/56, 22%) and >3 years (8/56, 14%) of birth. The congruence of the child growth assessments between parents and physicians are shown in Table 3. There were no statistically significant differences between the outcomes of the growth parameter assessments between parents and physicians in both the App and MCHH groups. The proportions of congruence in the App group for weight and head circumference were higher than those in the MCHH group but the differences between each group were not statistically significant (P=.65, P=.54, and P=.13 for weight for age, height for age, and head circumference between the App and MCHH group, respectively).

XSL•FO
RenderX

Table 2. Baseline characteristics of the parents and the children by group.

Demographics	App group, n=34	MCHH[a] group, n=22	*P* value
Parents			
Sex, female, n (%)	27 (79)	17 (77)	.85
Age (years)			.81
Mean (SD)	33.7 (5.4)	34.0 (5.8)	
Range	20-45	21-44	
Education, n (%)			.11
Less than bachelor's degree	3 (9)	5 (23)	
Bachelor's degree	12 (35)	11 (50)	
Master's or doctoral degree	19 (56)	6 (27)	
Monthly income <15000 Baht/month[b], n (%)	3 (9)	3 (14)	.65
Type of phone, n (%)			.12
iPhone	24 (71)	11 (50)	
Android	10 (29)	11 (50)	
Parents who used other apps to assess child's health in the past, n (%)	8 (24)	7 (3)	.49
Children			
Sex, female, n (%)	11 (32)	13 (59)	.048[c]
Age (years)			.002[c]
Mean (SD)	2.6 (1.3)	3.8 (1.6)	
Range	0.3-4.9	0.1-6	
Growth parameters, n (%)			
Weight			.82
Below normal	1 (3)	1 (5)	
Normal	31 (91)	19 (86)	
Above normal	2 (6)	2 (9)	
Length/height			.90
Below normal	1 (3)	1 (5)	
Normal	32 (94)	20 (90)	
Above normal	1 (3)	1 (5)	
Head circumference			.75
Below normal	0 (0)	0 (0)	
Normal	33 (97)	21 (95)	
Above normal	1 (3)	1 (5)	
MCHH version			
Year (range)	2005-2014	2007-2013	

[a]MCHH: maternal and child health handbook.

[b]1 USD=35.3 Baht.

[c]Differences were statistically significant at $P<.05$.

XSL•FO
RenderX

Table 3. Congruence of child growth assessments between parents and physicians by group.

Group, growth parameter assessment	Parent, normal assessment, n (%)	Physician, normal assessment, n (%)	Congruence of assessment outcome, n (%)	*P* value
App group (n=34)				
Weight	29 (85)	31 (91)	31 (91)	.16
Length/height	29 (85)	32 (94)	31 (91)	.22
Head circumference	32 (94)	33 (97)	33 (97)	.32
MCHH[a] group (n=22)				
Weight	17 (77)	19 (86)	20 (91)	.37
Length/height	20 (91)	20 (91)	21 (95)	.61
Head circumference	18 (82)	21 (95)	19 (86)	.08

[a]MCHH: maternal and child health handbook.

Phase 3: Parents' Evaluation of the KhunLook App

In this study, 356 parents from all regions of Thailand participated in the web-based survey. The baseline characteristics are shown in Table 4. Evaluation of the KhunLook app compared to the MCHH for feasibility in terms of health assessment and usage are shown in Table 5. Parents rated the feasibility of the app as "very easy to easy" to use at higher proportions than the MCHH in all domains with statistical significance ($P<.001$, Table 5). The majority (354/356, 99.4%) rated the installation of the KhunLook app as very easy to easy.

XSL•FO
RenderX

Table 4. Baseline characteristics of the parents and children in the web-based survey (n=356).

Demographics	Values
Parents, n=356	
Sex, female, n (%)	331 (93.0)
Age (years)	
Mean (SD)	28.1 (6.1)
Range (years)	18-45
Region of Thailand, n (%)	
North	50 (14.0)
Northeast	57 (16.0)
East	54 (15.2)
Central	147 (41.3)
South	25 (7.0)
West	23 (6.5)
Education, n (%)	
Less than bachelor's degree	31 (8.7)
Bachelor's degree	217 (61.0)
Master's or doctoral degree	108 (30.3)
Monthly income < 15,000 Baht/month[a], n (%)	25 (7.0)
Number of children, n (%)	
1	281 (78.9)
2	59 (16.6)
>3	16 (4.5)
Phone operating system, n (%)	
iPhone	229 (64.3)
Android	127 (35.7)
Parents who used other apps to assess child's health in the past, n (%)	24 (6.7)
Children, n=356	
Age (years), mean (SD)	2.92 (2.9)
Sex, female, n (%)	131 (36.7)

[a]1 USD=33 Baht.

XSL•FO
RenderX

Table 5. Parents' ratings of feasibility: health assessment and convenience of use of mobile app versus maternal and child health handbook (n=356).

Feasibility	App	MCHH[a]	*P* value
Health assessment (very easy to easy), n (%)			
Weight	352 (98.9)	268 (75.3)	<.001
Length/height	350 (98.3)	268 (75.3)	<.001
Head circumference	350 (98.3)	271 (76.1)	<.001
Development	353 (99.2)	279 (78.4)	<.001
Immunization	333 (93.5)	284 (79.8)	<.001
Convenience of use (easy to easy), n (%)			
Data input	352 (98.9)	287 (80.6)	<.001
Access to desired segment	340 (95.5)	248 (69.7)	<.001
Understandability of content	342 (96.1)	290 (81.7)	<.001
Applicability of content	346 (97.2)	282 (79.2)	<.001
Usefulness	354 (99.4)	261 (73.3)	<.001
Overall convenience	354 (99.4)	252 (70.8)	<.001

[a]MCHH: maternal and child health handbook.

For acceptability, on a scale of 10, the KhunLook app received a mean (SD) score of 8.59 (1.1) (range 3-10), which was significantly higher than that of the MCHH (7.6 [1.8], range 0-10; *P*<.001). The KhunLook app received a higher score than the MCHH from 199 (55.9%) parents, equal scores from 138 (38.7%) parents, and lower scores from 20 (5.6%) parents of the total population of 356 parents. If parents had to choose a between the app and MCHH, most (317/356, 89.0%) preferred to use the app. In addition, 93.5%, (333/356) of parents stated that they would continue to use the app and 96.9% (345/356) stated that they would recommend others to use it.

Discussion

Principal Results

This is the first study to describe the development and evaluation of a mobile app for child health supervision by comparing the congruence of parent-physician growth assessments and parental evaluation in Thailand. Our participants in phase 1 offered insight into the requirements of an ideal app. Participants suggested that the app could counter parents' infrequent use of the MCHH book. They also emphasized the importance of standard, up-to-date, understandable content, and offline functionality. Parents who used the app or the MCHH could comparably assess their child's growth status to that of a physician's assessment. Parents who used the app had higher proportions of congruence to the physician's assessment for weight and head circumference than parents who used the MCHH, but the differences were not statistically significant. For acceptability, parents rated the app significantly more feasible and acceptable in relation to the MCHH (*P*<.001).

mHealth technology interventions in maternal and child health are increasing worldwide [23]. There is a call for improvement in the content and quality of health care apps, including evidence-based information consistent with guidelines, supported and developed by health care providers with credible and reliable resources [24-26]. A study in 2015 from Australia found that only 40% of the apps involved health care professionals and provided evidence-based content, while 30% implemented user privacy security measures [24]. The KhunLook app was developed by health care providers and app developers in conjunction with parents since conception. It uses evidence-based content and honors user data privacy. Both themes from phase 1 support the development of a mobile app to assist with child health supervision.

Mobile apps are used to support interventions for maternal and child health care [27,28]. Studies have used mobile apps to connect and send data between health care providers and patients or use it for means of notifications, with positive outcomes [26-34]. In Kenya, the IFA app was found to shift the relationship between the caregiver and health care provider from feeling harassed for data to being genuinely interested [33]. For child health supervision in Thailand, mobile phones were successfully used to connect with parents for child immunization appointments via text messages, collect data via image capture, and conduct hearing screenings [29,30,34]. Findings from phase 2 add to the benefit of using a mobile app to ease the process of assessing growth parameters and instantly taking numbers to a deeper meaning of growth status. We found that 91.2%-97.1% of the parents who used the app could correctly assess their child's growth parameters, whereas only 86.4%-95.5% of the parents who used the MCHH could do so. Thus, our findings suggest that the KhunLook app yields the benefits of mHealth services and can support parents in assessing their child growth similar to the MCHH.

Successful benefits of mHealth services require acceptability and engagement from the user [23,35]. A study found that use of a mHealth intervention in patients with type 2 diabetes is associated with its acceptability [35]. According to our findings, KhunLook was well accepted. Results from a web-based survey rated the app to be feasible for parents for initial assessments and easier to use than the MCHH. Parents provided a higher

XSL•FO
RenderX

mean score for the KhunLook app than the MCHH with statistical significance (P<.001). Most of the parents (317/356, 89.0%) preferred to use the mobile app rather than the MCHH, 93.5% (333/356) stated that they would continue to use the app and 96.9% (345/356) would recommend others to use the app. Our findings support the KhunLook app as an acceptable mode of delivery for child health supervision for parents.

We developed the KhunLook app based on the belief that parental involvement would increase early detection and adequate childcare. Initially, the goal was to develop the app to compliment the MCHH. During the process, it has become eminent that the maintenance of content update, feature development, and user service is crucial. However, for the app to reach its potential, implementation at a wide scale is imperative. We started with a small team of health care providers, app developers, and users. Later on, we involved stakeholders on a wider scale. While drafting this paper (August 2020), the KhunLook app has been downloaded more than 320,000 times. It is endorsed by the Ministry of Public Health, and a link to download the app is currently provided in the MCHH.

Limitations

The results of this study are interpretations from a certain time and do not reflect the current app. The development of "KhunLook" in phase 1 and 2 involved a convenient sample of parents at a university hospital. Most of the parents were well educated and were willing to try new technology but the number of participants was too small to draw generalized conclusions.

Conclusions

In this study, KhunLook, a Thai mobile app for child health supervision, was developed, validated for growth assessments, and was found to be well accepted for ease-of-use by parents. The full potential of this mHealth app is yet to be defined. Further studies on parental and clinical use should be conducted such as a randomized study involving a wider scale of users or studies to evaluate its impact on health behaviors and health outcomes.

Acknowledgments

This study was funded by the Health Systems Research Institute, Ministry of Public Health, and the National Research Council of Thailand.

Conflicts of Interest

None declared.

References

1. Hagan J, Shaw J, Duncan P. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents. 4th edition. Elk Grove Village, IL: American Academy of Pediatrics; 2017.
2. Hagan J, Duncan P. Maximizing Children's Health: Screening, Anticipatory Guidance and Counseling. In: Kliegman RM, Stanton BF, St Geme JW, Schor NF. editors. Nelson Textbook of Pediatrics. 20th edition. Philadelphia: Elsevier, Inc; 2016.
3. Isaranurug S. Maternal and Child Health (MCH) handbook in the world maternal and child health handbook in Thailand. Kokusai Hoken Iryo (Journal of International Health);(24) 2009:61-66 [FREE Full text]
4. Jandee K, Kaewkungwal J, Khamsiriwatchara A, Lawpoolsri S, Wongwit W, Wansatid P. Effectiveness of Using Mobile Phone Image Capture for Collecting Secondary Data: A Case Study on Immunization History Data Among Children in Remote Areas of Thailand. JMIR Mhealth Uhealth 2015 Jul 20;3(3):e75 [FREE Full text] [doi: 10.2196/mhealth.4183] [Medline: 26194880]
5. Takeuchi J, Sakagami Y, Perez RC. The Mother and Child Health Handbook in Japan as a Health Promotion Tool: An Overview of Its History, Contents, Use, Benefits, and Global Influence. Glob Pediatr Health 2016;3:2333794X16649884. [doi: 10.1177/2333794X16649884] [Medline: 27336022]
6. de Onis M, Onyango A, Borghi E, Siyam A, Blössner M, Lutter C, WHO Multicentre Growth Reference Study Group. Worldwide implementation of the WHO Child Growth Standards. Public Health Nutr 2012 Sep;15(9):1603-1610. [doi: 10.1017/S136898001200105X] [Medline: 22717390]
7. Roberfroid D, Pelto GH, Kolsteren P. Plot and see! Maternal comprehension of growth charts worldwide. Trop Med Int Health 2007 Sep;12(9):1074-1086 [FREE Full text] [doi: 10.1111/j.1365-3156.2007.01890.x] [Medline: 17875018]
8. Ben-Joseph EP, Dowshen SA, Izenberg N. Public understanding of growth charts: a review of the literature. Patient Educ Couns 2007 Mar;65(3):288-295. [doi: 10.1016/j.pec.2006.09.001] [Medline: 17081719]
9. Wright CM, Sachs M, Short J, Sharp L, Cameron K, Moy RJ. Designing new UK-WHO growth charts: implications for health staff use and understanding of charts and growth monitoring. Matern Child Nutr 2012 Jul;8(3):371-379 [FREE Full text] [doi: 10.1111/j.1740-8709.2010.00296.x] [Medline: 21332642]
10. Pongsuthirak N. Evaluation of mother and child health book usage of parents at healthy child clinic Buddhachinaraj Phitsanulok Hospital. Thai Pediatric Journal;(19) 2012:165-173 [FREE Full text]
11. World Health Organization. mHealth: New horizons for health through mobile technologies: Second global survey on eHealth. Geneva, Switzerland: World Health Organization; 2011:9789241564250.

XSL•FO
RenderX

12. National Statistical Office of Thailand. 2007-2016 Statistics on Usage of Informational Technology by device in Thai Population above 6 years of age. Ministry of Information and Communication Technology. URL: http://service.nso.go.th/nso/web/statseries/statseries22.html [accessed 2019-03-27] [WebCite Cache ID 77CYi8T16]
13. National Statistical Office of Thailand. ICT household survey report 2018. Ministry of Information and Communication Technology. 2018. URL: http://stiic.sti.or.th/stat/ind-it/it-t012/ [accessed 2019-03-27]
14. Virani A, Duffett-Leger L, Letourneau N. Parenting apps review: in search of good quality apps. Mhealth 2019;5:44 [FREE Full text] [doi: 10.21037/mhealth.2019.08.10] [Medline: 31620471]
15. Entsieh AA, Emmelin M, Pettersson KO. Learning the ABCs of pregnancy and newborn care through mobile technology. Glob Health Action 2015;8:29340 [FREE Full text] [doi: 10.3402/gha.v8.29340] [Medline: 26673633]
16. Pope C, Mays N. Qualitative Research in Health Care (3rd edition). Malden: Blackwell Publishing Inc; 2013:978-971.
17. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: Content analysis of apps and user ratings. JMIR Mhealth Uhealth 2015 May 13;3(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]
18. Nutrition Division. National Growth References for Children Under 20 years of age. Bangkok: Department of Health, Ministry of Public Health, Thailand; 1999.
19. Royal College of Pediatricians of Thailand. Child developmental milestones (Thai). Bangkok: Royal College of Pediatricians of Thailand; 2014. URL: http://www.thaipediatrics.org/file_upload/files/pratom.pdf [accessed 2019-03-27] [WebCite Cache ID 76gxA6No1]
20. World Health Organization. Expanded Programme on Immunization (EPI) Fact Sheets-Thailand. 2015. URL: http://www.searo.who.int/immunization/data/thailand.pdf [accessed 2019-03-27] [WebCite Cache ID 76gs3JArx]
21. Pediatric Infectious Society of Thailand. Vaccination Table for Thai Children (Thai). Bangkok: Pediatric Infectious Society of Thailand; 2015. URL: http://www.pidst.or.th/A386.html [accessed 2019-03-27] [WebCite Cache ID 76hxvLPJQ]
22. Noipayak P, Piyasil V, Ningsanon W, Ungthavorn P, editors. The Royal College of Pediatricians of Thailand: Guideline in Child Health Supervision (Thai). Bangkok: Sapphasan Ltd; 2014.
23. Chen H, Chai Y, Dong L, Niu W, Zhang P. Effectiveness and Appropriateness of mHealth Interventions for Maternal and Child Health: Systematic Review. JMIR Mhealth Uhealth 2018 Jan 09;6(1):e7 [FREE Full text] [doi: 10.2196/mhealth.8998] [Medline: 29317380]
24. Scott K, Gome G, Richards D, Caldwell P. How trustworthy are apps for maternal and child health? Health Technol 2015;4:329-336 [FREE Full text] [doi: 10.1007/s12553-015-0099-x]
25. Cheng H, Tutt A, Llewellyn C, Size D, Jones J, Taki S, et al. Content and Quality of Infant Feeding Smartphone Apps: Five-Year Update on a Systematic Search and Evaluation. JMIR Mhealth Uhealth 2020 May 27;8(5):e17300 [FREE Full text] [doi: 10.2196/17300] [Medline: 32459187]
26. Davis D, Logsdon M, Vogt K, Rushton J, Myers J, Lauf A, et al. Parent Education is Changing. MCN, The American Journal of Maternal/Child Nursing 2017;42(5):248-256. [doi: 10.1097/nmc.0000000000000353]
27. Zhang P, Dong L, Chen H, Chai Y, Liu J. The Rise and Need for Mobile Apps for Maternal and Child Health Care in China: Survey Based on App Markets. JMIR Mhealth Uhealth 2018 Jun 08;6(6):e140 [FREE Full text] [doi: 10.2196/mhealth.9302] [Medline: 29884601]
28. Fernandez-Luque L, Labarta JI, Palmer E, Koledova E. Content Analysis of Apps for Growth Monitoring and Growth Hormone Treatment: Systematic Search in the Android App Store. JMIR Mhealth Uhealth 2020 Feb 18;8(2):e16208 [FREE Full text] [doi: 10.2196/16208] [Medline: 32130162]
29. Kaewkungwal J, Singhasivanon P, Khamsiriwatchara A, Sawang S, Meankaew P, Wechsart A. Application of smart phone in "Better Border Healthcare Program": a module for mother and child care. BMC Med Inform Decis Mak 2010 Nov 03;10:69 [FREE Full text] [doi: 10.1186/1472-6947-10-69] [Medline: 21047412]
30. Kaewkungwal J, Apidechkul T, Jandee K, Khamsiriwatchara A, Lawpoolsri S, Sawang S, et al. Application of mobile technology for improving expanded program on immunization among highland minority and stateless populations in northern Thailand border. JMIR Mhealth Uhealth 2015 Jan 14;3(1):e4 [FREE Full text] [doi: 10.2196/mhealth.3704] [Medline: 25589367]
31. Balakrishnan R, Gopichandran V, Chaturvedi S, Chatterjee R, Mahapatra T, Chaudhuri I. Continuum of Care Services for Maternal and Child Health using mobile technology - a health system strengthening strategy in low and middle income countries. BMC Med Inform Decis Mak 2016 Jul 07;16:84 [FREE Full text] [doi: 10.1186/s12911-016-0326-z] [Medline: 27387548]
32. Obasola OI, Mabawonku I, Lagunju I. A Review of e-Health Interventions for Maternal and Child Health in Sub-Sahara Africa. Matern Child Health J 2015 Aug;19(8):1813-1824. [doi: 10.1007/s10995-015-1695-0] [Medline: 25652059]
33. van Heerden A, Sen D, Desmond C, Louw J, Richter L. App-Supported Promotion of Child Growth and Development by Community Health Workers in Kenya: Feasibility and Acceptability Study. JMIR Mhealth Uhealth 2017 Dec 05;5(12):e182 [FREE Full text] [doi: 10.2196/mhealth.6911] [Medline: 29208588]
34. Yimtae K, Israsena P, Thanawirattananit P, Seesutas S, Saibua S, Kasemsiri P, et al. A Tablet-Based Mobile Hearing Screening System for Preschoolers: Design and Validation Study. JMIR Mhealth Uhealth 2018 Oct 23;6(10):e186 [FREE Full text] [doi: 10.2196/mhealth.9560] [Medline: 30355558]

XSL•FO
RenderX

35. Torbjørnsen A, Småstuen M, Jenum A, Årsand E, Ribu L. Acceptability of an mHealth App Intervention for Persons With Type 2 Diabetes and its Associations With Initial Self-Management: Randomized Controlled Trial. JMIR Mhealth Uhealth 2018 May 21;6(5):e125 [FREE Full text] [doi: 10.2196/mhealth.8824] [Medline: 29784635]

Abbreviations

MCHH: maternal and child health handbook
mHealth: mobile health

Edited by G Eysenbach, R Kukafka; submitted 26.06.19; peer-reviewed by P Zhang, S Lawpoolsri, J Kaewkungwal; comments to author 30.03.20; revised version received 22.08.20; accepted 03.10.20; published 30.10.20.

Please cite as:
Areemit R, Lumbiganon P, Suphakunpinyo C, Jetsrisuparb A, Sutra S, Sripanidkulchai K
A Mobile App, KhunLook, to Support Thai Parents and Caregivers With Child Health Supervision: Development, Validation, and Acceptability Study
JMIR Mhealth Uhealth 2020;8(10):e15116
URL: http://mhealth.jmir.org/2020/10/e15116/
doi:10.2196/15116
PMID:33124989

©Rosawan Areemit, Pagakrong Lumbiganon, Chanyut Suphakunpinyo, Arunee Jetsrisuparb, Sumitr Sutra, Kunwadee Sripanidkulchai. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 30.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Screening for Hearing Impairment in Older Adults by Smartphone-Based Audiometry, Self-Perception, HHIE Screening Questionnaire, and Free-Field Voice Test: Comparative Evaluation of the Screening Accuracy With Standard Pure-Tone Audiometry

Lok Yee Joyce Li[1,2,3], MD; Shin-Yi Wang[4], MSc; Cheng-Jung Wu[2,3,5,6], MD; Cheng-Yu Tsai[7], PhD; Te-Fang Wu[8], BA; Yaoh-Shiang Lin[6,8], MD

[1]Department of Medicine, Shin Kong Wu-Ho-Su Memorial Hospital, Taipei, Taiwan
[2]Department of Otolaryngology, Shuang Ho Hospital, Taipei Medical University, New Taipei City, Taiwan
[3]Department of Otolaryngology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan
[4]National Taiwan University Hospital Hsin-chu Branch, Hsin-chu, Taiwan
[5]PhD Degree Program of Biomedical Science and Engineering, National Chiao Tung University, Hsinchu, Taiwan
[6]Department of Otolaryngology-Head And Neck Surgery, Tri-Service General Hospital, Taipei, Taiwan
[7]Department of Civil and Environmental Engineering, Imperial College London, London, United Kingdom
[8]Department of Otolaryngology - Head And Neck Surgery, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan

Corresponding Author:
Cheng-Jung Wu, MD
Department of Otolaryngology, Shuang Ho Hospital, Taipei Medical University
No. 291, Zhongzheng Rd, Zhonghe District
New Taipei City, 23561
Taiwan
Phone: 886 22490088
Email: b101090126@tmu.edu.tw

Abstract

Background: Hearing impairment is the most frequent sensory deficit in humans, affecting more than 360 million people worldwide. In fact, hearing impairment is not merely a health problem, but it also has a great impact on the educational performance, economic income, and quality of life. Hearing impairment is therefore an important social concern.

Objective: We aimed to evaluate and compare the accuracy of self-perception, Hearing Handicap Inventory for the Elderly-Screening (HHIE-S) questionnaire, free-field voice test, and smartphone-based audiometry as tests for screening moderate hearing impairment in older adults in China.

Methods: In this study, 41 patients were recruited through a single otology practice. All patients were older than 65 years. Patients with otorrhea and cognitive impairment were excluded. Moderate hearing impairment was defined as mean hearing thresholds at 500, 1000, 2000, and 4000 Hz >40 dB hearing loss (pure-tone average > 40 dB hearing loss). All patients completed 5 hearing tests, namely, the self-perception test, HHIE-S questionnaire test, free-field voice test, smartphone-based audiometry test, and standard pure-tone audiometry by the same audiologist. We compared the results of these tests to the standard audiogram in the better-hearing ear.

Results: The sensitivity and the specificity of the self-perception test were 0.58 (95% CI 0.29-0.84) and 0.34 (95% CI 0.19-0.54), respectively. The sensitivity and the specificity of the HHIE-S questionnaire test were 0.67 (95% CI 0.35-0.89) and 0.31 (95% CI 0.316-0.51), respectively. The sensitivity and the specificity of the free-field voice test were 0.83 (95% CI 0.51-0.97) and 0.41 (95% CI 0.24-0.61), respectively. The sensitivity and the specificity of the smartphone-based audiometry test were 0.92 (95% CI 0.60-0.99) and 0.76 (95% CI 0.56-0.89), respectively. Smartphone-based audiometry correctly diagnosed the presence of hearing loss with high sensitivity and high specificity.

Conclusions: Smartphone-based audiometry may be a dependable screening test to rule out moderate hearing impairment in the older population.

XSL•FO
RenderX

(JMIR Mhealth Uhealth 2020;8(10):e17213) doi:10.2196/17213

KEYWORDS

hearing impairment; self-perception; HHIE-S questionnaire; free-field voice test; mobile phone; audiometry; mobile health

Introduction

Sensory deficit is defined as a condition wherein any one of the senses, that is, sight, hearing, touch, taste, or smell is no longer functioning normally. Based on the available data, the 2 most commonly encountered sensory impairments are blindness and deafness [1]. Hearing impairment is one of the most frequent sensory deficits in human beings, and it has a profound effect on the life of the affected persons, their families, and the society as a whole. Hearing impairment is actually not just a health problem, because it affects the educational opportunities, the economic situation, and the quality of life of individuals with these impairments. Hearing impairment affects more than 5% of the world's population. In 2012, the World Health Organization released new estimates on the magnitude of disabling hearing loss [2]. The estimates are based on a review of 42 population-based studies carried out up to 2010 [3]. Based on these studies, the World Health Organization estimated that there are 360 million persons in the world with disabling hearing loss. Approximately 328 million of these are adults and 32 million of these are children. About one-third of the persons older than 65 years are affected by hearing impairment [2]. The prevalence of hearing impairment in adults over 65 years is the highest in limited-income countries. However, over 50% of the causes for hearing impairment are preventable [4].

A large number of cases of hearing loss are preventable and many can be treated effectively and immediately. People with hearing losses due to other causes that cannot be treated effectively can be rehabilitated through various available measures, and the integration of such people into the society can be improved [1]. However, limited-income countries face many problems in order to achieve the global aim of preventing and rehabilitating hearing impairments in older adults [5]. The first challenge in these limited-income countries is that medical equipment are limited. Advanced diagnostic tools and standard audiometry tests are quite limited. Due to the limited medical supplies, primary health care professionals face difficult triage decisions such as "Who gets to see an audiologist and undergo standard audiometry tests and who will only see a nurse or a community health educator and follow up?" The second challenge in these limited-income countries is the high incidence of conditions that cause hearing loss compared to that in industrialized countries [6]. Chronic otitis media infections constitute the major disease burden in low-income countries. The third challenge is that human resources are also limited. In these countries, it is common for only 1 physician to see to more than 100,000 people. Besides, trained audiologists in these countries are lacking compared to those in industrialized countries.

The gold standard for diagnosing hearing impairment is the standard audiogram. However, there can be financial or geographic obstacles to receiving a timely audiogram test [5]. These problems may lead to delays in the diagnosis of hearing impairment. Delay in the diagnosis of hearing impairment may lead to delay in treatment, which is considered to be associated with low rates of hearing recovery. Owing to the above obstacles, we aimed to develop a simple, rapid, easily applicable, and cost-effective hearing test for the assessment of hearing conditions in low-income countries. With the development of mobile health technology, smartphone-based hearing tests have been developed as screening tools to identify patients with hearing loss. In this study, we evaluated smartphone-based audiometry as a test for screening moderate hearing impairment in older adults and we aimed to validate this test against standard pure-tone audiometry. This study also compared the usefulness of self-perception, the Hearing Handicap Inventory for the Elderly-Screening (HHIE-S) questionnaire, and free-field voice test to screen for moderate hearing impairment in the older adults in China. In this paper, we discuss the accuracy of the hearing-loss screening methods in older adults, including self-perception, HHIE-S questionnaire, free-field voice test, and smartphone-based audiometry.

Methods

Patient Selection

In this study, 41 patients were recruited through a single otology practice in the Kaohsiung Veterans General Hospital. All included patients were older than 65 years. All patients were fluent in Chinese and were able to read and write in Chinese. Patients with active otorrhea, cognitive impairment, Parkinson disease, clinically diagnosed dementia, and hand action tremor were excluded. Patients were excluded if they used hearing aids or received a standard pure-tone audiogram evaluation in the prior 24 months or were unable to complete questionnaires. Patients with conductive hearing loss based on standard pure-tone audiometry hearing test were also excluded as it would have made interpretation of an air-conduction smartphone-based audiometry hearing test impossible. The 5 hearing tests were performed for each patient in a randomized order during a single visit. Free-field voice test, smartphone-based audiometry test, and standard pure-tone audiometry were performed in the same soundproof room with average ambient noise level varying between 38 dBA and 39 dBA-weighted sound pressure level. Informed consent was obtained from each of the patients in this study.

Study Design

All patients completed 5 hearing evaluations, that is, the self-perception test, HHIE-S questionnaire test, free-field voice test, smartphone-based audiometry test, and a standard pure-tone audiometry test. All the hearing tests were conducted by the same audiologist. We compared the results of the self-perception test, HHIE-S questionnaire test, free-field voice test, and smartphone-based audiometry test to those of the standard pure-tone audiogram.

XSL•FO
RenderX

Screening Strategies

Self-Perception Hearing Screening Test

The self-perception hearing screening test was examined using a single question. The question was “Do you have a hearing problem now?” (您現在有聽力問題嗎?) This sentence was used as the subjective criterion of hearing impairment [7]. A yes or an equivalent response to this question was considered as a positive screen for hearing impairment [7]. Participating patients were asked to respond as yes or no. Patients with the answer “yes” were considered to have screened positive for moderate hearing loss [8].

HHIE-S Questionnaire

The HHIE-S questionnaire is a 10-item, self-administered questionnaire developed to assess how an older patient perceives the social and emotional effects of hearing impairment [9,10] (Figure 1 and Multimedia Appendix 1). The HHIE-S questionnaire was developed by the American Speech-Language-Hearing Association (ASHA). Each question was scored as yes (4 points), sometimes (2 points), or no (0 points). The possible scores ranged from 0 (no handicap) to 40 (maximum handicap). The higher the HHIE-S scores, the greater was the handicapping effect of the hearing loss. The questionnaire takes 5 minutes to complete. Patients with scores of 10 or above were considered to have screened positive for moderate hearing loss [9].

Figure 1. Hearing Handicap Inventory for the Elderly Screening Version.

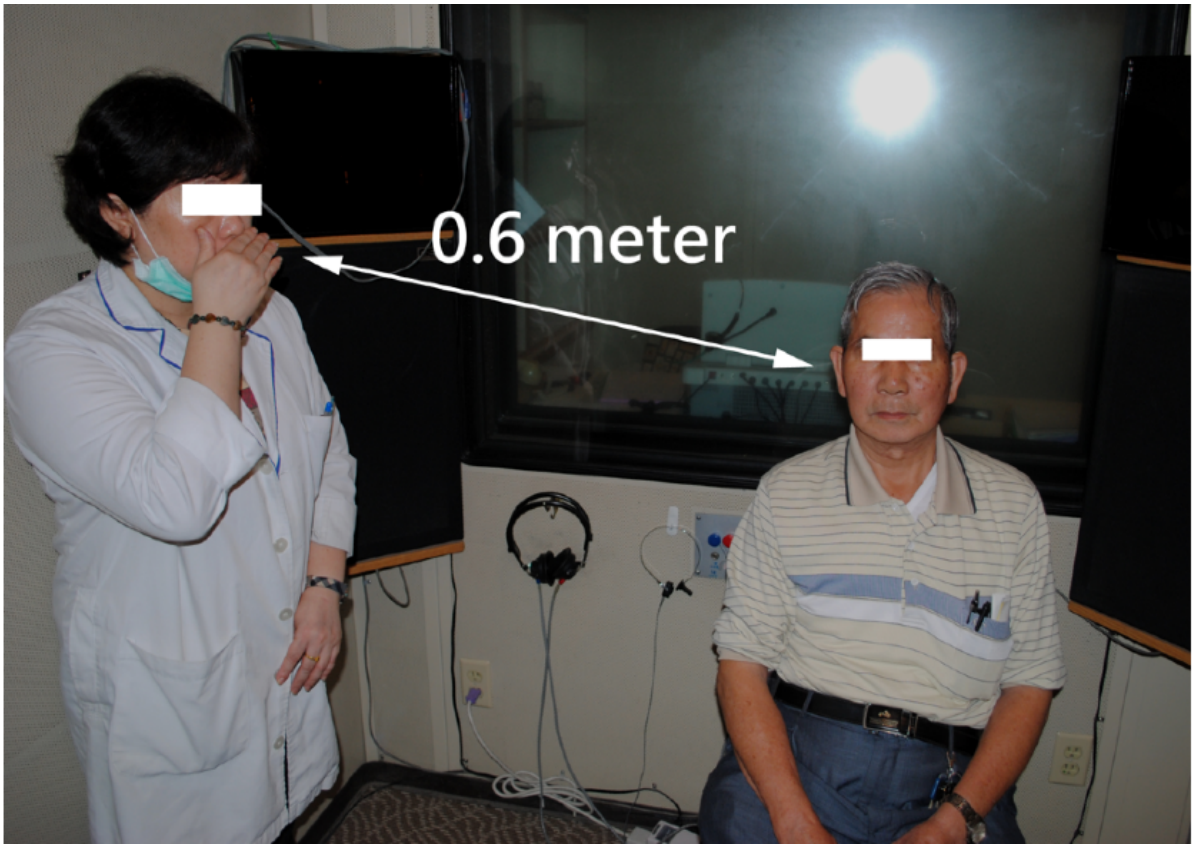
	Item	Yes (4 pts)	Sometimes (2 pts)	No (0 pts)
E	Does a hearing problem cause you to feel embarrassed when meeting new people?	_____	_____	_____
E	Does a hearing problem cause you to feel frustrated when talking to members of your family?	_____	_____	_____
S	Do you have difficulty hearing when someone speaks in a whisper?	_____	_____	_____
E	Do you feel handicapped by a hearing problem?	_____	_____	_____
S	Does a hearing problem cause you difficulty when visiting friends, relatives, or neighbors?	_____	_____	_____
S	Does a hearing problem cause you to attend religious services less often than you would like?	_____	_____	_____
E	Does a hearing problem cause you to have arguments with family members?	_____	_____	_____
S	Does a hearing problem cause you difficulty when listening to TV or radio?	_____	_____	_____
E	Do you feel that any difficulty with your hearing limits or hampers your personal or social life?	_____	_____	_____
S	Does a hearing problem cause you difficulty when in a restaurant with relatives or friends?	_____	_____	_____
	TOTAL SCORE = _____ (sum of the points assigned to each of the items)			

Free-Field Voice Test

This test was performed using a sound level meter to ensure that the whispered voice corresponded to a sound level of 30-45 dB, conversational voice to a sound level of 45-60 dB, and loud voice to a sound level of 60-80 dB [6]. The examiner was asked to practice her voice levels and compare it with the sound level meter every day to ensure that her whispered voice level, conversational voice level, and loud voice level were consistent. In this test, the examiner stands 0.6 meter away from the seated patient and whispers a combination of number and letters (eg, 2-R-9) and asks the patient to repeat the sequence [11] (Figure 2). If the patient repeats correctly, hearing is considered normal; if the patient repeats incorrectly, the test is repeated using a different number/letter combination. The patient is considered to have passed the test if he/she repeated at least 3 out of the possible 6 numbers or letters correctly. If the patient did not pass the whispered voice test, the patient was considered to have mild hearing impairment and the test was repeated with a conversational voice. If the patient could not pass the conversational voice test, the patient was considered to have moderate hearing impairment and the test was repeated with a loud voice. The patients were divided into 3 groups in the free-field voice test; understanding the whispered voice equated with normal hearing (25-dB threshold), understanding the conversational voice level (25-40 dB threshold) was considered as mild hearing impairment, and understanding loud voice levels (41-60 dB threshold) was considered as moderate or severe (61-80 dB threshold) hearing impairment. Patients with no response to any of these voice levels (>80 dB threshold) were considered to have profound hearing impairment [6,9,11-13]. Free-field voice test was performed by the same audiologist in the same soundproof room. The patient was asked to mask his/her nontest ear during the examination. We compared the results of the free-field voice test to those of the standard pure-tone audiometry in the better-hearing ear.

XSL•FO
RenderX

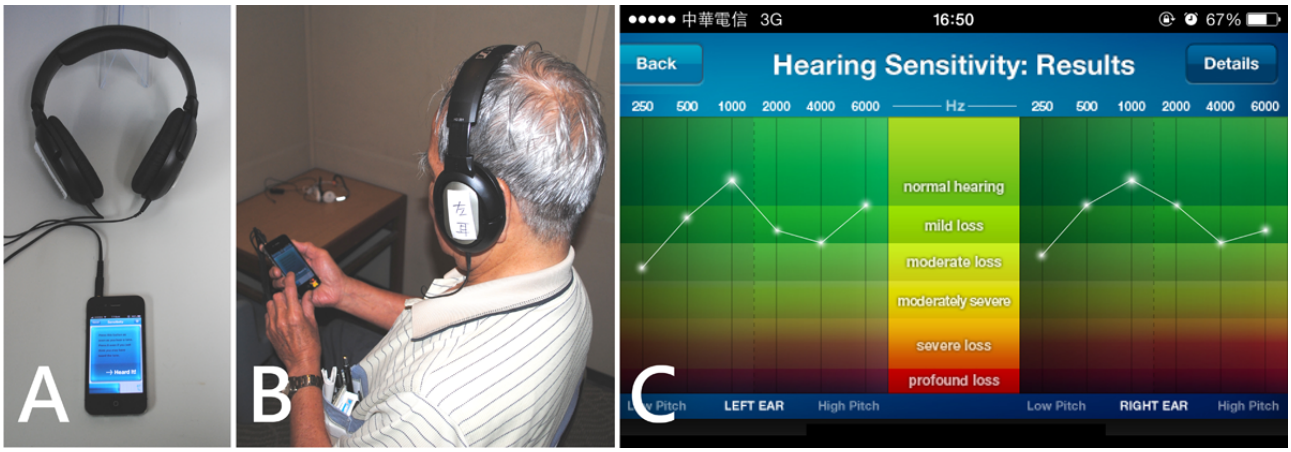
Figure 2. Demonstration of the free-field voice hearing screening test.



Smartphone-Based Audiometry

An iPhone 4S (Apple Inc) was used in our study. In order to avoid any possible effects of smartphone capabilities, Wi-Fi and 3G connectivity were turned off during the test. A free hearing app (uHear) was downloaded in the iPhone 4S from the iTune app store. The software enables patients to test their pure-tone air conduction hearing sensitivity. The software employs a simple “10 dB down and 5 dB up” approach [5,14]. The uHear app has the ability to determines the air-conducted sound at 250, 500, 1000, 2000, 4000, and 6000 Hz in both the left and right ears. The lowest threshold with 2 positive responses of 3 excursions was recorded as the hearing test. The hearing test can be completed in 6 minutes. Completing the hearing test required no learning curve [5,14]. At the end of the test, hearing sensitivity was shown in a typical audiogram format (Figure 3). The self-administered smartphone-based audiometry test was done in a soundproof room with an average ambient noise level of less than 35 dB hearing loss. Sennheiser HD201 headphones were used for all the patients. The patient was asked to press a large button on the touch screen to indicate when a sound was heard. Verbal instructions for the self-administered smartphone-based audiometry were presented by the same audiologist at the beginning of the test [5]. Smartphone-based audiometry test was performed by the patient in the same soundproof room. We compared the results of the smartphone-based audiometry test and those of the standard pure-tone audiometry in the better-hearing ear.

Figure 3. Smartphone-based audiometry hearing screening test. A: Smartphone-based audiometry device. B: Demonstration of smartphone-based audiometry. C: Results of hearing sensitivity shown in a typical audiogram format.



XSL•FO
RenderX

Standard Pure-Tone Audiometry

The Grason-Stadler GSI-61 Clinical Audiometer (AIC Medical Audiometric Instruments Corporation) and the Telephonics TDH-50P Audiometric Headphones were used to perform standard pure-tone audiometry. All clinical audiometers and accessory devices were calibrated as per ANSI S3.6, 1996 [15]. The standard pure-tone audiometry test was performed by the same audiologist in the same soundproof room. The degree of hearing loss was defined as the hearing thresholds in the better ear of the patients.

Data Analysis

Moderate hearing impairment was defined as the mean of the hearing thresholds at 500, 1000, 2000, and 4000 Hz >40 dB hearing loss (pure-tone average [PTA]>40 dB hearing loss). The results of the hearing tests were entered into 2 × 2 tables. Sensitivity, specificity, negative likelihood ratio, and the positive likelihood ratio were calculated.

Results

Demographic Data of the Patients

In our study, 41 patients were recruited, of which 27 were men and 14 were women. The mean (SD) age of the patients was 72.32 (6.81) years. The mean (SD) hearing thresholds at 500, 1000, 2000, and 4000 Hz of standard pure-tone audiometry was 36.29 (15.57) dB. Table 1 summarizes the demographic characteristics of the patients.

Table 1. Demographic data of the patients (N=41).

Characteristics of the patients	Values
Age (years)	
66-70, n (%)	22 (54)
71-75, n (%)	7 (17)
76-80, n (%)	6 (15)
81-85, n (%)	3 (7)
86-90, n (%)	3 (7)
Mean (SD) age	72.32 (6.81)
Gender, n (%)	
Male	27 (66)
Female	14 (34)
Better ear with pure-tone audiometry (dB)	
≤25, n (%)	12 (29)
26-40, n (%)	17 (42)
41-55, n (%)	6 (15)
56-70, n (%)	2 (5)
71-90, n (%)	3 (7)
≥91, n (%)	1 (2)
Standard pure-tone audiometry, mean (SD)	36.29 (15.57)
Smartphone-based audiometry, mean (SD)	43.4 (15.75)

Self-Perception Hearing Screening Test Findings

Table 2 compares the results of the self-perception hearing screening test and those of standard pure-tone audiometry. Of the 12 patients with moderate hearing impairment in the better ear (PTA>40 dB hearing loss) documented on standard pure-tone audiometry, 7 had a PTA>40 dB in the self-perception hearing screening test. This translates to a sensitivity of 0.58 (95% CI 0.29-0.84). Of the 29 patients without moderate hearing impairment in the better ear (PTA≤40 dB hearing loss) documented on standard pure-tone audiometry, 10 had a PTA≤40 dB in the self-perception hearing screening test. This translates to a specificity of 0.34 (95% CI 0.19-0.54). The positive likelihood ratio was 0.89 (95% CI 0.52-1.54). The negative likelihood ratio was 1.21 (95% CI 0.54-2.68).

XSL•FO
RenderX

Table 2. Accuracy of self-perception as a screening test compared to the standard pure-tone audiometry (N=41).

Self-perception test	Standard pure-tone audiometry	
	Patients with PTA[a]≤40 dB (n)	Patients with PTA>40 dB (n)
Patients with PTA≤40 dB (n)	10	5
Patients with PTA>40 dB (n)	19	7

[a]PTA: pure-tone average.

HHIE-S Questionnaire Results

Table 3 compares the results of the HHIE-S questionnaire hearing screening test and those of standard pure-tone audiometry. Of the 12 patients with moderate hearing impairment in better ear (PTA>40 dB hearing loss) documented on standard pure-tone audiometry, 8 had a PTA>40 dB in the HHIE-S questionnaire hearing screening test. This translates to a sensitivity of 0.67 (95% CI 0.35-0.89). Of the 29 patient without moderate hearing impairment in the better ear (PTA≤40 dB hearing loss) documented on standard pure-tone audiometry, 9 had a PTA≤0 dB in the HHIE-S questionnaire hearing screening test. This translates to a specificity of 0.31 (95% CI 0.316-0.51). The positive likelihood ratio was 0.97 (95% CI 0.60-1.54). The negative likelihood ratio was 1.07 (95% CI 0.42-2.78).

Table 3. Accuracy of the HHIE-S questionnaire as a screening test compared to the standard pure-tone audiometry (N=41).

HHIE-S questionnaire	Standard pure-tone audiometry	
	Patients with PTA[a]≤40 dB (n)	Patients with PTA>40 dB (n)
Patients with PTA≤40 dB (n)	9	4
Patients with PTA>40 dB (n)	20	8

[a]PTA: pure-tone average.

Free-Field Voice Test Findings

Table 4 compares the results of the free-field voice hearing screening test and those of standard pure-tone audiometry. Of the 12 patients with moderate hearing impairment in better ear (PTA>40 dB hearing loss) documented on standard pure-tone audiometry, 10 had a PTA>40 dB in the free-field voice hearing screening test. This translates to a sensitivity of 0.83 (95% CI 0.51-0.97). Of the 29 patients without moderate hearing impairment in better ear (PTA≤40 dB hearing loss) documented on standard pure-tone audiometry, 12 had a PTA≤40 dB in the free-field voice test hearing screening test. This translates to a specificity of 0.41 (95% CI 0.24-0.61). The positive likelihood ratio was 1.42 (95% CI 0.96-2.11). The negative likelihood ratio was 0.40 (95% CI 0.10-1.57).

Table 4. Accuracy of free-field voice test as a screening test compared to the standard pure-tone audiometry (N=41).

Free-field voice test	Standard pure-tone audiometry	
	Patients with PTA[a]≤40 dB (n)	Patients with PTA>40 dB (n)
Patients with PTA≤40 dB (n)	12	2
Patients with PTA>40 dB (n)	17	10

[a]PTA: pure-tone average.

Smartphone-Based Audiometry Findings

Table 5 compares the results of the smartphone-based audiometry hearing screening test and those of standard pure-tone audiometry. Of the 12 patients with moderate hearing impairment in the better ear (PTA>40 dB hearing loss) documented on standard pure-tone audiometry, 11 had a PTA>40 dB in the smartphone-based audiometry hearing screening test. This translates to a sensitivity of 0.92 (95% CI 0.60-0.99). Of the 29 patients without moderate hearing impairment in the better ear (PTA≤40 dB hearing loss) documented on standard pure-tone audiometry, 22 had a PTA≤40 dB in the smartphone-based audiometry hearing screening test. This translates to a specificity of 0.76 (95% CI 0.56-0.89). The positive likelihood ratio was 3.80 (95% CI 1.95-7.4). The negative likelihood ratio was 0.11 (95% CI 0.02-0.73).

XSL•FO
RenderX

Table 5. Accuracy of smartphone-based audiometry as a screening test compared to the standard pure-tone audiometry (N=41).

Smartphone-based audiometry	Standard pure-tone audiometry (n)	
	Patients with PTA[a]≤40 dB (n)	Patients with PTA>40 dB (n)
Patients with PTA≤40 dB (n)	22	1
Patients with PTA>40 dB (n)	7	11

[a]PTA: pure-tone average.

Discussion

More than 50% of the world's population live in low-income countries. People with hearing impairments in low-income countries account for 80% of the world's associated population [16]. Thus, most patients with hearing impairments live in the low-income countries. However, audiology services are overlooked in low-income countries, because these countries are struggling to provide even the basic medical services in order to avoid other life-threatening diseases. There is a surge of need for audiology services in these limited-income countries, because most people in these countries have neither access to an audiologist nor any form of hearing health care. According to Fagan [16], audiology services are nonexistent in most African countries or there is only a single audiologist attending to millions of people. It is estimated that there are 4 audiologists for 100,000 people in the United States. The ratio is almost the same in the United Kingdom. In Taiwan, the data from the Mackay Medical College estimated that there is only 1 audiologist per 100,000 people. Thus, audiology services are unequally distributed across the world.

In industrialized countries, well-established audiology equipment is available for audiology practice. The huge gap in the provision of audiology services between industrialized and low-income countries can barely be filled by volunteer audiologists or by developing new hearing screening methods. Therefore, we aimed to develop some fast, easy-to-use, and reliable methods for low-cost hearing screening tests. All the different methods shown in our study for screening hearing impairments have advantages and disadvantages. Smartphone-based audiometry has higher sensitivity and higher specificity than the other screening methods for detecting moderate hearing impairment in the older adults.

The first screening method for hearing impairment in this study was the self-perception hearing screening test. Self-reported data to assess the presence of diseases or disorders have been used frequently in large-scale epidemiologic survey studies in the past [7]. Studies [17] have shown that patients have been screened for hearing loss by using self-perception screening questions, which involves asking the patient whether they feel they have hearing loss, and these questions are used in the measurements of global health. Several questions can be asked, for example, "Do you have a hearing problem now?" or "Would you say you have any difficulty hearing now?" In this paper, we recommend using the question, "Do you have a hearing problem now?" as a measure of the global health in the annual medical history forms in geriatric practices [17]. A positive response to this question is considered as a positive screening for hearing loss. A total of 26 patients thought that they had hearing loss, of which 7 were tested positive for moderate hearing impairment. This single-question self-perception hearing screening test had a sensitivity of 0.58 (95% CI 0.29-0.84) and specificity of 0.34 (95% CI 0.19-0.54), with a positive likelihood ratio of 0.89 (95% CI 0.52-1.54) and negative likelihood ratio of 1.21 (95% CI 0.54-2.68). This result implied that an older person's self-perception of a hearing problem could not reliably indicate the presence of a hearing impairment. Compared to the results in our study, the results of Gates et al [17] had higher sensitivity (0.71, 95% CI 0.63-0.78) and higher specificity (0.72, 95% CI 0.67-0.76). Cultural views of hearing impairment and disability can affect the answer to the question "Do you have a hearing problem now?" Attitudes to hearing impairment are influenced by social behavior, economic situation, and education levels, and these factors would affect the results of the self-perception hearing screening test.

The second screening method for hearing impairment in this study was the HHIE-S questionnaire hearing screening test. The screening version of the HHIE-S questionnaire is the most widely applied test that is validated. The HHIE-S questionnaire is a 10-item, self-administered questionnaire that was developed to measure the emotional and social handicap [7]. The ASHA-suggested fail-criteria of 10 points or more equals to moderate hearing impairment. A total of 28 patients scored more than 10 points in the HHIE-S questionnaire, of which 8 tested positive for moderate hearing impairment. This HHIE-S questionnaire hearing screening test had a sensitivity of 0.67 (95% CI 0.35-0.89) and specificity of 0.31 (95% CI 0.316-0.51), with a positive likelihood ratio of 0.97 (95% CI 0.60-1.54) and negative likelihood ratio of 1.07 (95% CI 0.42-2.78). Yueh et al showed that the sensitivity of the HHIE-S questionnaire hearing screening test ranges from 0.53 to 0.80 and the specificity ranges from 0.67 to 0.75 [18]. The sensitivity of the single-question self-perception hearing screening test was only 0.58 in our study. The 10-item HHIE-S questionnaire hearing screening test, which is related to daily activities, includes only a single question, which increases the sensitivity of the HHIE-S questionnaire to 0.67. Both single-question self-perception and the 10-item HHIE-S questionnaire hearing screening tests are used for assessing the subjective hearing status. However, our results suggest that these tests may be less effective in screening early stages of hearing impairment.

The use of speech to determine hearing impairment level has been used for a long time. The hearing impairment level can be assessed by a primary health care doctor by using the free-field voice hearing screening test. It is surprising that the free-field voice hearing screening test is still performed nowadays. In an industrialized country like United Kingdom, primary health care doctors are obliged to screen the older adult population for

XSL•FO
RenderX

hearing impairment, as listed in the 1990 National Health Service contract [19]. The primary health doctor uses the free-field voice hearing screening test as the first examination for hearing impairment in the United Kingdom. We used a sound level meter to calibrate the voice of the examiner. In order to ensure that the whispered voice level, conversational voice level, and loud voice level were consistent, the examiner was asked to calibrate her voice levels every day. A total of 27 patients were considered to have moderate hearing impairment in the free-field voice hearing screening test, of which 10 were tested positive for moderate hearing impairment. This free-field voice hearing screening test had a sensitivity of 0.83 (95% CI 0.51-0.97) and specificity of 0.41 (95% CI 0.24-0.61), with a positive likelihood ratio of 1.42 (95% CI 0.96-2.11) and negative likelihood ratio of 0.40 (95% CI 0.10-1.57). McShefferty et al [20] showed that the sensitivity of the free-field voice hearing screening test was 0.56 and that the specificity was 0.65. The results of our study showed higher sensitivity and lower specificity. The potential reason for our findings could be explained by the calibration of the sound level, which resulted in higher sensitivity.

In recent years, there has been a rapid evolution in the development of mobile health services. Smartphone-based audiometry is considered as a fast, easy, and reliable technique for cost-effective screening of hearing impairments. Free apps can be downloaded from the Android market and installed in any smartphone. Smartphones have the ability to control the audio output and frequency. However, given that there are difficulties with calibration, in terms of sound output levels and timing uncertainty, the applicability of smartphones and standard pure-tone audiometry is limited. A total of 18 patients were considered to have moderate hearing impairment in the smartphone-based audiometry hearing screening test, of which 11 tested positive for moderate hearing impairment. The smartphone-based audiometry hearing screening test had a sensitivity of 0.92 (95% CI 0.60-0.99) and specificity of 0.76 (95% CI 0.56-0.89), with a positive likelihood ratio of 3.80 (95% CI 1.95-7.4) and negative likelihood ratio of 0.11 (95% CI 0.02-0.73). Szudek et al [5] showed that the sensitivity of smartphone-based audiometry was 0.98 and the specificity was 0.82. The results of our study showed lower sensitivity and specificity compared to those reported in previous studies [21-27], probably because our study population consisted of adults older than 65 years. Older populations may not be good at operating a smartphone. Despite the above limitations, our study shows that smartphone-based audiometry is a reasonable hearing test for diagnosing moderate hearing impairment. Smartphone-based audiometry could be the next trend for screening hearing impairment in older adults, especially in low-income countries. Our study specifically excluded people with hearing aids or those who received standard pure-tone audiogram evaluation in the prior 24 months, because the purpose of this study was to detect older adults with unrecognized moderate hearing impairment.

In conclusion, all the different methods mentioned in this study are available for hearing screening but they all have advantages and disadvantages. Smartphone-based audiometry has higher sensitivity and higher specificity for detecting moderate hearing impairment in older populations. Prospective randomized clinical studies are still needed to test the application of smartphone-based audiometry in the field of screening for hearing impairment in older populations. Nevertheless, we recommend that in low-income countries where specialized audiological services and proven audiometric equipment are not available, primary health care givers should be trained to administer simple smartphone-based audiometry as an acceptable alternative hearing assessment for older adults who have complaints with regard to hearing impairment.

Acknowledgments

We wish to thank all the patients and their families for participating in this study.

Authors' Contributions

CW conceived and designed the experiments. CW and TW performed the experiments. LYJL analyzed the data. LYJL, SW, TW, and CT contributed the reagents/materials/analysis tools. LYJL drafted the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Hearing Handicap Inventory for the Elderly-Screening Questionnaire in Chinese.
[PNG File , 303 KB - mhealth_v8i10e17213_app1.png]

References

1. Prevention of blindness and deafness. World Health Organization. URL: http://www.who.int/pbd/en [accessed 2020-10-03]
2. WHO global estimates on prevalence of hearing loss. World Health Organization. URL: http://www.who.int/pbd/deafness/WHO_GE_HL.pdf [accessed 2020-10-03]
3. Stevens G, Flaxman S, Brunskill E, Mascarenhas M, Mathers CD, Finucane M, Global Burden of Disease Hearing Loss Expert Group. Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. Eur J Public Health 2013 Feb;23(1):146-152. [doi: 10.1093/eurpub/ckr176] [Medline: 22197756]

XSL•FO
RenderX

4. Sound Hearing 2030 Report 2012. Sound Hearing 2030 website. URL: http://www.soundhearing2030.org [accessed 2020-10-03]
5. Szudek J, Ostevik A, Dziegielewski P, Robinson-Anagor J, Gomaa N, Hodgetts B, et al. Can Uhear me now? Validation of an iPod-based hearing loss screening test. J Otolaryngol Head Neck Surg 2012 Apr;41 Suppl 1:S78-S84. [Medline: 22569055]
6. Prescott C, Omoding S, Fermor J, Ogilvy D. An evaluation of the 'voice test' as a method for assessing hearing in children with particular reference to the situation in developing countries. International Journal of Pediatric Otorhinolaryngology 1999 Dec;51(3):165-170 [FREE Full text] [doi: 10.1016/s0165-5876(99)00263-3]
7. Bagai A, Thavendiranathan P, Detsky AS. Does this patient have hearing impairment? JAMA 2006 Jan 25;295(4):416-428. [doi: 10.1001/jama.295.4.416] [Medline: 16434632]
8. Benson V, Marano M. Current estimates from the National Health Interview Survey, 1992. Vital Health Stat 10 1994 Jan(189):1-269 [FREE Full text] [Medline: 8154108]
9. Yueh B, Collins M, Souza P, Boyko E, Loovis C, Heagerty P, et al. Long-term effectiveness of screening for hearing loss: the screening for auditory impairment--which hearing assessment test (SAI-WHAT) randomized trial. J Am Geriatr Soc 2010 Mar;58(3):427-434. [doi: 10.1111/j.1532-5415.2010.02738.x] [Medline: 20398111]
10. Wu HY, Chin JJ, Tong HMH. Screening for hearing impairment in a cohort of elderly patients attending a hospital geriatric medicine service. Singapore Med J 2004 Feb;45(2):79-84 [FREE Full text] [Medline: 14985847]
11. Pirozzo S, Papinczak T, Glasziou P. Whispered voice test for screening for hearing impairment in adults and children: systematic review. BMJ 2003 Oct 25;327(7421):967 [FREE Full text] [doi: 10.1136/bmj.327.7421.967] [Medline: 14576249]
12. Boatman DF, Miglioretti DL, Eberwein C, Alidoost M, Reich SG. How accurate are bedside hearing tests? Neurology 2007 Apr 17;68(16):1311-1314. [doi: 10.1212/01.wnl.0000259524.08148.16] [Medline: 17438223]
13. Chou R, Dana T, Bougatsos C, Fleming C, Beil T. Screening adults aged 50 years or older for hearing loss: a review of the evidence for the U.S. preventive services task force. Ann Intern Med 2011 Mar 01;154(5):347-355. [doi: 10.7326/0003-4819-154-5-201103010-00009] [Medline: 21357912]
14. Handzel O, Ben-Ari O, Damian D, Priel MM, Cohen J, Himmelfarb M. Smartphone-based hearing test as an aid in the initial evaluation of unilateral sudden sensorineural hearing loss. Audiol Neurootol 2013;18(4):201-207. [doi: 10.1159/000349913] [Medline: 23689282]
15. American NSI. American National Standards on Acoustics. In: American National Standard Specification for Audiometers. Washington (DC), USA: National Academies Press; 1996:A.
16. Fagan JJ. Developing World ENT: a global responsibility. J. Laryngol. Otol 2012 Mar 30;126(6):544-547 [FREE Full text] [doi: 10.1017/s0022215112000345]
17. Gates GA, Murphy M, Rees TS, Fraher A. Screening for handicapping hearing loss in the elderly. J Fam Pract 2003 Jan;52(1):56-62. [Medline: 12540314]
18. Yueh B, Shapiro N, MacLean CH, Shekelle PG. Screening and management of adult hearing loss in primary care: scientific review. JAMA 2003 Apr 16;289(15):1976-1985. [doi: 10.1001/jama.289.15.1976] [Medline: 12697801]
19. Williams EI, Wallace P. Health checks for people aged 75 and over. Occas Pap R Coll Gen Pract 1993 Apr(59):1-30 [FREE Full text] [Medline: 8199581]
20. McShefferty D, Whitmer WM, Swan IRC, Akeroyd MA. The effect of experience on the sensitivity and specificity of the whispered voice test: a diagnostic accuracy study. BMJ Open 2013;3(4). [doi: 10.1136/bmjopen-2012-002394] [Medline: 23604349]
21. Chu Y, Cheng Y, Lai Y, Tsao Y, Tu T, Young ST, et al. A Mobile Phone-Based Approach for Hearing Screening of School-Age Children: Cross-Sectional Validation Study. JMIR Mhealth Uhealth 2019 Apr 01;7(4):e12033 [FREE Full text] [doi: 10.2196/12033] [Medline: 30932870]
22. Yimtae K, Israsena P, Thanawirattananit P, Seesutas S, Saibua S, Kasemsiri P, et al. A Tablet-Based Mobile Hearing Screening System for Preschoolers: Design and Validation Study. JMIR Mhealth Uhealth 2018 Oct 23;6(10):e186 [FREE Full text] [doi: 10.2196/mhealth.9560] [Medline: 30355558]
23. Bright T, Pallawela D. Validated Smartphone-Based Apps for Ear and Hearing Assessments: A Review. JMIR Rehabil Assist Technol 2016 Dec 23;3(2):e13 [FREE Full text] [doi: 10.2196/rehab.6074] [Medline: 28582261]
24. Masalski M, Kipiński L, Grysiński T, Kręcicki T. Hearing Tests on Mobile Devices: Evaluation of the Reference Sound Level by Means of Biological Calibration. J Med Internet Res 2016 May 30;18(5):e130 [FREE Full text] [doi: 10.2196/jmir.4987] [Medline: 27241793]
25. Masalski M, Grysiński T, Kręcicki T. Biological calibration for web-based hearing tests: evaluation of the methods. J Med Internet Res 2014 Jan 15;16(1):e11 [FREE Full text] [doi: 10.2196/jmir.2798] [Medline: 24429353]
26. Masalski M, Grysiński T, Kręcicki T. Hearing Tests Based on Biologically Calibrated Mobile Devices: Comparison With Pure-Tone Audiometry. JMIR Mhealth Uhealth 2018 Jan 10;6(1):e10 [FREE Full text] [doi: 10.2196/mhealth.7800] [Medline: 29321124]
27. Wu C, Wu S, Chen P, Lin Y. An innovative smartphone-based otorhinoendoscope and its application in mobile health and teleotolaryngology. J Med Internet Res 2014 Mar 03;16(3):e71 [FREE Full text] [doi: 10.2196/jmir.2959] [Medline: 24590187]

XSL•FO
RenderX

Abbreviations

ASHA: American Speech-Language-Hearing Association
HHIE-S: Hearing Handicap Inventory for the Elderly-Screening
PTA: pure-tone average

Edited by G Eysenbach; submitted 26.11.19; peer-reviewed by LT Li, S Akbarian; comments to author 24.12.19; revised version received 30.01.20; accepted 17.09.20; published 27.10.20.

Please cite as:
Li LYJ, Wang SY, Wu CJ, Tsai CY, Wu TF, Lin YS
Screening for Hearing Impairment in Older Adults by Smartphone-Based Audiometry, Self-Perception, HHIE Screening Questionnaire, and Free-Field Voice Test: Comparative Evaluation of the Screening Accuracy With Standard Pure-Tone Audiometry
JMIR Mhealth Uhealth 2020;8(10):e17213
URL: http://mhealth.jmir.org/2020/10/e17213/
doi:10.2196/17213
PMID:33107828

©Lok Yee Joyce Li, Shin-Yi Wang, Cheng-Jung Wu, Cheng-Yu Tsai, Te-Fang Wu, Yaoh-Shiang Lin. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 27.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An Interactive Text Messaging Intervention to Improve Adherence to Option B+ Prevention of Mother-to-Child HIV Transmission in Kenya: Cost Analysis

Yilin Chen[1], MPH; Keshet Ronen[1], MPH, PhD; Daniel Matemo[2], MPH; Jennifer A Unger[1], MPH, MD; John Kinuthia[2], MMed, MPH; Grace John-Stewart[1], MPH, MD, PhD; Carol Levin[1], MSc, PhD

[1]Department of Global Health, University of Washington, Seattle, WA, United States

[2]Kenyatta National Hospital, Nairobi, Kenya

Corresponding Author:
Carol Levin, MSc, PhD
Department of Global Health, University of Washington
1959 NE Pacific St
Seattle, WA, 98105
United States
Phone: 1 206 744 3790
Email: clevin@uw.edu

Abstract

Background: Mobile health (mHealth) approaches offer potentially affordable ways to support the care of HIV-infected patients in overstretched health care systems. However, only few studies have analyzed the costs associated with mHealth solutions for HIV care.

Objective: The aim of this study was to estimate the total incremental costs and incremental cost per beneficiary of an interactive SMS text messaging support intervention within a clinical trial.

Methods: The Mobile WAChX trial (NCT02400671) evaluates an interactive semiautomated SMS text messaging intervention to improve adherence to antiretroviral therapy and retention in care among peripartum women infected with HIV in Kenya to reduce the mother-to-child transmission of HIV. Women were randomized to receive one-way versus two-way SMS text messages. Messages were sent weekly, and these messages included motivational and educational content and visit reminders; two-way messaging enabled prompt consultation with the nurse as needed. Microcosting methods were used to collect resource-use data related to implementing the Mobile WAChX SMS text messaging intervention. At 2 sites (Nairobi and Western Kenya), we conducted semistructured interviews with health personnel to identify startup and recurrent activities by obtaining information on the personnel, supplies, and equipment. Data on expenditures and prices from project expense reports, administrative records, and published government salary data were included to estimate the total incremental costs. Using a public provider perspective, we estimated incremental unit costs per beneficiary and per contact during 2017.

Results: The weighted average annual incremental costs for the two-way SMS text messaging group were US $3725 per facility, US $62 per beneficiary, and US $0.85 per contact to reach 115 beneficiaries. For the one-way SMS text messaging group, the weighted average annual incremental costs were US $2542 per facility, US $41 per beneficiary, and US $0.66 per contact to reach 117 beneficiaries. The largest cost shares were for the personnel: 48.2% (US $1794/US $3725) in two-way and 32.4% (US $825/US $2542) in one-way SMS text messaging groups. Costs associated with software development and communication accounted for 29.9% (US $1872/US $6267) of the costs in both intervention arms (US $1042 vs US $830, respectively).

Conclusions: Cost information for budgeting and financial planning is relevant for implementing mHealth interventions in national health plans. Given the proportion of costs related to systems development, it is likely that costs per beneficiary will decline with the scale-up of the interventions.

(JMIR Mhealth Uhealth 2020;8(10):e18351) doi:10.2196/18351

KEYWORDS

mHealth; cost analysis; prevention of mother-to-child transmission; antiretroviral therapy adherence; Kenya; mobile phone

XSL•FO
RenderX

Introduction

In 2017, an estimated 180,000 children became infected with HIV [1]. The Joint United Nations Program on HIV and AIDS has set the ambitious target of reducing mother-to-child transmission (MTCT) to 20,000 cases by 2020. A critical factor required to achieve this goal is sustained adherence to antiretroviral therapy (ART) by pregnant and breastfeeding women living with HIV. Suboptimal adherence and virologic suppression in pregnant and postpartum women have been documented in several contexts, and these factors have significantly increased the risk of MTCT and poor maternal outcomes. Thus, development of strategies to support ART adherence and engagement in care has been identified as a priority.

Use of SMS text messaging communication is a promising approach for improving ART adherence in peripartum women. Several studies and meta-analyses have shown that regularly delivered SMS text messages can improve ART adherence and retention in care outside of pregnancy [2-5] and that interactive, two-way SMS text messaging between patients and health care providers is more efficacious than one-way informational messaging [6,7]. Based on these data, SMS text messaging has been identified as a recommended intervention to promote ART adherence by the World Health Organization [8]. In 2016, the Ministry of Health (MOH) of Kenya rolled out a mobile health (mHealth) service "Ushauri" in 105 facilities to manage appointments and deliver standardized messages and reminders to patients with HIV, which was shown to be a promising strategy in improving viral suppression and retention in routine HIV care [9].

While SMS text messaging interventions are supported by policymakers to improve ART adherence, little is known about the value for money of such technologies. In a global survey conducted by the World Health Organization, lack of evidence on economic evaluation was identified as a major barrier to implementation of mHealth solutions in resource-constrained settings [10]. One of the reasons was the limited availability of data on the cost of implementing SMS text messaging interventions [11,12]. Only 1 study has performed a cost analysis of one-way SMS text messaging promoting the availability of HIV self-testing kits in Kenyan clinics [13]. We found no cost analysis data on implementing two-way SMS text messaging interventions during our literature review. While the Kenyan MOH is currently planning nationwide scale-up of mHealth interventions to meet universal health coverage (UHC) goals [9], it is essential to understand the costs for budgeting purposes and for designing efficient and affordable programs that can be scaled nationally [14]. We therefore conducted a cost analysis of one-way and two-way SMS text messaging interventions (vs control) in the Mobile WAChX study, which is a randomized controlled trial of semiautomated SMS text messaging interventions to improve ART adherence and retention in care in peripartum women at 6 facilities in Nairobi and Western Kenya [15]. The primary objective of this study was to estimate the total incremental costs, incremental cost per beneficiary, and cost per contact associated with the SMS text messaging interventions. The secondary objective was to estimate the costs of alternative implementation and usage scenarios for future scale-up of the interventions.

Methods

Study Design

The Mobile WAChX trial (NCT02400671) evaluates one-way versus two-way communication versions of a semiautomated SMS text messaging intervention to improve ART adherence and retention in care in peripartum women at 6 facilities in Nairobi and Western Kenya. One-way SMS text messaging consisted of weekly automated motivational and educational SMS text messages and clinic visit reminders. Participants randomized to the two-way SMS text messaging arm additionally had the capability of communicating with a nurse through the SMS text messaging system. Participants randomized to the control arm received no SMS text messages (standard of care).

The study procedures for the randomized controlled trial are described in detail in a previous paper [15]. For costing analysis, convenience sampling was used to collect the cost data from 2 of the 6 sites: an urban health center in Nairobi (facility A) and a rural subcounty hospital in Western Kenya (facility B). We used a provider perspective to analyze the incremental financial and economic costs of implementing Mobile WAChX alongside existing maternal and child health (MCH) services from January 2017 to December 2017, as well as the average incremental cost per beneficiary and average incremental cost per contact. Beneficiaries are service users who accessed Mobile WAChX SMS text messaging communication, including automated messages and personalized messaging with nurses, through the Mobile WAChX system. The number of contacts is equal to the number of total messages sent successfully to the beneficiaries. Data on the number of beneficiaries and contacts at each facility were extracted from ongoing project monitoring and evaluation reports and electronic databases of SMS text messages.

Data Collection Method

We used an activity-based ingredients approach to identify all activities undertaken to deliver the Mobile WAChX project. Activities included intervention planning, project partner sensitization, staff training, system development, and delivery of SMS text messages. After identifying all intervention-related activities, we quantified the resources used and valued these by using the best available data on salaries and commodity prices. We used a combination of data collection methods to collect primary resource-use and cost data, including obtaining prices from project expense reports, administrative records, and published government salaries, and conducting semistructured interviews with facility-based health workers and project administrators. Time-motion studies were conducted in both facilities to record staff time spent on intervention activities (eg, recruiting, screening, and registering participants, sending SMS text messages to users). Data collection was conducted between October 2017 and January 2018.

XSL•FO
RenderX

Cost Categories

We organized cost data into one-time fixed costs and variable costs (Table 1). Fixed costs were categorized according to the following activities: intervention planning, preparation of intervention sites, development of the Mobile WAChX SMS text messaging management system, initial training, and sensitization meetings with facility staff and Kenyan MOH officials. All fixed costs were used for one-time start-up activities, where we assumed a 5-year useful life. Variable costs included recurrent costs, which were required to sustain the intervention. They were divided into mutually exclusive input categories for personnel, communication, equipment, and overhead. Personnel costs included salaries, benefits, and allowances of facility-based health workers as well as of staff in charge of personnel supervision and coordination. A study nurse and a retention officer delivered Mobile WAChX-related activities. Shared facility costs included a study coordinator and a data manager who were responsible for supervising all the 6 sites and some shared communications costs such as the costs of developing a Mobile WAChX system and cost of the system-hosting platform. We allocated these shared facility costs based on the annual share of clients served at the facility, as a percentage of the total annual number of clients served for all 6 facilities. In addition to the cost of a system-hosting platform, other communication costs included cost of internet data bundles for using the internet-based Mobile WAChX system, airtime cost for making patient follow-up phone calls, costs of sending SMS text messages to users, and short code toll-free numbers used by participants to deliver messages to the nurse. Equipment costs (mobile phones, laptops, and furniture) were annuitized over the useful life of 10 years by using a discount rate of 3%. All costs in this evaluation were expressed in USD, using the official exchange rate of 1 USD to 103.25 Kenyan Shillings (2017 exchange rate). In addition, research time and other research costs were removed from the costing analysis. International staff time was also excluded to better reflect the costs of the program when implemented and scaled locally.

Table 1. Activity and input cost categories and description.

Cost categories, subcategories	Description
Fixed costs	
Planning/microplanning	Planning activities for project implementation during the start-up period.
System development	Resources and inputs to design the Mobile WAChX system and activities to collaborate with a local mobile technology company to obtain SMS text messaging packages for participants.
Initial training	Expenses for conducting 2 training workshops during the start-up period for all project staff, including development of relevant training materials.
Sensitization	Stakeholder workshops and activities at facility level.
Variable costs	
Personnel	Value of personnel time
Service delivery	Activities for delivering the Mobile WAChX intervention, such as recruiting participants, screening and registering participants, and sending SMS text messages to users.
Personnel supervision and coordination	Meetings to supervise staff and coordinate and monitor implementation of activities at all sites.
Communication	Resources and inputs to deliver SMS text messages to participants, including data bundles for internet, an online platform for hosting the Mobile WAChX system, airtime cost for phone calls, and SMS text messaging cost.
Equipment	Investments that last longer than 1 year, including mobile phones, laptops, and furniture.
Overhead	Clinic collaboration fee and indirect costs.

Data Analysis

We developed an Excel-based model (Microsoft Excel version 15.28, Redmond) to estimate total incremental costs and incremental unit costs. The sum of all the activity cost categories reflects all the resources required to deliver the Mobile WAChX intervention. All activities were mutually exclusive, thereby avoiding double counting. We used project output data on the number of beneficiaries per month per facility and number of messages sent by the system and nurses per month. These data were collected as part of the Mobile WAChX monitoring and evaluation strategy. We first estimated the total incremental cost for each facility to deliver the Mobile WAChX intervention and divided this by the number of women receiving the intervention to determine the cost per beneficiary. We also estimated the cost per contact, defined as the total cost divided by the total number of messages sent to participants during 2017, for each facility. We analyzed the costs of one-way and

XSL•FO
RenderX

two-way SMS text messaging interventions separately. We also estimated the average weighted costs for both facilities by using project output data on the number of beneficiaries per facility.

Scenario Analysis

In addition to estimating intervention costs, we estimated 3 scenarios. The first scenario was to estimate the cost when the two-way Mobile WAChX intervention was implemented in all 6 facilities in this project. In this scenario, the same intervention and personnel were applied to every facility where they shared most start-up costs from system development, cost of system-hosting platform, as well as personnel supervision and coordination costs. The number of participants was the total number of participants receiving intervention in all 6 sites. The second scenario was to estimate a more typical scenario where the MOH of Kenya supports these activities after the pilot phase. We applied the MOH salary scale to service delivery health workers in the current project. Finally, we calculated the incremental cost effectiveness ratios (ICERs) of comparing two-way SMS text messages to no intervention in the 2 facilities. Due to unavailable efficacy data in the Mobile WAChX trial, we used clinical outcome results from a similar randomized controlled trial, which assessed whether two-way SMS text messaging interventions improved plasma HIV-1 viral RNA load suppression at 12 months in 3 clinics in Kenya [7]. The efficacy of mobile SMS text messaging intervention for medication adherence was obtained from a meta-analysis study [16]. Other inputs were using project output data and our analyses.

Results

Project Output

Among the 152 HIV-infected women in facility A who received the Mobile WAChX intervention, 76 (50.0%) were randomized to the two-way SMS text messaging group and the other 76 (50.0%) were randomized to the one-way SMS text messaging group. A total of 80 women participated in facility B, of whom 39 were randomized to the two-way SMS text messaging intervention group and 41 were randomized to the one-way SMS text messaging intervention group (Table 2). The beneficiaries in these 2 facilities accounted for 27.7% (152/548) and 14.6% (80/548) of the total number of project beneficiaries at the time of data collection in facilities A and B, respectively. The total number of automated SMS text messages from the system sent to beneficiaries in the 2 facilities in 2017 was 6924 in the two-way SMS text messaging intervention group and 7318 in the one-way SMS text messaging intervention group. The nurses in the 2 facilities sent 1386 personalized messages in response to two-way participant messages during 2017.

Table 2. Summary of the beneficiaries and the total points of contact by health facility in 2017.

Health facility	Two-way SMS text messaging intervention			One-way SMS text messaging intervention		
	Beneficiaries, n=115, n (%)	Total automated SMS text messages, n=6924, n (%)	Total nurse SMS text messages, n=1386, n (%)	Beneficiaries, n=117, n (%)	Total automated SMS text messages, n=7318, n (%)	Total nurse SMS text messages, n=0, n (%)
Facility A (Urban health center)	76 (66.1)	4425 (63.9)	993 (71.6)	76 (65.0)	4604 (62.9)	0 (0)
Facility B (Rural subcounty hospital)	39 (33.9)	2499 (36.1)	393 (28.4)	41 (35.0)	2714 (37.1)	0 (0)

Total Costs and Unit Costs

Table 3 summarizes the weighted average total annual incremental costs and incremental unit costs for beneficiaries in 2017. The weighted average total cost of the Mobile WAChX intervention was estimated at US $3725 for the two-way SMS text messaging group and US $2542 for the one-way SMS text messaging group. Fixed costs were US $2936 for two-way SMS text messaging and US $1757 for one-way SMS text messaging intervention groups (Figure 1), while variable costs were similar across the 2 groups (US $789 and US $785, respectively). The weighted average cost per beneficiary for the two-way group was higher than that for the one-way SMS text messaging intervention group (US $62 and US $41, respectively) (Figure 2). In addition, the weighted average cost per contact was estimated at US $0.85 for the two-way and US $0.66 for the one-way SMS text messaging intervention group (Figure 3). The detailed annual incremental costs and cost per beneficiary by input categories for each facility and the weighted average estimates are presented in Multimedia Appendix 1, Multimedia Appendix 2, and Multimedia Appendix 3.

Table 3. Weighted average total annual incremental costs and unit costs for beneficiaries.

Intervention group	Total costs and unit costs (USD)		
	Total annual cost	Cost per beneficiary	Cost per contact
Two-way SMS text messaging	$3725	$62	$0.85
One-way SMS text messaging	$2542	$41	$0.66

XSL•FO
RenderX

Figure 1. Total annual incremental costs by fixed and variable costs.

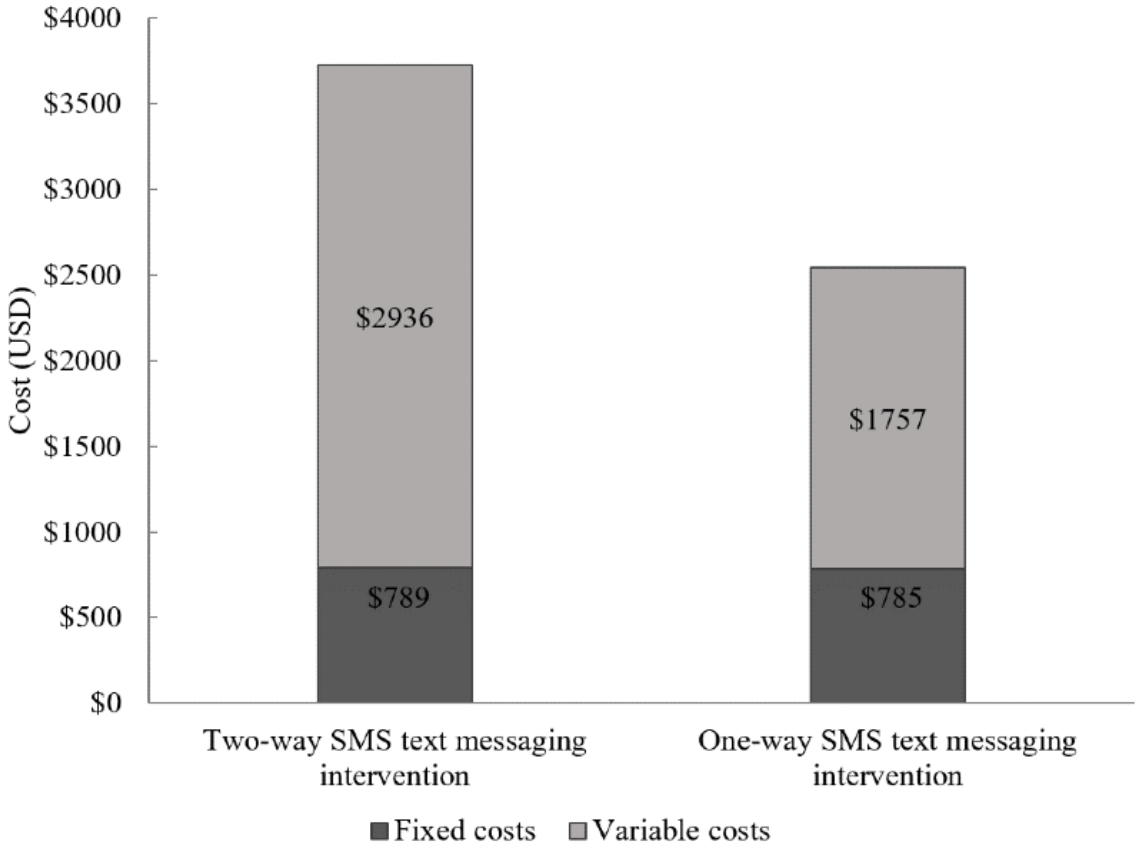
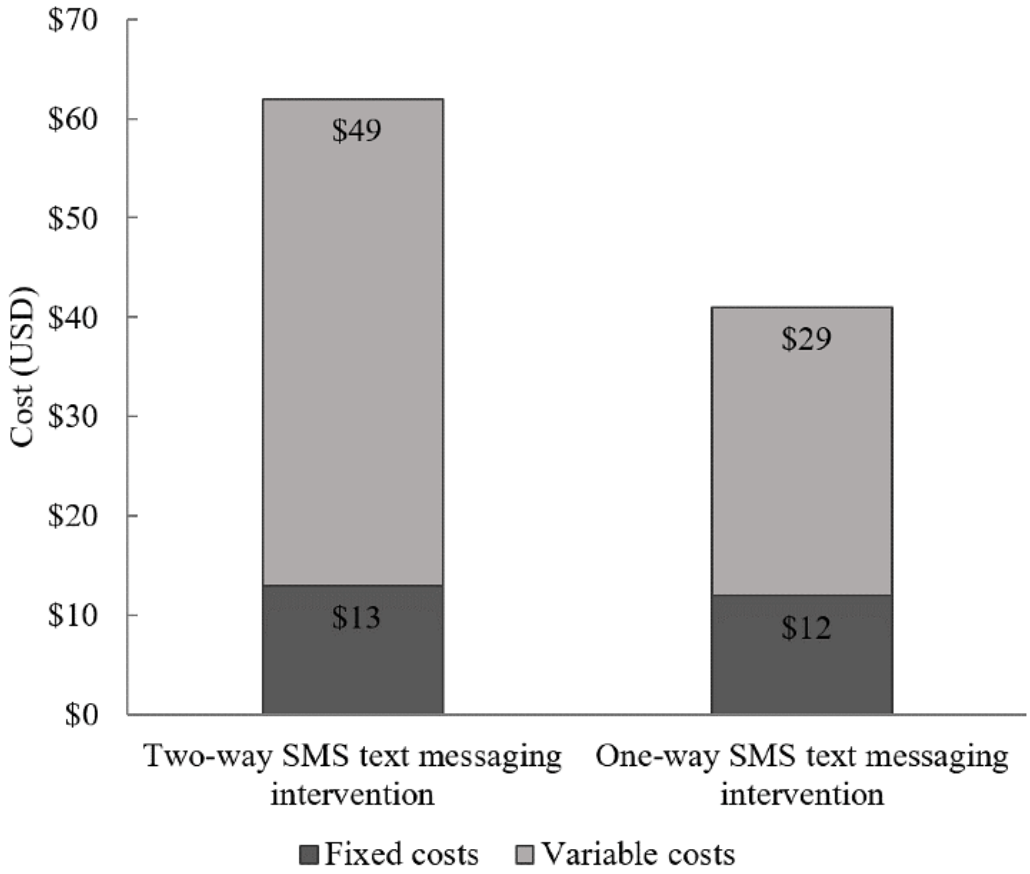
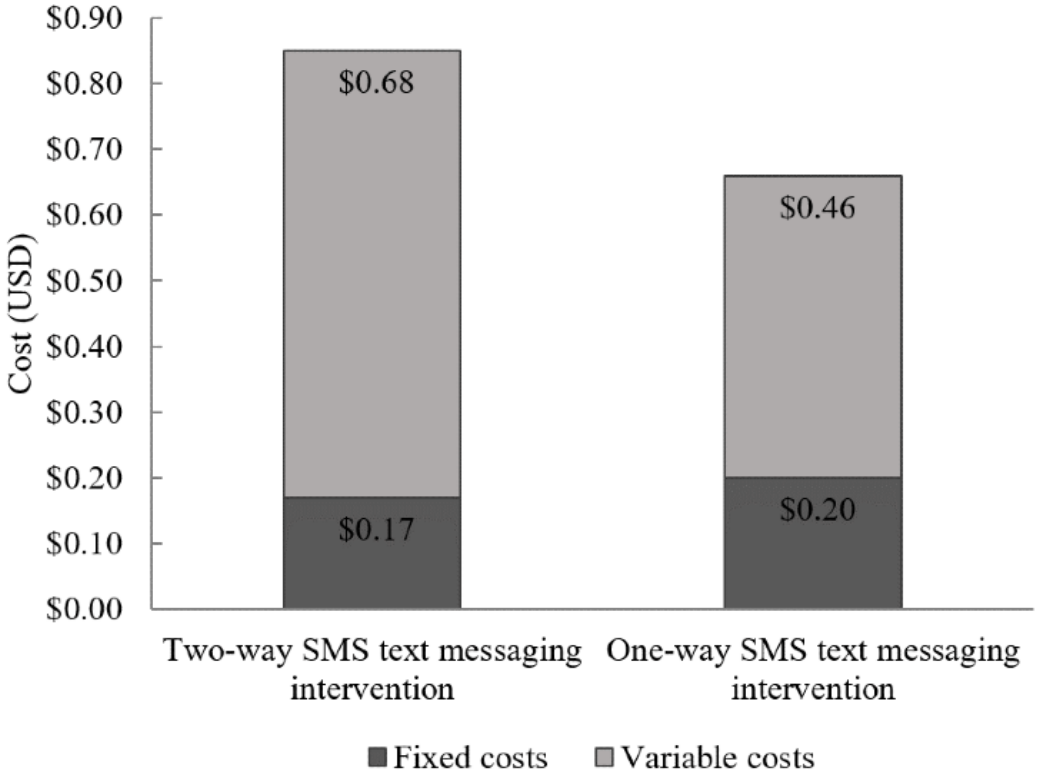


Figure 2. Cost per beneficiary by fixed and variable costs.



XSL•FO
RenderX

Figure 3. Cost per contact by fixed and variable costs.



Cost Profiles

Figure 4 presents the cost shares by activity and input cost categories. Personnel cost accounted for the largest share of the total costs. The personnel cost in the one-way SMS text messaging group (US $825/US $2542, 32.4%) was lower than that of the two-way SMS text messaging group (US $1794/US $3725, 48.2%), as one-way messaging required no personnel time responding to messages. The second largest share of the total costs was related to software development of the SMS text messaging management system (two-way: US $573/US $3725, 15.4% vs one-way: US $569/US $2542, 22.4%) and communication (two-way: US $470/US $3725, 12.6% vs one-way: US $261/US $2542, 10.3%) such as flat rate fees for internet usage and mobile phone minutes. In the two-way SMS text messaging group, an estimated 10.2% (US $381/US $3725) of the total costs was used for purchasing equipment and mobile devices to set up the intervention site in the facility and to implement the Mobile WAChX intervention. Overhead costs accounted for 7.8% (US $291/US $3725) of the total costs in the two-way SMS text messaging group, in the form of a clinic collaboration fee to use a clinic room for study activities (US $48 per month). Two trainings were conducted at the beginning of the project to teach nurses and retention officers how to use the Mobile WAChX system and orient them to implement the intervention, which costs US $98 per facility (3% of the total costs). The remaining budget was allocated to sensitization (US $64) and microplanning (US $54), which accounted for 3.2% (US $118/US $3725) of the total costs. The main cost drivers were similar in both two-way and one-way SMS text messaging intervention groups. The variable cost shares by input categories are shown in Figure 5. The majority of the total costs were related to personnel in both two-way and one-way SMS text messaging intervention groups (US $1794/US $2936, 61.1% vs US $825/US $1757, 46.9%, respectively), followed by communication (US $470/US $2936, 16.0% vs US $261/US $1757, 14.9%, respectively), equipment (US $381/US $2936, 13.0% vs US $381/US $1757, 21.7%, respectively), and overhead (US $291/US $2936, 9.9% vs US $291/US $1757, 16.5%, respectively).

XSL•FO
RenderX

Figure 4. Cost shares by input categories for all costs.

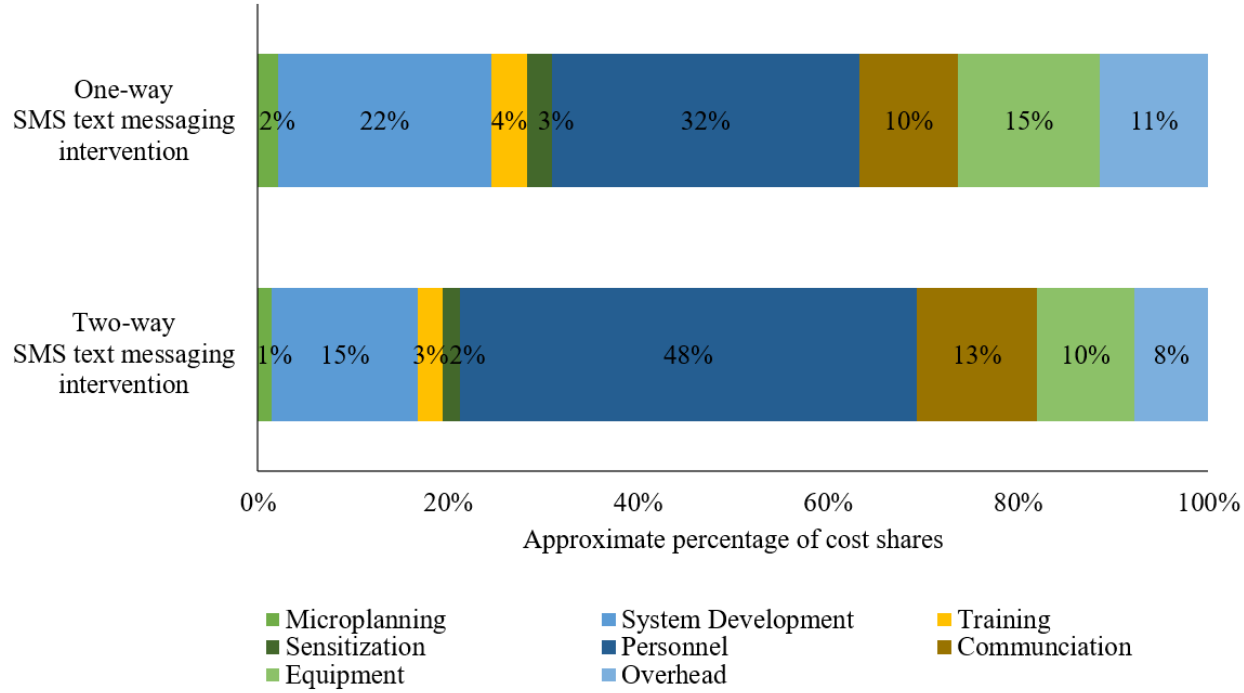
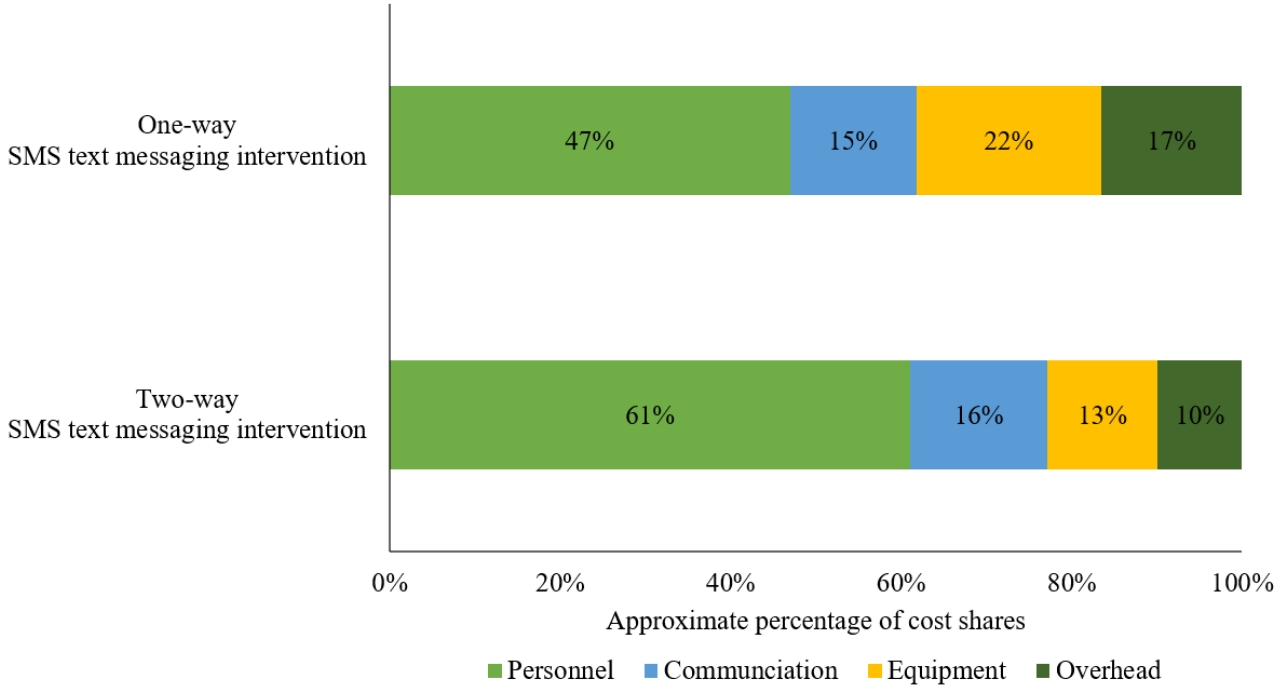


Figure 5. Cost shares for variable input categories only.



Sensitivity and Scenario Analysis

Cost estimates in the 2 scenarios compared to the baseline scenario estimates reflected in Table 3 are presented in Multimedia Appendix 4. Expanding the two-way SMS text messaging intervention to all 6 facilities would have the effect of decreasing the estimated cost per beneficiary by 31% (from US $62 to US $43; US $19/US $62) and cost per contact by 26% (from US $0.85 to US $0.63; US $0.22/US $0.85). When building on scenario 1 and replacing research project health worker salaries with the more typical salaries for MOH staff, the cost per beneficiary decreased by 50% (from US $62 to US $31, US $31/US $62) and the cost per contact decreased by 46% (from US $0.85 to US $0.46, US $0.39/US $0.85).

The ICERs of comparing two-way SMS text messaging versus no intervention in the 2 facilities and the input parameters are summarized in Multimedia Appendix 5 and Multimedia Appendix 6, respectively. According to published RCT and meta-analysis estimates, implementing two-way SMS text

XSL•FO
RenderX

messaging intervention was estimated to have suppressed viral load in 11 patients and achieved medication adherence among 15 patients, with total annual incremental costs of US $7084 [7,16]. Thus, the ICERs were US $644 per viral load suppression and US $472 per medication adherence.

Discussion

This is the first study to estimate the costs of an mHealth intervention to promote prevention of MTCT (PMTCT)-ART adherence among peripartum women. Previous studies have shown that SMS text messaging interventions have a positive impact on ART adherence and maternal and neonatal health outcomes in low-and-middle income settings [4,17-20]. However, data regarding the programmatic costs of implementing these mHealth interventions are scant [4,18,21]. We estimate the average total incremental costs for 1 year of project implementation to be US $3725 per facility, US $62 per beneficiary, and US $0.85 per contact for the two-way SMS text messaging intervention. For the one-way SMS text messaging intervention, the average total incremental costs are US $2542 per facility, US $41 per beneficiary, and US $0.66 per contact. The higher costs for the two-way SMS text messaging group is due to the personnel time spent responding to SMS text messages, which is not provided in the one-way SMS text messaging group.

Only a few studies have provided the cost estimates of mHealth interventions, including SMS text messaging interventions, in low-income countries. The MAMA program was initiated in 2012 in South Africa to enhance the utilization of MCH services among pregnant and postpartum women by sending registered users SMS text messages twice per week. The estimated program costs over 5 years was US $1.2 million, 17% of which was incurred by costs on program development and 31% on SMS text message delivery costs [22]. The Chipatala cha pa foni (CCPF) project consisted of a toll-free hotline and a mobile phone–based tips and reminders service seeking to improve maternal and neonatal health in Balaka District, Malawi [11]. The tips and reminders service was a one-way messaging system, wherein community health workers sent weekly text or voice messages to participants. Service users could call the hotline as well. The costs during a 2-year period (2011-2012) were estimated to be US $29.33 per user and US $4.33 per successful contact. The ReMiND (Reducing Maternal and Newborn Deaths) Project was designed to improve the quality of counseling of community health workers in India [12]. The mHealth app was implemented through 259 accredited social health activist workers. The total program costs over 3 years (2012-2015) were estimated at US $191,894, with US $20.50 per registered woman. Labor costs accounted for 57% of the total costs, followed by mobile phone purchases and data/internet charges (6%). The annual number of beneficiaries were 9798 and 9390 in CCPF and ReMiND projects, respectively. Both CCPF and ReMiND studies had lower cost per user compared to our study because of the higher number of beneficiaries. Moreover, the ReMiND project did not involve any SMS text message exchange between patients and health workers, which decreases data usage and airtime-related costs. In addition, we found that the higher costs for the two-way SMS text messaging intervention largely resulted from higher personnel costs. As previous studies have demonstrated that two-way SMS text messaging interventions are likely to be more efficacious than one-way SMS text messaging interventions, the higher cost may contribute to health service utilization and better health outcomes. Compared with the cost estimates in the CCPF project, our estimated costs per contact were much lower, both in the two-way and one-way SMS text messaging interventions (US $0.85 and US $0.66, respectively), thereby suggesting overall higher utilization of messaging interventions in our study.

Personnel costs accounted for the largest share of the total costs in our project (US $1794/US $3725, 48.2% in two-way and US $825/US $2542, 32.4% in one-way SMS text messaging intervention groups), followed by software development of the SMS text messaging management system and communication costs. Our results are consistent with findings from previous studies that labor costs for delivering other SMS text messaging interventions were the main drivers of the total program costs, followed by SMS text messaging program development [11,12,22]. How could we potentially reduce system development and communication-related costs as well as personnel costs? Our scenario analysis results suggest that the costs would be reduced significantly by expanding the two-way intervention to more beneficiaries (>500), with cost per beneficiary decreasing from US $62 to US $43 and cost per contact from US $0.85 to US $0.63. Although recurrent communication costs such as airtime and cost of SMS text messages would increase due to increased number of beneficiaries and messages exchanged, the total unit cost would eventually go down because the fixed start-up costs associated with system development and the system-hosting platform would be shared across a larger number of beneficiaries. In addition, the shared program costs for personnel supervision and coordination costs (>60% of total costs) would be allocated across a higher number of beneficiaries. In the second scenario analysis, we explored what a more typical government-sponsored program might cost. Keeping the activities the same, if we replaced health worker salaries in our project with the more typical salaries for MOH staff, the unit costs would go down further to US $31 per beneficiary and US $0.46 per contact. This means moving from a research-focused project to part of a routine government-supported program will result in much lower cost per beneficiary through greater economies of scale and through a more typical mode of service delivery and supervision.

Our findings provide important costing information for budgeting and financial planning for implementing mHealth interventions to achieve UHC in Kenya. mHealth interventions, as a part of the broader eHealth interventions, may become transformational strategies in addressing public health challenges and striving toward UHC in Kenya. This need has also been reaffirmed by the Kenya's Health Policy (2014-2030), National eHealth Policy (2016-2030), Information Technology Master Plan, and the Health Bill [7-10]. While the Kenyan MOH is currently planning for nationwide scale-up of mHealth systems, affordability cannot be ignored. Assuming 80% mobile phone penetration and that 59,000 women were offered PMTCT

XSL•FO
RenderX

services in Kenya [23,24], we could estimate that scaling two-way SMS text messaging intervention to increase adherence would result in total spending of US $1.46 million with the average cost per beneficiary of US $31. This is likely to be affordable as it only comprises 3% of the 2015-2016 PMTCT budget of US $46 million. In addition, our estimated ICERs per patient with viral load suppressed (US $644) and per patient achieving medication adherence (US $472) were both much lower than the 2017 gross domestic product per capita in Kenya (US $1568). Our study showed that, with expanded coverage, the total unit cost will decrease significantly due to shared fixed start-up costs associated with system development and the system-hosting platform. In the future, market forces and private sector could also be harnessed to achieve affordability and sustainability [11]. With the increasingly active participation of the private sector in public health in Kenya, public-private partnership could be explored to leverage the infrastructure and resources of private sectors to help Kenya achieve UHC. For example, there are innovative ways to lower the costs of communication such as by obtaining discounted SMS text messaging packages from local or foreign telecommunication companies.

To our knowledge, this is the first study to estimate the costs of an mHealth intervention targeting PMTCT and MCH in Kenya. Our study has limitations. First, we did not include the labor costs of international collaborators who contributed to the program design and installation and other start-up activities such as work planning meetings. Second, we had to make assumptions about the allocation of communication costs to research and implementation activities, which may influence the total implementation costs. Third, we did not allocate equipment costs to research activities, which may lead to overestimation of the total implementation costs. Equipment such as laptops and phones were mainly used for service delivery, and we do not have detailed information on how much equipment was used to support the research-related activities. Therefore, the evidence should be used with the consideration that after taking into account the equipment costs allocated to research activities, the total incremental cost of implementing SMS text messaging interventions would be lower than the estimates in this study.

In conclusion, this study fills the knowledge gap on the costs of mHealth approaches for improving PMTCT-ART adherence among pregnant women in Kenya. When operating at scale, there may be opportunities to reduce the costs per beneficiary. As the Mobile WAChX intervention is scaled up, further research is needed to understand the economic impact from different perspectives, including cost-utility analyses to assess the value for money compared with alternative approaches to improve women's clinical outcomes and adherence to HIV treatment as part of PMTCT services.

Acknowledgments

We would like to acknowledge the project staff and health care workers who participated in the data collection and supported us during this activity. This research received administrative support and mentorship from the Global Center for Integrated Health of Women, Adolescents and Children (Global WACh), which is jointly supported by the Departments of Global Health, Pediatrics, and Obstetrics and Gynecology. This study was supported by the following grants: National Institute of Allergy and Infectious Diseases (P30AI027757; Principal Investigator: Dr. Jared Baeten and K01AI116298; Principal Investigator: Dr. Alison Drake). Support was also received from the National Institute of Child Health and Human Development (R01HD080460; Principal Investigator: Dr. Grace John-Stewart and K24HD054314; Principal Investigator: Dr. Grace John-Stewart).

Authors' Contributions

YC, CL, GJS, KR, and JU contributed to the study design. YC conducted primary data collection and data analyses. JK and DM coordinated data collection in the field. YC, CL, and KR wrote and revised the initial drafts of the manuscript. All authors contributed to manuscript revisions and interpretation of study results.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Annual incremental costs and cost per beneficiary by activity at facility A.
[DOCX File , 19 KB - mhealth_v8i10e18351_app1.docx]

Multimedia Appendix 2
Annual incremental costs and cost per beneficiary by activity at facility B.
[DOCX File , 19 KB - mhealth_v8i10e18351_app2.docx]

Multimedia Appendix 3
Weighted average annual incremental costs and cost per beneficiary by activity at facility A and facility B.
[DOCX File , 19 KB - mhealth_v8i10e18351_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Sensitivity and scenario analysis: total annual costs and unit costs for beneficiaries.
[DOCX File , 14 KB - mhealth_v8i10e18351_app4.docx]

Multimedia Appendix 5
Incremental cost-effectiveness ratios of two-way SMS versus no intervention in viral load suppression and adherence.
[DOCX File , 17 KB - mhealth_v8i10e18351_app5.docx]

Multimedia Appendix 6
Input parameters for incremental cost-effectiveness ratios of two-way SMS versus no intervention in viral load suppression and adherence.
[DOCX File , 16 KB - mhealth_v8i10e18351_app6.docx]

References

1. UNAIDS. Miles to go—closing gaps, breaking barriers, righting injustices. 2018 Aug 13. URL: https://www.unaids.org/en/20180718_GR2018 [accessed 2020-09-12] [WebCite Cache ID https://www.unaids.org/en/20180718_GR2018]
2. Mayer JE, Fontelo P. Meta-analysis on the effect of text message reminders for HIV-related compliance. AIDS Care 2017 Apr;29(4):409-417 [FREE Full text] [doi: 10.1080/09540121.2016.1214674] [Medline: 27477580]
3. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database Syst Rev 2012 Mar 14(3):CD009756 [FREE Full text] [doi: 10.1002/14651858.CD009756] [Medline: 22419345]
4. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014;9(2):e88166 [FREE Full text] [doi: 10.1371/journal.pone.0088166] [Medline: 24505411]
5. Mills EJ, Lester R, Thorlund K, Lorenzi M, Muldoon K, Kanters S, et al. Interventions to promote adherence to antiretroviral therapy in Africa: a network meta-analysis. Lancet HIV 2014 Dec;1(3):e104-e111 [FREE Full text] [doi: 10.1016/S2352-3018(14)00003-4] [Medline: 26424119]
6. Wald DS, Butt S, Bestwick JP. One-way versus two-way text messaging on improving medication adherence: meta-analysis of randomized trials. Am J Med 2015 Oct;128(10):1139.e1-1139.e5. [doi: 10.1016/j.amjmed.2015.05.035] [Medline: 26087045]
7. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
8. WHO. 2016. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection. URL: https://www.who.int/hiv/pub/arv/arv-2016/en/ [accessed 2020-09-12]
9. National AIDS & STI Control Programme. Implementing a National mHealth System Improves Viral Suppression and Retention in Routine HIV Care in Kenya. URL: http://regist2.virology-education.com/presentations/2019/13INTEREST/12_Wachira.pdf [accessed 2020-08-12]
10. WHO. mHealth: New Horizons for Health through Mobile Technologies. Geneva: Global Observatory for eHealth Series; 2011. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf? [accessed 2020-08-12]
11. Larsen-Cooper E, Bancroft E, Rajagopal S, O'Toole M, Levin A. Scale Matters: A Cost-Outcome Analysis of an m-Health Intervention in Malawi. Telemed J E Health 2016 Apr;22(4):317-324 [FREE Full text] [doi: 10.1089/tmj.2015.0060] [Medline: 26348994]
12. Prinja S, Bahuguna P, Gupta A, Nimesh R, Gupta M, Thakur JS. Cost effectiveness of mHealth intervention by community health workers for reducing maternal and newborn mortality in rural Uttar Pradesh, India. Cost Eff Resour Alloc 2018;16:25 [FREE Full text] [doi: 10.1186/s12962-018-0110-2] [Medline: 29983645]
13. George G, Chetty T, Strauss M, Inoti S, Kinyanjui S, Mwai E, et al. Costing analysis of an SMS-based intervention to promote HIV self-testing amongst truckers and sex workers in Kenya. PLoS One 2018;13(7):e0197305 [FREE Full text] [doi: 10.1371/journal.pone.0197305] [Medline: 29979704]
14. LeFevre AE, Shillcutt SD, Broomhead S, Labrique AB, Jones T. Defining a staged-based process for economic and financial evaluations of mHealth programs. Cost Eff Resour Alloc 2017;15:5 [FREE Full text] [doi: 10.1186/s12962-017-0067-6] [Medline: 28428734]
15. Drake AL, Unger JA, Ronen K, Matemo D, Perrier T, DeRenzi B, et al. Evaluation of mHealth strategies to optimize adherence and efficacy of Option B+ prevention of mother-to-child HIV transmission: Rationale, design and methods of a 3-armed randomized controlled trial. Contemp Clin Trials 2017 Jun;57:44-50 [FREE Full text] [doi: 10.1016/j.cct.2017.03.007] [Medline: 28315480]

XSL•FO
RenderX

16. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile Telephone Text Messaging for Medication Adherence in Chronic Disease: A Meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
17. Chen H, Chai Y, Dong L, Niu W, Zhang P. Effectiveness and Appropriateness of mHealth Interventions for Maternal and Child Health: Systematic Review. JMIR Mhealth Uhealth 2018 Jan 09;6(1):e7 [FREE Full text] [doi: 10.2196/mhealth.8998] [Medline: 29317380]
18. Cooper V, Clatworthy J, Whetham J, Consortium E. mHealth Interventions To Support Self-Management In HIV: A Systematic Review. Open AIDS J 2017;11:119-132 [FREE Full text] [doi: 10.2174/1874613601711010119] [Medline: 29290888]
19. Hurt K, Walker RJ, Campbell JA, Egede LE. mHealth Interventions in Low and Middle-Income Countries: A Systematic Review. Glob J Health Sci 2016 Sep 01;8(9):54429 [FREE Full text] [doi: 10.5539/gjhs.v8n9p183] [Medline: 27157176]
20. Sondaal SFV, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the Effect of mHealth Interventions in Improving Maternal and Neonatal Care in Low- and Middle-Income Countries: A Systematic Review. PLoS One 2016;11(5):e0154664. [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
21. de la Torre-Díez I, López-Coronado M, Vaca C, Aguado JS, de Castro C. Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: a systematic review. Telemed J E Health 2015 Feb;21(2):81-85 [FREE Full text] [doi: 10.1089/tmj.2014.0053] [Medline: 25474190]
22. LeFevre A, Cabrera-Escobar MA, Mohan D, Eriksen J, Rogers D, Neo Parsons A, et al. Forecasting the Value for Money of Mobile Maternal Health Information Messages on Improving Utilization of Maternal and Child Health Services in Gauteng, South Africa: Cost-Effectiveness Analysis. JMIR Mhealth Uhealth 2018 Jul 27;6(7):e153 [FREE Full text] [doi: 10.2196/mhealth.8185] [Medline: 30054263]
23. Roxana Elliott. Mobile Phone Penetration Throughout Sub-Saharan Africa. 2019 Jul 08. URL: https://www.geopoll.com/blog/mobile-phone-penetration-africa/ [accessed 2020-08-11]
24. UNAIDS. Prevention Gap Report. 2016. URL: https://www.unaids.org/sites/default/files/media_asset/2016-prevention-gap-report_en.pdf [accessed 2020-08-11]

Abbreviations

ART: antiretroviral therapy
CCPF: Chipatala cha pa foni
ICER: incremental cost effectiveness ratio
MCH: maternal and child health
mHealth: mobile health
MOH: Ministry of Health
MTCT: mother-to-child transmission
PMTCT: prevention of mother-to-child transmission
ReMiND: Reducing Maternal and Newborn Deaths
UHC: universal health coverage

Edited by G Eysenbach; submitted 21.02.20; peer-reviewed by H Ide, P Zhang; comments to author 12.06.20; revised version received 06.08.20; accepted 17.08.20; published 02.10.20.

Please cite as:
Chen Y, Ronen K, Matemo D, Unger JA, Kinuthia J, John-Stewart G, Levin C
An Interactive Text Messaging Intervention to Improve Adherence to Option B+ Prevention of Mother-to-Child HIV Transmission in Kenya: Cost Analysis
JMIR Mhealth Uhealth 2020;8(10):e18351
URL: https://mhealth.jmir.org/2020/10/e18351
doi:10.2196/18351
PMID:33006562

©Yilin Chen, Keshet Ronen, Daniel Matemo, Jennifer A Unger, John Kinuthia, Grace John-Stewart, Carol Levin. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 02.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development and Usability of a Novel Interactive Tablet App (PediAppRREST) to Support the Management of Pediatric Cardiac Arrest: Pilot High-Fidelity Simulation-Based Study

Francesco Corazza[1], MD; Deborah Snijders[1], MD; Marta Arpone[1], PhD; Valentina Stritoni[2], MD, PhD; Francesco Martinolli[1], MD; Marco Daverio[2], MD; Maria Giulia Losi[3], MA; Luca Soldi[3], BSc; Francesco Tesauri[3], PhD; Liviana Da Dalt[1], MD; Silvia Bressan[1], MD, PhD

[1]Division of Pediatric Emergency Medicine, Department of Women's and Children's Health, University of Padova, Padova, Italy

[2]Pediatric Intensive Care Unit, Department of Women's and Children's Health, University of Padova, Padova, Italy

[3]RE:Lab S.r.l., Reggio Emilia, Italy

Corresponding Author:
Francesco Corazza, MD
Division of Pediatric Emergency Medicine
Department of Women's and Children's Health
University of Padova
Via Giustiniani 3
Padova, 35128
Italy
Phone: 39 0498218015
Email: francesco.corazza@unipd.it

Abstract

Background: Pediatric cardiac arrest (PCA), although rare, is associated with high mortality. Deviations from international management guidelines are frequent and associated with poorer outcomes. Different strategies/devices have been developed to improve the management of cardiac arrest, including cognitive aids. However, there is very limited experience on the usefulness of interactive cognitive aids in the format of an app in PCA. No app has so far been tested for its usability and effectiveness in guiding the management of PCA.

Objective: To develop a new audiovisual interactive app for tablets, named PediAppRREST, to support the management of PCA and to test its usability in a high-fidelity simulation-based setting.

Methods: A research team at the University of Padova (Italy) and human–machine interface designers, as well as app developers, from an Italian company (RE:Lab S.r.l.) developed the app between March and October 2019, by applying an iterative design approach (ie, design–prototyping–evaluation iterative loops). In October–November 2019, a single-center nonrandomized controlled simulation–based pilot study was conducted including 48 pediatric residents divided into teams of 3. The same nonshockable PCA scenario was managed by 11 teams with and 5 without the app. The app user's experience and interaction patterns were documented through video recording of scenarios, debriefing sessions, and questionnaires. App usability was evaluated with the User Experience Questionnaire (UEQ) (scores range from –3 to +3 for each scale) and open-ended questions, whereas participants' workload was measured using the NASA Raw-Task Load Index (NASA RTLX).

Results: Users' difficulties in interacting with the app during the simulations were identified using a structured framework. The app usability, in terms of mean UEQ scores, was as follows: attractiveness 1.71 (SD 1.43), perspicuity 1.75 (SD 0.88), efficiency 1.93 (SD 0.93), dependability 1.57 (SD 1.10), stimulation 1.60 (SD 1.33), and novelty 2.21 (SD 0.74). Team leaders' perceived workload was comparable (P=.57) between the 2 groups; median NASA RTLX score was 67.5 (interquartile range [IQR] 65.0-81.7) for the control group and 66.7 (IQR 54.2-76.7) for the intervention group. A preliminary evaluation of the effectiveness of the app in reducing deviations from guidelines showed that median time to epinephrine administration was significantly longer in the group that used the app compared with the control group (254 seconds versus 165 seconds; P=.015).

Conclusions: The PediAppRREST app received a good usability evaluation and did not appear to increase team leaders' workload. Based on the feedback collected from the participants and the preliminary results of the evaluation of its effects on the management of the simulated scenario, the app has been further refined. The effectiveness of the new version of the app in

XSL•FO
RenderX

reducing deviations from guidelines recommendations in the management of PCA and its impact on time to critical actions will be evaluated in an upcoming multicenter simulation-based randomized controlled trial.

(JMIR Mhealth Uhealth 2020;8(10):e19070) doi:10.2196/19070

KEYWORDS

cardiac arrest; resuscitation; mobile app; high-fidelity simulation training; cognitive aid; pediatrics; emergency medicine

Introduction

Pediatric cardiac arrest (PCA), although rare, is an important public health issue due to its high mortality and morbidity, its complex time-dependent management and emotional burden, its social and economic costs, and differences with adult cardiac arrest (CA) [1-5].

International guidelines by relevant societies are periodically updated to help health professionals provide the best evidence-based basic and advanced care to improve the management and outcome of PCA [6-10]. Nevertheless, deviations from guideline recommendations occur frequently in the management of CA [11-19] and are associated with poorer clinical outcomes [20,21].

Different cognitive support tools have been developed and tested in order to improve adherence to guideline-recommended management of both adult CA and PCA, with variable results. Most of these tools are devices that provide real-time audiovisual feedback on the quality of chest compressions. Such tools have shown to be effective in improving the quality of compressions [22-25]. Several tools have been conceived to guide bystanders' management of out-of-hospital CA (OHCA) through audio/video support by means of mobile phones [26]. Augmented reality glasses have also been studied to communicate with a remote intensivist to support the management of in-hospital PCA [27] or to display Pediatric Advanced Life Support (PALS) guidelines to the team leader during a PCA scenario [28] with partial benefit in improving resuscitation performance. In addition, numerous apps have been created to support the management of CA, mainly for OHCA. However, a recent systematic review outlined how the majority of these apps are not tested for content, usability, and effectiveness, even though many have already been released and are available on app stores [29].

Research on apps developed and tested to guide the management of in-hospital PCA is very limited [30], and to the best of our knowledge, no app has so far been tested for both its usability and its effectiveness in guiding the management of PCA.

Based on the deviations from guidelines recorded on a prior study conducted by our research team in PCA simulation scenarios [31], we set out to design, develop, and test a new interactive multimodal (audio–visual) cognitive aid in the format of a tablet app.

The primary aim of this pilot study was to refine the app and to test its usability and impact on team leader's workload using high-fidelity simulation. As a secondary aim, we explored the trend in the occurrence of deviations from guidelines.

Methods

App Development

We designed and developed an app for tablet that we named PediAppRREST, which is the result of the collaboration between a pediatric research team, including physicians and researchers from the Pediatric Emergency Department and the Pediatric Intensive Care Unit of the University Hospital of Padova (Italy), and human–machine interface designers, human factor experts, and app/software developers of RE:Lab S.r.l., an Interaction Engineering company (Reggio Emilia, Italy).

The app was designed to guide the team leader to perform resuscitation interventions in the sequence/timing and modality reported by the American Heart Association (AHA) PALS 2015 guidelines [6-8]. In the design process we took into account the results of a prior study conducted by our research team [31], which assessed deviations from guidelines in PCA simulation scenarios managed by pediatric residents without the use of any cognitive aid.

We developed the app between March and October 2019. As a first step, the research team defined the actions to be displayed in separate screens, the flow/pathways, and the additional features that were deemed helpful to guide resuscitation and achieve a high-quality cardiopulmonary resuscitation (CPR), based on recommended PALS guidelines/algorithms. We then progressively refined and validated the cognitive aid following an iterative prototyping development approach [32]. In the development phase, serial testing of the app by research staff revealed bugs and highlighted the need for refinements concerning the information layout and organization, the user interface navigation flow, and the naming conventions. Bugs and re-design suggestions were implemented by the Interaction Engineering company.

App Description

Directions on recommended interventions, following the order reported in the PALS algorithms, are sequentially displayed in the app which has been designed as a *checklist* app. Indeed, progression to the next screen is allowed once the recommended actions are tapped by the user, to indicate they have been read and likely performed.

The main criteria applied in the user experience design phase of the app have been (1) timely information (each screen gathers only the necessary information for each phase of the PALS algorithm, communicating it both visually and acoustically, with the aim of reducing the load on the team leader's working memory and relying on a multichannel communication); (2) priority (actions [ie, epinephrine administration] triggered by timers have priority on other actions displayed on the screen);

XSL•FO
RenderX

and (3) sequential versus alternative choices (decisions that team leaders must take into consideration concern actions and choices to be performed sequentially or alternatively). Hence, sequential actions are displayed with rectangular buttons, aligned vertically on the page, whereas alternative choices are organized with square buttons, aligned horizontally on the screen (Figure 1). The app is currently in Italian, but a multilingual version is under development (an English translation of the screen content is herein provided to ease the understanding of this article).

Each screen is structured into 3 zones (Figure 2): zone 1 (top bar), zone 2 (main area), and zone 3 (bottom bar).

Figure 1. Sequential versus alternative choices. CPR: cardiopulmonary resuscitation; ROSC: return of spontaneous circulation.

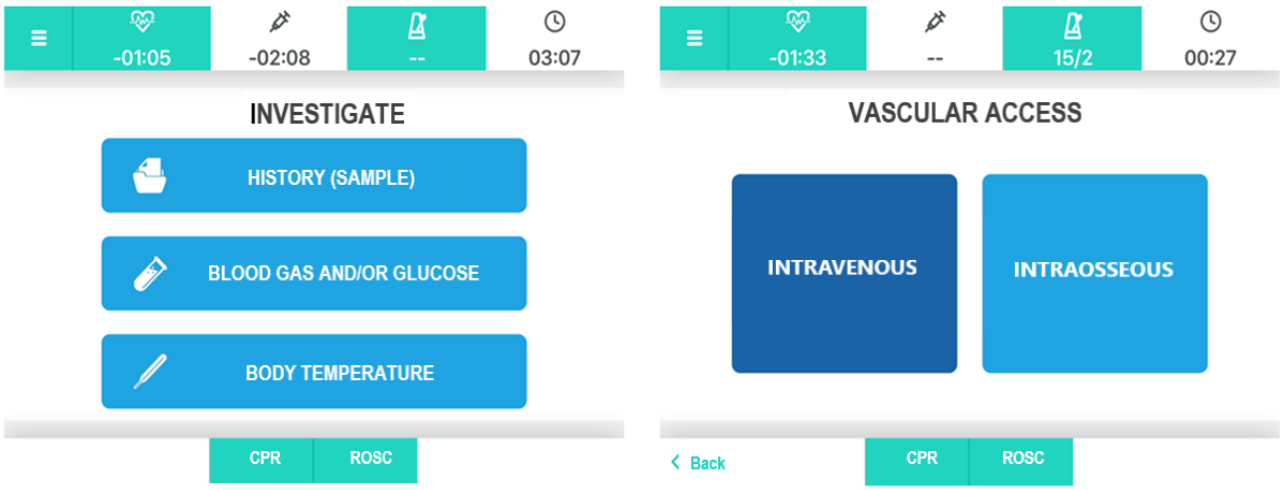
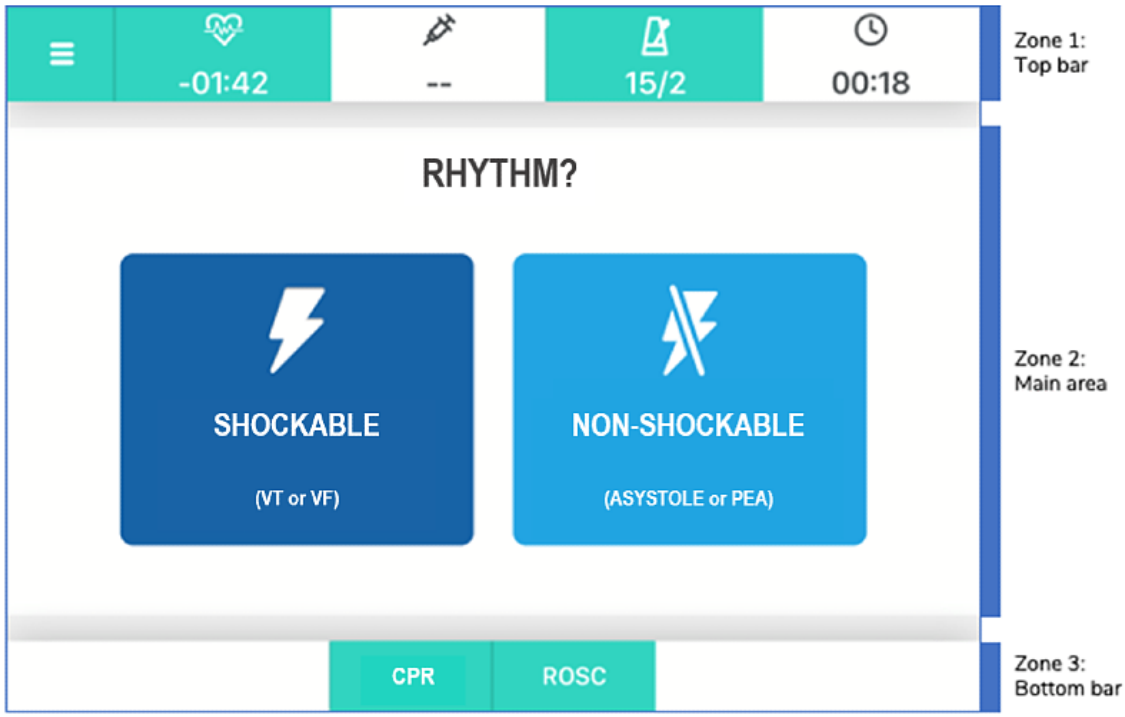


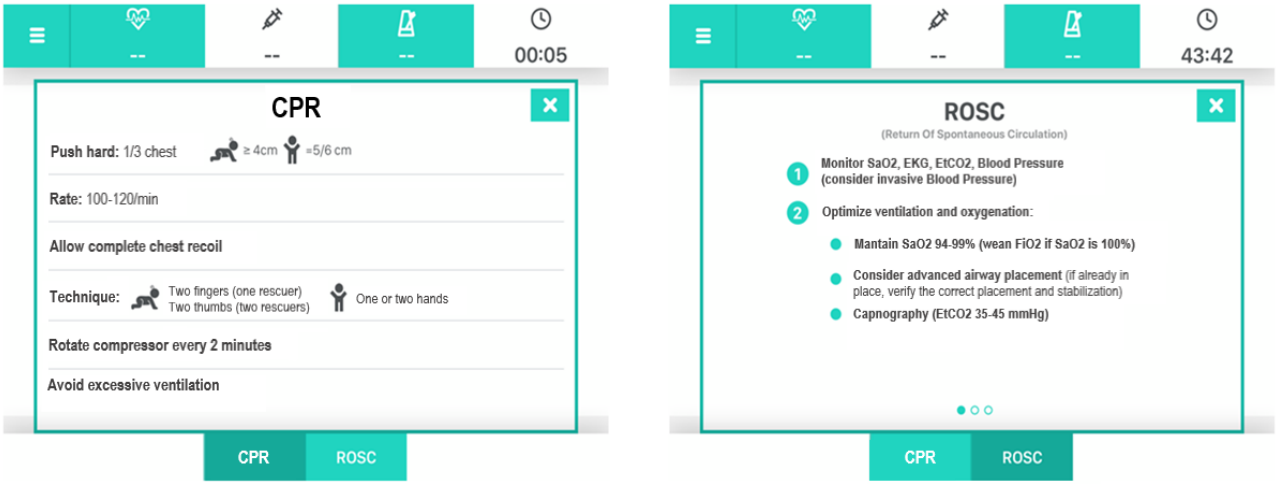
Figure 2. User interface main areas. CPR: cardiopulmonary resuscitation; ROSC: return of spontaneous circulation; VF: ventricular fibrillation; VT: ventricular tachycardia.



The main area presents the actions to be performed (with buttons of the same shape and color) or a question with different choices (buttons with different colors). Once the user taps on a button, the flow of prompts will progress following the user's choices. On the top bar a menu-log button, a 2-minute countdown clock for repeat rhythm check button, a button with countdown for medications, a metronome button, and a total counter are displayed. The metronome button can be activated by a tap: this is a sound guide to perform compressions at the recommended rate (100-120/minute). On the bottom bar, CPR and Return of Spontaneous Circulation (ROSC) buttons are available at any time: the CPR button opens up a recap of the characteristics of a high-quality CPR, whereas the "ROSC" button summarizes the recommended management when ROSC is achieved (Figure 3).

XSL•FO
RenderX

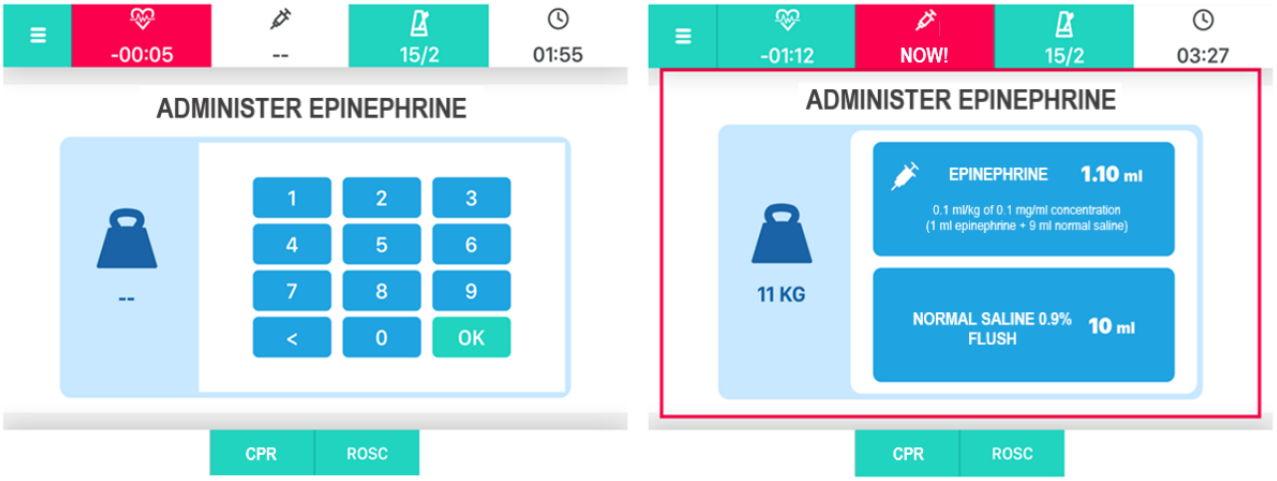
Figure 3. Cardiopulmonary resuscitation (CPR) and return of spontaneous circulation (ROSC) information. EKG: electrocardiogram; EtCO2: end-tidal CO2; FiO2: fraction of inspired oxygen; SaO2: oxygen saturation measured with pulse oximetry.



The flow of actions that pops up in the app main area follows 2 different pathways based on the identified cardiac rhythm on the monitor (shockable versus nonshockable rhythms), as per PALS algorithms.

The app provides assistance with shock delivery, in case of a shockable rhythm, and the preparation/administration of medications, prompting the correct doses (automatically calculated on patients' weight) and time intervals of administration (Figure 4).

Figure 4. Epinephrine administration screens. CPR: cardiopulmonary resuscitation; ROSC: return of spontaneous circulation.



Every 2 minutes the app acoustically and visually reminds the user to check the rhythm and, in case of a shockable rhythm, to deliver a shock.

Finally, the app prompts to search/treat reversible causes of CA and to correctly manage the airway (Figure 5).

The app also gives audio prompts, suggesting the user to perform the actions shown on the main area of the screen. The user can navigate the app only using touch gestures. Voice interaction has not been integrated due to the characteristics of the resuscitation environment, which would impede accurate recognition of vocal commands.

All actions done by the user are sequentially saved on the device in the *log* function to store information that can be retrieved for any documentation purpose.

XSL•FO
RenderX

Figure 5. Reversible causes and airway management screens. CPR: cardiopulmonary resuscitation; ROSC: return of spontaneous circulation.

Pilot Study

We conducted a single-center simulation-based pilot nonrandomized controlled study in October–November 2019 at the University Hospital of Padova, Italy. Although our study is not a randomized controlled trial (RCT), we followed the guidelines for reporting simulation-based studies as far as applicable [33]. All the teams of the intervention group had to manage a standard simulated PCA scenario with the use of the same version of the PediAppRREST app, while the teams of the control group managed the same scenario without the app, following usual practice. We chose a case of nonshockable PCA because asystole/pulseless electrical activity are the most frequent initial CA rhythms in children [4]. Participants were pediatric residents in their third/fourth/fifth year of their pediatric residency program who had AHA-PALS provider certification. To obtain this certification providers have to undertake a standardized resuscitation course with theorical and simulation-based education [34]. Residents unable to attend the simulations because of maternity/sick/personal leave or training abroad were excluded from the study.

Further details regarding the study methodology and procedures are described in Multimedia Appendix 1.

Written informed consent for participation was obtained from all the participants. The study was approved by the Hospital Ethics Committee as an educational project.

Outcomes

The primary outcomes of our study were the usability of the app and the team leader's workload. They were measured by 2 validated questionnaires, the User Experience Questionnaire (UEQ) [35-38] and the National Aeronautics and Space Administration Raw-Task Load Index (NASA RTLX) questionnaire [39,40], respectively.

Secondary outcomes were qualitative feedback on the app provided by participants, preliminary data on deviations in management from PALS guidelines recommendations, time to epinephrine administration, and resuscitation performance of the teams evaluated with the validated Clinical Performance Tool (CPT) [41,42].

Research Measures

The User Experience Questionnaire

The UEQ is a validated questionnaire which comprises 26 items. Each item is represented by 2 terms with opposite meanings that the user evaluates on a 7-point Likert-type scale (from –3 to +3). The 26 items are grouped into 6 scales that cover both classical usability aspects (efficiency, perspicuity, dependability) and user experience aspects (attractiveness, stimulation, novelty). The range of each scale is also between –3 and +3. The standard interpretation of the scale is that values between –0.8 and 0.8 represent a neutral evaluation of the corresponding scale, values over 0.8 represent a positive evaluation, and values less than –0.8 a negative evaluation [35-38].

The NASA Raw-Task Load Index

The NASA RTLX is a simplified version of the NASA-Task Load Index which is a subjective multidimensional tool designed to assess workload. Six subscales represent different domains of the perceived workload: mental demand, physical demand, temporal demand, frustration, effort, and performance. Each domain is clearly defined and rated by participants through a 0 to 100 scale with 5-point steps. The ratings of the 6 subscales are simply averaged to create an estimate of overall workload, defined as low (<40), moderate (between 40 and 60), and high (>60) [39,40].

Qualitative Feedback

The qualitative feedback on user app interaction was collected through open-ended questions in the postscenario questionnaire ("What are the main difficulties you have encountered in the use of the app and/or tablet?", "Do you have any suggestions to improve the app or its use?") and through the postscenario debriefing. Feedback from participants was categorized by common themes.

Deviations From PALS Guidelines

Deviations from PALS guidelines recommendations were defined as delays and errors according to a novel checklist adapted to our intervention and scenario. We derived this new measure from the checklist, denominated c-DEV, published by Wolfe et al [20] and we integrated it with evidence-based

XSL•FO
RenderX

guidelines [6-8], previously reported scoring tools [41-44], and checklists [45,46]; we named our new modified checklist c-DEVplus (Multimedia Appendix 2). It includes 16 items which represent correct critical actions for pediatric resuscitation. Each item is associated with a score as follows: 0 (action correctly and timely performed, as described in the item) or 1 point (action not undertaken or undertaken incorrectly or with wrong timing). The points of each action were summed and expressed as a total score ranging from 0 to 16. A higher c-DEVplus total score corresponds to more deviations from the guidelines. Time to epinephrine administration was measured in seconds from the recognition of CA to the moment epinephrine was administered.

The Clinical Performance Tool

The CPT is a performance assessment tool and a validated scoring system designed based on PALS algorithms comprising different tasks. Each task is scored as follows: not performed (0 points), performed partially, incorrectly, or late (1 point); and performed completely, correctly, and timely (2 points). Thus, the tool assesses sequence, timing, and quality of specific actions during different simulated scenarios [41,42]. In our study we used the section related to asystole and the reviewer assigned a score from 0 to 13 to each scenario.

Statistical Analysis

The characteristics of the study participants, stratified by group allocation (control vs intervention), and the outcome variables, were summarized using descriptive statistics, and compared between the 2 groups using Mann–Whitney *U* tests for continuous variables, and chi-squared tests or Fisher exact tests for categorical variables. Data were entered into an Excel database (Microsoft) and were analyzed using Stata (version 13; StataCorp). *P*-values were two-sided, and differences were considered significant if *P* was <.05. The statistical analysis on the app usability, obtained through the UEQ, was conducted using the UEQ dedicated software, which provides descriptive statistics, and Cronbach α coefficients for each subscale, indexing their internal consistency [38].

Results

Characteristics of Participants

During the study period, 63 pediatric residents were assessed for eligibility, of whom 48 (16 for each one of the 3 years of residency program involved in the study) were included in the study and divided into teams of 3. Five teams managed the case following usual care (control group), whereas 11 teams (intervention group) conducted the scenario using the support of the PediAppRREST app (Multimedia Appendix 3).

Participants' demographic characteristics, as well as training and clinical experience of resuscitation, were comparable between the 2 groups (Multimedia Appendix 4).

Primary Outcomes

The PediAppRREST app attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty were on average evaluated positively; mean UEQ scale scores were substantially above the 0.8 cutoff. UEQ scales internal consistency varied from poor (perspicuity) to acceptable (efficiency and dependability), good (novelty), and excellent (attractiveness and stimulation); see Table 1.

Table 1. PediAppRREST app usability.

UEQ[a] scale	Mean (SD)	95% CI	Cronbach α coefficient
Attractiveness	1.712 (1.434)	0.865-2.559	.98
Perspicuity	1.750 (0.880)	1.230-2.270	.55
Efficiency	1.932 (0.929)	1.383-2.481	.72
Dependability	1.568 (1.102)	0.917-2.219	.76
Stimulation	1.598 (1.333)	0.811-2.386	.91
Novelty	2.205 (0.740)	1.767-2.642	.86

[a]UEQ: User Experience Questionnaire.

Team leaders' perceived workload was comparable between the 2 groups; median NASA RTLX score was 67.5 (IQR 65.0-81.7) for the control group and 66.7 (IQR 54.2-76.7) for the intervention group (*P*=.57).

Secondary Outcomes

Based on the qualitative feedback provided by participants, the most frequently highlighted difficulties were (1) interacting with the screens flow because information delivery was unclear about recommendation on performance of an action versus suggestion to perform an action based on the team skillset, that is, advanced airway management (n=8 team leaders); (2) information overload in the reversible causes screen, which was perceived as too dense (n=5); (3) understanding whether the selection of an icon for a recommended action had to occur at the beginning of the action or after the action was completed (ie, users did not understand whether to select the epinephrine icon at the time of preparation or administration; n=5); and (4) interacting with the app while leading the teamwork (n=4). A less frequently reported difficulty was the lack of a traditional PALS algorithm embedded within the app (n=2). Lastly, it also emerged that longer training and familiarization with the app before the simulated scenario would have been beneficial to interact more efficiently and effectively with the app.

With respect to deviations from the guidelines, the frequency of (1) incorrect compressions-to-ventilations ratio during CPR, (2) prescription of incorrect doses/dilutions of epinephrine, and (3) lack of search/treatment of reversible causes of CA (ie, hypovolemia) were higher in the control group in comparison

XSL•FO
RenderX

to the intervention group; however, these differences did not reach statistical significance (Table 2). Furthermore, although the median number of deviations (c-DEVplus score) from PALS guidelines recommendations was similar between the 2 groups, a statistically significant delay (*P*=.015) in epinephrine administration was observed for the intervention group compared with the control group (Table 2). Nevertheless, the overall clinical performance of the teams, evaluated through the validated CPT, was comparable between the 2 groups (Table 2).

Table 2. Resuscitation performance of the teams.

Performance	Control group (N=5)	Intervention group (N=11)	*P*-value
Incorrect compressions-to-ventilation ratio, n (%)	1 (20)	0 (0)	.31
Incorrect dose or dilution of epinephrine, n (%)	1 (20)	0 (0)	.31
Lack of search and treatment of reversible causes of cardiac arrest, n (%)	2 (40)	2 (18)	.55
c-DEVplus score, median (IQR)	6 (6-7)	6 (4-7)	.27
Time (seconds) to first epinephrine administration, median (IQR)	165 (139-173)	254 (204-290)	.015
CPT[a] scores, median (IQR)	9 (8-10)	9 (9-10)	.77

[a]CPT: Clinical Performance Tool.

Discussion

Principal Results

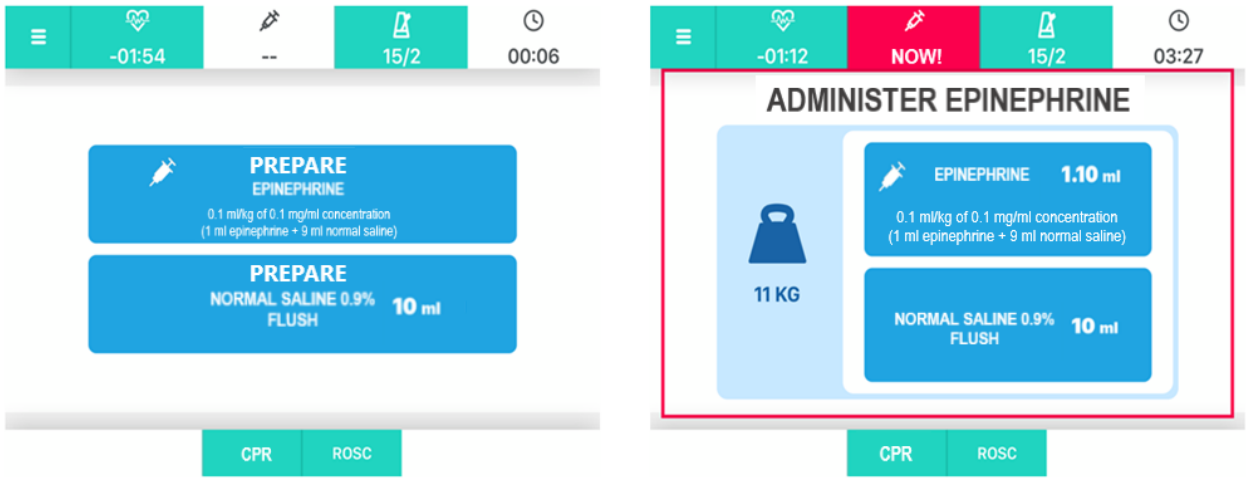
We developed a novel cognitive aid, an app for tablet, which aims to optimize the management of PCA by facilitating increased adherence to guideline recommendations. In our pilot study, the app showed a good usability profile and its use was not associated with increased team leaders' workload. These findings are encouraging and in contrast to data on previously developed cognitive support tools which are shown to increase users' workload [47]. Our results lay the ground to further test in an RCT the effectiveness of the PediAppRREST app which has now been refined. In fact, based on the feedback provided by participants in this pilot study, we have modified the app to improve the app–user interaction, and integrate better the use of the tool within the scenario management flow, by minimizing possible distractions related to its use, as well as possible related interference in team communication. Participants' feedback has also guided us in better organizing the presentation of information/prompts (wording, content per screen, definition of single management steps, type of prompts) and the flow of information in the app. We have also reduced information load per screen and endeavored to facilitate the user's understanding of the prompts presented by the app. Lastly, participants expressed the need for a longer training and testing of the app to better familiarize with the tool. This will be taken into account for the design of the RCT protocol.

Our preliminary results, although based on a very limited sample size, highlighted the potential benefits, as well as the drawbacks, of using the app to guide resuscitation. Nevertheless, we are confident that the refined version of the app based on the feedback received in this pilot study and a better familiarization with it prior to its use have the potential to significantly reduce deviations from guidelines, which correlate with clinical outcomes [20], while limiting drawbacks, such as the delay in epinephrine administration. Our pilot study showed a median time to epinephrine administration approximately 90 seconds longer in the group using the app in comparison to the control group. This would be an unacceptable side effect as a recent study showed how survival decreased by 5% for every minute delay in administration of epinephrine [48]. For this reason, we have re-designed the information flow and presentation of information for the delivery of epinephrine with the aim to better reflect the management flow. Thus, we have separated the information on the preparation of epinephrine from the information on its administration (Figure 6).

We also measured team performance by means of the validated CPT score [41,42] and found similar results for the control and intervention groups. However, although this tool has been widely used in simulation research, its score has not shown to be associated with change in clinical outcomes. In addition, its items are limited and only partially reflect the potential of the app in reducing deviations from guidelines. Nevertheless, our findings are an important starting point for the design and development of an RCT to test the effectiveness of the app.

XSL•FO
RenderX

Figure 6. Epinephrine preparation and administration screens. CPR: cardiopulmonary resuscitation; ROSC: return of spontaneous circulation.



Limitations

The results of our pilot study are preliminary and are the first step of a larger project that aims at testing the effectiveness of the PediAppRREST app. As such, they cannot be considered definitive, as the sample size is very limited. However, this pilot experience was essential to refine the newly developed tool and to verify there were the premises for a larger comparative study.

Participants of our study were exclusively pediatric residents. Although this may limit the generalizability of study findings to other clinicians, trainees are in a unique process of learning and are more used to incorporate digital assistive tools in their clinical practice. While in our setting it is easier to get trainees involved in simulation projects, than experienced clinicians, the app could be seen as a training tool in itself and future studies will be designed to test its effectiveness in knowledge and skill retention.

Our resuscitation team composition differs from other settings, where a co-team leader, a respiratory therapist, or a CPR coach is often part of the team, which includes a higher number of members. We chose a team composition that reflects the actual management at our institution during night shifts, where trainees are in the front line in the management of the first few minutes of pediatric emergencies. We felt it was important to test the device in this highly stressful and staff-limited scenario where the app could be most useful.

The intervention and control groups were unbalanced with respect to the intervention group as we needed to test the device in the first place, and a few control teams were necessary to test the overall procedures for the RCT. Similarly, although we randomly allocated residents to each team and randomly selected the teams who were performing the scenario without the app, the timing of app availability was the main determinant of our pilot study procedure and a proper randomization process will be performed for the future RCT.

A single trained reviewer rated the videos as the preliminary evaluation of team performance and deviations from guidelines were a secondary aim of this study. Two independent and trained reviewers will be ensured for the RCT, and interrater reliability will be reported and monitored. Blinding of participants and research staff was not possible because of the nature of simulation-based study. Blinding of video reviewer was not applied as video recording of the team leader using the app and the tablet was necessary to detect possible difficulties with its use. Blinding of the statistician performing data analysis will be ensured for the RCT.

The high-fidelity simulation setting during the last decades has established itself as a way to investigate rare but high-risk medical conditions. Although it does not provide data on actual patient outcomes, it is the best available way to reproduce and study rare high-stake emergencies and test novel devices developed to improve their management without compromising patients' safety.

Comparison With Prior Work

Several researchers have tried to create and test software products and apps to improve the quality of resuscitation. Different products have been conceived, mostly dealing with OHCA and in-hospital CA (IHCA) in adults. For instance, to help lay rescuers to manage adult cases of OHCA, different tools have been developed, such as the M-AID (an app for mobile phones [49]); a handheld personal computer software (personal digital assistant) [50]; and a voice activated decision support system, which is installed on a smartphone [51]. Conversely, other products have been developed and tested to aid professionals in training or qualified health professionals in the management of adult IHCA. These instruments comprise

XSL•FO
RenderX

mobile apps, such as the iResus app developed by the Resuscitation Council UK [52] and the Medical Assistance eXpert (MAX) smartphone app [53,54], as well as an iPod Touch software (decision support) [55]. Very variable results were obtained when these tools were tested in simulation-based RCTs with only partial benefit observed in those studies that achieved positive results [50-52,54,55].

With regard to PCA, a mobile app was developed to help adolescent lay bystanders to manage an infant OHCA scenario, but an RCT showed that the participants who used the app only partially improved their performance [56]. Siebert et al [28] adapted PALS guidelines to augmented reality glasses and tested the novel cognitive aid through a simulation-based RCT on a case of shockable pediatric IHCA. The trial did not show a significant difference in time to defibrillation when using the augmented reality glasses compared with the PALS pocket reference card. However, the intervention group showed a reduction in the number of errors in defibrillation doses.

An app to help nurses prepare and administer drugs for infusion during in-hospital pediatric resuscitation has been recently developed and tested in a simulation-based RCT. The app was effective in reducing errors and time to preparation/delivery of medications compared with conventional methods [57]. Another simulation-based RCT, from the same research group, has shown a reduction in the time to critical actions and in the deviations from guidelines recommendations in the management of a shockable PCA in the group guided by an app for tablet compared with the group that used the PALS pocket reference card [30]. However, the app had not been previously tested for its usability and the sample size was limited (13 residents per group playing the team leader role, whereas the other team members were part of the research staff) [30].

To our knowledge, no app similar to the PediAppRREST has been tested in a pilot study to be refined, and to evaluate its usability and related workload before being tested in an adequately powered RCT.

Conclusion

We developed and refined a novel interactive tablet app (PediAppRREST) for the management of PCA that has potential to reduce deviations from guidelines recommendations. The app showed a good usability profile and was not associated with higher team leaders' workload. After this pilot testing its effectiveness will be evaluated in an adequately powered simulation-based RCT.

Acknowledgments

The authors thank Professor Anna Chiara Frigo for her contribution in randomly assigning residents in teams of 3, the Department of Informatics and Telecommunications of University of Padova and its Director Dr Andrea Baraldo for supporting the project, and all pediatric residents who took part into this study. This work was funded and supported by the Department of Women's and Children's Health and the University of Padova, Italy.

Authors' Contributions

FC and SB conceived the development of the app and the pilot testing study. ML, LS, and FT were responsible for designing and developing the app prototypes. FC, SB, VS, FM, MD, DS, and LDD tested the app prototypes, organized, and conducted the simulation sessions. FC reviewed the videotapes of the simulations and collected data. MA performed the statistical analysis. FC, SB, and MA drafted the manuscript. SB supervised the different stages of the study. All authors revised and approved the final version of the manuscript.

Conflicts of Interest

The authors affiliated with RE:Lab had no direct involvement in the project, other than the technical activities related to the development of the app. All the other authors have no conflict of interest to declare neither with the RE:Lab company nor with other companies.

Multimedia Appendix 1
Pilot study methodology.
[DOCX File , 23 KB - mhealth_v8i10e19070_app1.docx]

Multimedia Appendix 2
c-DEVplus score calculation grid.
[DOCX File , 19 KB - mhealth_v8i10e19070_app2.docx]

Multimedia Appendix 3
Flowchart of participant recruitment and study group allocation.
[DOCX File , 34 KB - mhealth_v8i10e19070_app3.docx]

Multimedia Appendix 4

XSL•FO
RenderX

Characteristics of participants: demographics, training and clinical experience on resuscitation.
[DOCX File , 22 KB - mhealth_v8i10e19070_app4.docx]

References

1. Tibballs J, Kinney S. A prospective study of outcome of in-patient paediatric cardiopulmonary arrest. Resuscitation 2006 Dec;71(3):310-318. [doi: 10.1016/j.resuscitation.2006.05.009] [Medline: 17069956]
2. Reis AG, Nadkarni V, Perondi MB, Grisi S, Berg RA. A prospective investigation into the epidemiology of in-hospital pediatric cardiopulmonary resuscitation using the international Utstein reporting style. Pediatrics 2002 Feb;109(2):200-209. [doi: 10.1542/peds.109.2.200] [Medline: 11826196]
3. Suominen P, Olkkola KT, Voipio V, Korpela R, Palo R, Räsänen J. Utstein style reporting of in-hospital paediatric cardiopulmonary resuscitation. Resuscitation 2000 Jun;45(1):17-25 [FREE Full text] [doi: 10.1016/s0300-9572(00)00167-2] [Medline: 10838235]
4. Nehme Z, Namachivayam S, Forrest A, Butt W, Bernard S, Smith K. Trends in the incidence and outcome of paediatric out-of-hospital cardiac arrest: A 17-year observational study. Resuscitation 2018 Jul;128:43-50. [doi: 10.1016/j.resuscitation.2018.04.030] [Medline: 29704520]
5. Knudson JD, Neish SR, Cabrera AG, Lowry AW, Shamszad P, Morales DLS, et al. Prevalence and outcomes of pediatric in-hospital cardiopulmonary resuscitation in the United States: an analysis of the Kids' Inpatient Database*. Crit Care Med 2012 Nov;40(11):2940-2944. [doi: 10.1097/CCM.0b013e31825feb3f] [Medline: 22932398]
6. de Caen AR, Berg MD, Chameides L, Gooden CK, Hickey RW, Scott HF, et al. Part 12: Pediatric Advanced Life Support: 2015 American Heart Association Guidelines Update for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2015 Nov 3;132(18 Suppl 2):S526-S542. [doi: 10.1161/CIR.0000000000000266] [Medline: 26473000]
7. Atkins DL, Berger S, Duff JP, Gonzales JC, Hunt EA, Joyner BL, et al. Part 11: Pediatric Basic Life Support and Cardiopulmonary Resuscitation Quality: 2015 American Heart Association Guidelines Update for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2015 Nov 03;132(18 Suppl 2):S519-S525. [doi: 10.1161/CIR.0000000000000265] [Medline: 26472999]
8. Duff JP, Topjian AA, Berg MD, Chan M, Haskell SE, Joyner BL, et al. 2019 American Heart Association Focused Update on Pediatric Advanced Life Support: An Update to the American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Pediatrics 2020 Jan;145(1). [doi: 10.1542/peds.2019-1361] [Medline: 31727859]
9. Maconochie IK, Bingham R, Eich C, López-Herce J, Rodríguez-Núñez A, Rajka T, Paediatric life support section Collaborators. European Resuscitation Council Guidelines for Resuscitation 2015: Section 6. Paediatric life support. Resuscitation 2015 Oct;95:223-248. [doi: 10.1016/j.resuscitation.2015.07.028] [Medline: 26477414]
10. Maconochie IK, de Caen AR, Aickin R, Atkins DL, Biarent D, Guerguerian A, Pediatric Basic Life SupportPediatric Advanced Life Support Chapter Collaborators. Part 6: Pediatric basic life support and pediatric advanced life support: 2015 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science with Treatment Recommendations. Resuscitation 2015 Oct;95:e147-e168. [doi: 10.1016/j.resuscitation.2015.07.044] [Medline: 26477423]
11. Abella BS, Sandbo N, Vassilatos P, Alvarado JP, O'Hearn N, Wigder HN, et al. Chest compression rates during cardiopulmonary resuscitation are suboptimal: a prospective study during in-hospital cardiac arrest. Circulation 2005 Feb 01;111(4):428-434. [doi: 10.1161/01.CIR.0000153811.84257.59] [Medline: 15687130]
12. Abella BS, Alvarado JP, Myklebust H, Edelson DP, Barry A, O'Hearn N, et al. Quality of cardiopulmonary resuscitation during in-hospital cardiac arrest. JAMA 2005 Jan 19;293(3):305-310. [doi: 10.1001/jama.293.3.305] [Medline: 15657323]
13. Wik L, Kramer-Johansen J, Myklebust H, Sørebø H, Svensson L, Fellows B, et al. Quality of cardiopulmonary resuscitation during out-of-hospital cardiac arrest. JAMA 2005 Jan 19;293(3):299-304. [doi: 10.1001/jama.293.3.299] [Medline: 15657322]
14. Sutton RM, Niles D, French B, Maltese MR, Leffelman J, Eilevstjønn J, et al. First quantitative analysis of cardiopulmonary resuscitation quality during in-hospital cardiac arrests of young children. Resuscitation 2014 Jan;85(1):70-74 [FREE Full text] [doi: 10.1016/j.resuscitation.2013.08.014] [Medline: 23994802]
15. Sutton RM, Case E, Brown SP, Atkins DL, Nadkarni VM, Kaltman J, et al. A quantitative analysis of out-of-hospital pediatric and adolescent resuscitation quality--A report from the ROC epistry-cardiac arrest. Resuscitation 2015 Aug;93:150-157 [FREE Full text] [doi: 10.1016/j.resuscitation.2015.04.010] [Medline: 25917262]
16. Niles DE, Duval-Arnould J, Skellett S, Knight L, Su F, Raymond TT, pediatric Resuscitation Quality (pediRES-Q) Collaborative Investigators. Characterization of Pediatric In-Hospital Cardiopulmonary Resuscitation Quality Metrics Across an International Resuscitation Collaborative. Pediatr Crit Care Med 2018 May;19(5):421-432. [doi: 10.1097/PCC.0000000000001520] [Medline: 29533355]
17. Cheng A, Hunt EA, Grant D, Lin Y, Grant V, Duff JP, International Network for Simulation-based Pediatric Innovation, Research,Education CPR Investigators. Variability in quality of chest compressions provided during simulated cardiac arrest across nine pediatric institutions. Resuscitation 2015 Dec;97:13-19. [doi: 10.1016/j.resuscitation.2015.08.024] [Medline: 26417701]

XSL•FO
RenderX

18. Hunt EA, Vera K, Diener-West M, Haggerty JA, Nelson KL, Shaffner DH, et al. Delays and errors in cardiopulmonary resuscitation and defibrillation by pediatric residents during simulated cardiopulmonary arrests. Resuscitation 2009 Jul;80(7):819-825. [doi: 10.1016/j.resuscitation.2009.03.020] [Medline: 19423210]
19. Labrosse M, Levy A, Donoghue A, Gravel J. Delays and errors among pediatric residents during simulated resuscitation scenarios using Pediatric Advanced Life Support (PALS) algorithms. Am J Emerg Med 2015 Oct;33(10):1516-1518. [doi: 10.1016/j.ajem.2015.07.049] [Medline: 26298054]
20. Wolfe HA, Morgan RW, Zhang B, Topjian AA, Fink EL, Berg RA, American Heart Association's Get With the Guidelines-Resuscitation Investigator. Deviations from AHA guidelines during pediatric cardiopulmonary resuscitation are associated with decreased event survival. Resuscitation 2020 Apr;149:89-99. [doi: 10.1016/j.resuscitation.2020.01.035] [Medline: 32057946]
21. McEvoy MD, Field LC, Moore HE, Smalley JC, Nietert PJ, Scarbrough SH. The effect of adherence to ACLS protocols on survival of event in the setting of in-hospital cardiac arrest. Resuscitation 2014 Jan;85(1):82-87 [FREE Full text] [doi: 10.1016/j.resuscitation.2013.09.019] [Medline: 24103233]
22. Lin Y, Cheng A, Grant VJ, Currie GR, Hecker KG. Improving CPR quality with distributed practice and real-time feedback in pediatric healthcare providers - A randomized controlled trial. Resuscitation 2018 Sep;130:6-12. [doi: 10.1016/j.resuscitation.2018.06.025] [Medline: 29944894]
23. Lakomek F, Lukas R, Brinkrolf P, Mennewisch A, Steinsiek N, Gutendorf P, et al. Real-time feedback improves chest compression quality in out-of-hospital cardiac arrest: A prospective cohort study. PLoS One 2020;15(2):e0229431 [FREE Full text] [doi: 10.1371/journal.pone.0229431] [Medline: 32092113]
24. Wagner M, Bibl K, Hrdliczka E, Steinbauer P, Stiller M, Gröpel P, et al. Effects of Feedback on Chest Compression Quality: A Randomized Simulation Study. Pediatrics 2019 Feb;143(2):2018-2441. [doi: 10.1542/peds.2018-2441] [Medline: 30700565]
25. Kirkbright S, Finn J, Tohira H, Bremner A, Jacobs I, Celenza A. Audiovisual feedback device use by health care professionals during CPR: a systematic review and meta-analysis of randomised and non-randomised trials. Resuscitation 2014 Apr;85(4):460-471. [doi: 10.1016/j.resuscitation.2013.12.012] [Medline: 24361457]
26. Lin Y, Chiang W, Hsieh M, Sun J, Chang Y, Ma MH. Quality of audio-assisted versus video-assisted dispatcher-instructed bystander cardiopulmonary resuscitation: A systematic review and meta-analysis. Resuscitation 2018 Feb;123:77-85. [doi: 10.1016/j.resuscitation.2017.12.010] [Medline: 29242057]
27. Drummond D, Arnaud C, Guedj R, Duguet A, de Suremain N, Petit A. Google Glass for Residents Dealing With Pediatric Cardiopulmonary Arrest: A Randomized, Controlled, Simulation-Based Study. Pediatr Crit Care Med 2017 Feb;18(2):120-127. [doi: 10.1097/PCC.0000000000000977] [Medline: 28165347]
28. Siebert JN, Ehrler F, Gervaix A, Haddad K, Lacroix L, Schrurs P, et al. Adherence to AHA Guidelines When Adapted for Augmented Reality Glasses for Assisted Pediatric Cardiopulmonary Resuscitation: A Randomized Controlled Trial. J Med Internet Res 2017 May 29;19(5):e183 [FREE Full text] [doi: 10.2196/jmir.7379] [Medline: 28554878]
29. Metelmann B, Metelmann C, Schuffert L, Hahnenkamp K, Brinkrolf P. Medical Correctness and User Friendliness of Available Apps for Cardiopulmonary Resuscitation: Systematic Search Combined With Guideline Adherence and Usability Evaluation. JMIR Mhealth Uhealth 2018 Nov 06;6(11):e190 [FREE Full text] [doi: 10.2196/mhealth.9651] [Medline: 30401673]
30. Siebert JN, Lacroix L, Cantais A, Manzano S, Ehrler F. The Impact of a Tablet App on Adherence to American Heart Association Guidelines During Simulated Pediatric Cardiopulmonary Resuscitation: Randomized Controlled Trial. J Med Internet Res 2020 May 27;22(5):e17792 [FREE Full text] [doi: 10.2196/17792] [Medline: 32292179]
31. Stritoni V, Martinolli F, Daverio M, Diogo M, Binotti M, Genoni G, et al. Gaps in the management of nonshockable paediatric cardiac arrest by paediatric residents in simulated scenario. 2019 Presented at: Proceedings of the Third European Pediatric Resuscitation and Emergency Medicine Conference (PREM); 2019 May 23-24; Ghent, Belgium p. 23-24.
32. Ehrler F, Del Zotto M, Rouyer F, Weinhold T, Lovis C, Siebert J. Design of InterFACE: A Tool to Improve Collaborative Work and Decision Making During Rescucitation. Stud Health Technol Inform 2018;255:117-121. [Medline: 30306919]
33. Cheng A, Kessler D, Mackinnon R, Chang TP, Nadkarni VM, Hunt EA, International Network for Simulation-based Pediatric Innovation, Research, and Education (INSPIRE) Reporting Guidelines Investigators. Reporting Guidelines for Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements. Simul Healthc 2016 Aug;11(4):238-248. [doi: 10.1097/SIH.0000000000000150] [Medline: 27465839]
34. Bhanji F, Donoghue AJ, Wolff MS, Flores GE, Halamek LP, Berman JM, et al. Part 14: Education: 2015 American Heart Association Guidelines Update for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2015 Nov 03;132(18 Suppl 2):S561-S573. [doi: 10.1161/CIR.0000000000000268] [Medline: 26473002]
35. Laugwitz B, Schrepp M, Held T. Construction and evaluation of a user experience questionnaire. In: Holzinger A, editor. HCI and Usability for Education and Work. USAB 2008. Lecture Notes in Computer Science, vol 5298. Berlin, Heidelberg: Springer; 2008:63-76.
36. Schrepp M, Hinderks A, Thomaschewski J. Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. In: Marcus A, editor. Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Lecture Notes in Computer Science, vol 8517. Springer: Design, User Experience, and Usability. Theories,

XSL•FO
RenderX

Methods, and Tools for Designing the User Experience. Lecture Notes in Computer Science. Springer International Publishing; 2014:383-392.
37. Schrepp M, Hinderks A, Thomaschewski J. Construction of a Benchmark for the User Experience Questionnaire (UEQ). IJIMAI 2017;4(4):40-44. [doi: 10.9781/ijimai.2017.445]
38. User Experience Questionnaire. Hockenheim, Germany: Team UEQ URL: https://www.ueq-online.org/ [accessed 2020-09-08]
39. Hart SG. Nasa-Task Load Index (NASA-TLX); 20 Years Later. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2006 Oct 16 Presented at: 50th Annual Meeting of the Human Factors and Ergonomics Society; October 16-20, 2006; San Francisco, CA p. 904-908 URL: http://pro.sagepub.com/content/50/9/904.full.pdf+html [doi: 10.1177/154193120605000909]
40. Cao A, Chintamani KK, Pandya AK, Ellis RD. NASA TLX: software for assessing subjective mental workload. Behav Res Methods 2009 Feb;41(1):113-117. [doi: 10.3758/BRM.41.1.113] [Medline: 19182130]
41. Levy A, Donoghue A, Bailey B, Thompson N, Jamoulle O, Gagnon R, et al. External validation of scoring instruments for evaluating pediatric resuscitation. Simul Healthc 2014 Dec;9(6):360-369. [doi: 10.1097/SIH.0000000000000052] [Medline: 25503530]
42. Donoghue A, Ventre K, Boulet J, Brett-Fleegler M, Nishisaki A, Overly F, EXPRESS Pediatric Simulation Research Investigators. Design, implementation, and psychometric analysis of a scoring instrument for simulated pediatric resuscitation: a report from the EXPRESS pediatric investigators. Simul Healthc 2011 Apr;6(2):71-77. [doi: 10.1097/SIH.0b013e31820c44da] [Medline: 21358564]
43. Brett-Fleegler MB, Vinci RJ, Weiner DL, Harris SK, Shih M, Kleinman ME. A simulator-based tool that assesses pediatric resident resuscitation competency. Pediatrics 2008 Mar;121(3):e597-e603. [doi: 10.1542/peds.2005-1259] [Medline: 18283069]
44. Reid J, Stone K, Brown J, Caglar D, Kobayashi A, Lewis-Newby M, et al. The Simulation Team Assessment Tool (STAT): development, reliability and validation. Resuscitation 2012 Jul;83(7):879-886. [doi: 10.1016/j.resuscitation.2011.12.012] [Medline: 22198422]
45. Arriaga AF, Bader AM, Wong JM, Lipsitz SR, Berry WR, Ziewacz JE, et al. Simulation-based trial of surgical-crisis checklists. N Engl J Med 2013 Jan 17;368(3):246-253. [doi: 10.1056/NEJMsa1204720] [Medline: 23323901]
46. McEvoy MD, Smalley JC, Nietert PJ, Field LC, Furse CM, Blenko JW, et al. Validation of a detailed scoring checklist for use during advanced cardiac life support certification. Simul Healthc 2012 Aug;7(4):222-235 [FREE Full text] [doi: 10.1097/SIH.0b013e3182590b07] [Medline: 22863996]
47. Brown LL, Lin Y, Tofil NM, Overly F, Duff JP, Bhanji F, International Network for Simulation-based Pediatric Innovation, Research, Education CPR Investigators (INSPIRE). Impact of a CPR feedback device on healthcare provider workload during simulated cardiac arrest. Resuscitation 2018 Sep;130:111-117. [doi: 10.1016/j.resuscitation.2018.06.035] [Medline: 30049656]
48. Andersen LW, Berg KM, Saindon BZ, Massaro JM, Raymond TT, Berg RA, et al. Time to Epinephrine and Survival After Pediatric In-Hospital Cardiac Arrest. JAMA 2015 Aug 25;314(8):802-810. [doi: 10.1001/jama.2015.9678] [Medline: 26305650]
49. Zanner R, Wilhelm D, Feussner H, Schneider G. Evaluation of M-AID, a first aid application for mobile phones. Resuscitation 2007 Sep;74(3):487-494. [doi: 10.1016/j.resuscitation.2007.02.004] [Medline: 17452068]
50. Ertl L, Christ F. Significant improvement of the quality of bystander first aid using an expert system with a mobile multimedia device. Resuscitation 2007 Aug;74(2):286-295. [doi: 10.1016/j.resuscitation.2007.01.006] [Medline: 17376581]
51. Hunt EA, Heine M, Shilkofski NS, Bradshaw JH, Nelson-McMillan K, Duval-Arnould J, et al. Exploration of the impact of a voice activated decision support system (VADSS) with video on resuscitation performance by lay rescuers during simulated cardiopulmonary arrest. Emerg Med J 2015 Mar;32(3):189-194. [doi: 10.1136/emermed-2013-202867] [Medline: 24243484]
52. Low D, Clark N, Soar J, Padkin A, Stoneham A, Perkins GD, et al. A randomised control trial to determine if use of the iResus© application on a smart phone improves the performance of an advanced life support provider in a simulated medical emergency. Anaesthesia 2011 Apr;66(4):255-262 [FREE Full text] [doi: 10.1111/j.1365-2044.2011.06649.x] [Medline: 21401537]
53. Lelaidier R, Balança B, Boet S, Faure A, Lilot M, Lecomte F, et al. Use of a hand-held digital cognitive aid in simulated crises: the MAX randomized controlled trial. Br J Anaesth 2017 Nov 01;119(5):1015-1021 [FREE Full text] [doi: 10.1093/bja/aex256] [Medline: 29028930]
54. Donzé P, Balanca B, Lilot M, Faure A, Lecomte F, Boet S, et al. 'Read-and-do' response to a digital cognitive aid in simulated cardiac arrest: the Medical Assistance eXpert 2 randomised controlled trial. Br J Anaesth 2019 Aug;123(2):e160-e163 [FREE Full text] [doi: 10.1016/j.bja.2019.04.049] [Medline: 31153630]
55. Field LC, McEvoy MD, Smalley JC, Clark CA, McEvoy MB, Rieke H, et al. Use of an electronic decision support tool improves management of simulated in-hospital cardiac arrest. Resuscitation 2014 Jan;85(1):138-142 [FREE Full text] [doi: 10.1016/j.resuscitation.2013.09.013] [Medline: 24056391]
56. Hawkes GA, Murphy G, Dempsey EM, Ryan AC. Randomised controlled trial of a mobile phone infant resuscitation guide. J Paediatr Child Health 2015 Nov;51(11):1084-1088. [doi: 10.1111/jpc.12968] [Medline: 26289403]

XSL•FO
RenderX

57. Siebert JN, Ehrler F, Combescure C, Lovis C, Haddad K, Hugon F, PedAMINES Trial Group. A mobile device application to reduce medication errors and time to drug delivery during simulated paediatric cardiopulmonary resuscitation: a multicentre, randomised, controlled, crossover trial. Lancet Child Adolesc Health 2019 May;3(5):303-311. [doi: 10.1016/S2352-4642(19)30003-3] [Medline: 30797722]

Abbreviations

AHA: American Heart Association
CA: cardiac arrest
CPT: Clinical Performance Tool
CPR: cardiopulmonary resuscitation
IHCA: in-hospital cardiac arrest
NASA: National Aeronautics and Space Administration
OHCA: out-of-hospital cardiac arrest
PALS: Pediatric Advanced Life Support
PCA: pediatric cardiac arrest
RCT: randomized controlled trial
ROSC: Return of Spontaneous Circulation
RTLX: Raw Task Load Index
UEQ: Usability Experience Questionnaire

Edited by G Eysenbach; submitted 02.04.20; peer-reviewed by J Siebert, S Erdley; comments to author 12.06.20; revised version received 26.06.20; accepted 26.07.20; published 01.10.20.

Please cite as:
Corazza F, Snijders D, Arpone M, Stritoni V, Martinolli F, Daverio M, Losi MG, Soldi L, Tesauri F, Da Dalt L, Bressan S
Development and Usability of a Novel Interactive Tablet App (PediAppRREST) to Support the Management of Pediatric Cardiac Arrest: Pilot High-Fidelity Simulation-Based Study
JMIR Mhealth Uhealth 2020;8(10):e19070
URL: https://mhealth.jmir.org/2020/10/e19070
doi:10.2196/19070
PMID:32788142

©Francesco Corazza, Deborah Snijders, Marta Arpone, Valentina Stritoni, Francesco Martinolli, Marco Daverio, Maria Giulia Losi, Luca Soldi, Francesco Tesauri, Liviana Da Dalt, Silvia Bressan. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 01.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development and Acceptability of a Method to Investigate Prescription Drug Misuse in Daily Life: Ecological Momentary Assessment Study

Lauren M Papp[1], PhD; Alexandra Barringer[1], MSc; Shari M Blumenstock[1], PhD; Pamela Gu[1], BA; Madison Blaydes[1], MSc; Jaime Lam[1], BA; Chrystyna D Kouros[2], PhD

[1]Department of Human Development and Family Studies, University of Wisconsin-Madison, Madison, WI, United States

[2]Department of Psychology, Southern Methodist University, Dallas, TX, United States

Corresponding Author:
Lauren M Papp, PhD
Department of Human Development and Family Studies
University of Wisconsin-Madison
1300 Linden Drive
Madison, WI, 53706
United States
Phone: 1 608 262 8611
Email: papp@wisc.edu

Abstract

Background: Prescription drug misuse and abuse is an established public health challenge, and young adults are particularly affected. There is a striking lack of real-time, naturalistic data collection assessing intentions to misuse and other precipitating factors at the time of actual misuse, leaving the conditions under which individuals are most likely to misuse prescription medications unknown. Ecological momentary assessment (EMA) apps and protocols designed to capture this information would accelerate and expand the knowledge base and could directly contribute to prevention and treatment efforts.

Objective: The objectives of this study are to describe the development and administration of a mobile app and the EMA protocol designed to collect real-time factors associated with college students' prescription drug misuse intentions and behaviors in daily life; present completion rates, compliance, acceptability, and reactivity associated with the EMA protocol for participants who endorsed recent prescription drug misuse at screening (ie, risk group; n=300) and those who did not (ie, nonrisk group; n=55); and establish initial construct validity by linking the reports of misuse behaviors in daily life collected via the EMA app to prescription drug misuse reported on a standard survey.

Methods: An EMA data collection app and protocol were designed specifically to capture hypothesized contextual factors along with prescription drug misuse intentions and behaviors in daily life. Using this protocol, young adult college students (N=352) completed signal- and event-contingent reports over a 28-day period. When the intention to misuse a prescription drug was endorsed, a brief follow-up prompt was sent 15 min later to collect participants' indications of whether or not misuse had occurred.

Results: Risk-group participants were significantly more likely than nonrisk counterparts to endorse any prescription drug misuse intentions in daily life (P<.001), to complete one or more follow-up reports (P<.001), and to endorse any prescription drug misuse behavior in daily life on the follow-ups (P<.001). Overall, participants demonstrated consistent engagement with the EMA procedures and returned an average of 74.5 (SD 23.82; range 10-122) reports. Participants in the risk and nonrisk groups did not differ in the number of reports they completed (P=.12), the number of their reporting days (P=.32), or their average completion rates (P=.14). The results indicated some evidence of reactivity to the momentary reporting procedure. Participants reported uniformly positive experiences and remained highly engaged throughout the reporting protocol and broader study.

Conclusions: The novel EMA app and protocol provide an effective way to assess real-time factors associated with prescription drug misuse intentions and behaviors in daily life. The resulting investigations offer the potential to provide highly translatable information for research and prevention efforts.

(JMIR Mhealth Uhealth 2020;8(10):e21676) doi:10.2196/21676

XSL•FO
RenderX

KEYWORDS

compliance; ecological momentary assessment; prescription drug misuse; young adult

Introduction

Background

Prescription drug misuse is an established public health concern in the United States and beyond [1,2]. Although available evidence, based on epidemiological studies of group averages, documents harmful correlates [3], there is a striking absence of identified antecedents related to the occurrence of prescription drug misuse in daily life by individuals with elevated risk of engaging in the behavior. This gap is surprising in light of the successful application of rigorous momentary assessment protocols to other legal and illegal substances, including alcohol, tobacco, heroin, cocaine, and cannabis [4-6]. Nonetheless, on the basis of several fundamental differences between prescription drugs and other abusable substances [7], explicit investigation of prescription medication misuse is highly justified. First, prescription medications are precisely designed for other purposes and have obvious health benefits, leading to a potential false sense of safety among people considering their misuse [8]. Moreover, prescription drugs can be taken throughout the day in nearly any setting, often without raising concerns, and their outward effects on the individual may be less noticeable. In addition, medications are frequently misused simultaneously with other substances [9], and misuse of prescriptions is identified as a major risk factor for other substance abuse [10]. Therefore, a necessary next step in this area of research is to develop a sound method to investigate the factors that contribute to prescription drug misuse in populations with elevated risks of engaging in such behavior in natural environments in daily life.

Prescription Drug Misuse Among Young Adults: Background Research and Preliminary Study

Prescription drug misuse itself is potentially harmful and poses additional costs to individuals and society for its established links with illicit drug use, alcohol abuse, mental health problems, risky sexual behaviors, and overdose-related deaths [11-14]. College students comprise a particularly high-risk group for this behavior, with prevalence rates ranging from 5% to 43% [15,16]. College students are more likely than young adult peers to misuse stimulant medications [17] and to routinely be in situations, such as final exam periods, which involve known motivators for misuse (eg, improving study habits and grades) [18-20]. Notably, the college years during young adulthood reflect a vulnerable point in the life course for experiencing substance problems [21], and health-related functioning in this period has a unique and lasting impact on the quality of adult development [22]. Identifying the situational characteristics that predict the occurrence of prescription drug misuse would offer translatable scientific knowledge for improving young adult health.

The status quo in the literature on college students' serious medication practices has been largely cross-sectional and based on retrospective surveys of prescription drug misuse over an extended time frame. For example, the available research predominantly involves participants indicating whether they have misused any prescription drugs over the past 12 months or their lifetime [23-26]. This global approach has offered extensive evidence on who is most likely to misuse prescription drugs and the concurrent risks associated with the misuse. However, there is a striking lack of real-time misuse data collected in natural environments on precipitating factors at the time of the misuse. This type of information allows researchers to assess the conditions under which individuals are most likely to misuse prescription medication. Findings would be critical to accelerating and expanding the knowledge base and could directly contribute to prevention and treatment efforts.

Preliminary findings have demonstrated the utility of collecting young adults' momentary prescription drug misuse reports in real-world settings. We obtained data from 49 mixed-sex dating couples reporting 3 times per day for 10 days [27]. The approach captured misuse instances that occurred since the last report was submitted, along with participant ratings of their emotions, sexual experiences, and alcohol and other drug use. Multilevel modeling of dyadic data found that females' prescription misuse was more likely to occur concurrently with their higher-than-average negative affect and sexual regret, whereas males' misuse was not reliably associated with these momentary experiences. Males and females with relatively greater prescription misuse across the reporting period were more likely to engage in heavy drinking in daily life, and females with greater misuse further showed lower levels of sexual enjoyment and higher risk of unprotected sex. Thus, documenting within- and between-person correlates of young adults' prescription drug misuse in daily life supported the present efforts to develop a more temporally precise ecological momentary assessment (EMA) design to identify real-time triggers of misuse at the time of the substance behavior and to examine initial evidence for construct validity in studies using this method.

Objectives

Building on our preliminary work, the objectives in this study are to (1) describe the development and administration of a mobile app that was designed to collect real-time factors associated with college students' prescription drug misuse in daily life; (2) present completion rates, compliance, acceptability, and reactivity associated with the EMA method for both risk group (n=300; endorsed recent prescription drug misuse of one or more medications in the past 3 months) and nonrisk group (n=55; did not endorse recent prescription drug misuse) participants; and (3) establish initial construct validity by linking the reports of misuse behaviors in daily life collected via the EMA app to prescription drug misuse reported on a standard survey.

Methods

Overview

Participants attended two laboratory sessions that were scheduled an average of 35 days apart, with the EMA protocol

XSL•FO
RenderX

implemented in between these laboratory sessions. Before this study, university institutional review board approval and the National Institutes of Health Certificate of Confidentiality were obtained. Participants completed questionnaires in each laboratory session, including an EMA feedback form to describe their reporting experiences.

Participants and Procedures

Between September 2017 and September 2019, participants at a large university in the Midwestern United States were continuously enrolled into an ongoing longitudinal study on daily behaviors and health in college life. This study is drawn from the baseline phase of a broader study [28]. Participants were recruited via flyers and announcements (eg, newspaper advertisements, emails to enrolled students) that stated, "We are particularly interested in how people use prescription medications." Prospective participants completed a web-based screening, and a telephone call was scheduled to confirm their eligibility. Inclusion criteria for all participants were (1) being enrolled as first- and second-year college students (verified against the campus Registrar database) and (2) being 18-21 years of age; eligibility for the risk subsample also required (3) endorsing recent prescription drug misuse of one or more medications.

The web-based screening obtained prospective participants' indications that they were in a private location and agreed to complete the screening questions, which asked about potentially sensitive day-to-day behaviors. Participants were instructed to think back over the past 3 months and indicate whether they used the medications listed in any way a doctor did not intend, such as use without a prescription, increased amounts, more often, or longer than directed. The screener presented 4 prescription medication classes adapted from the 2015 National Survey on Drug Use and Health (NSDUH) [29]: pain relievers, tranquilizers, stimulants, and sedatives or barbiturates, and common examples were provided. Prospective participants responded to each medication question (0=no; 1=yes), and multiple classes could be endorsed.

Given the main objective of capturing prescription drug misuse in daily life during this study, we oversampled for participants who endorsed recent prescription drug misuse (ie, risk group). This study included 300 participants from the risk group and 55 participants who did not endorse recent prescription drug misuse in the screener (ie, the nonrisk comparison group). To minimize any differences in recruitment between groups, the nonrisk participants were enrolled simultaneously with the risk participants throughout the recruitment period. Participants completed informed consent procedures at the start of their first laboratory session. The typical EMA reporting period was scheduled for 28 days following the first laboratory session. The length of the reporting period was adjusted for some participants due to scheduling conflicts (eg, campus recess periods) or device issues. During the second laboratory session, participants returned their devices and completed additional measures. Participants received their choice of electronic or check payments; compensation included US $75 for session 1, US $84 for reporting in daily life (prorated for partial completion), US $55 for session 2, and a US $36 bonus for maintaining compliance across the planned reporting period.

App Development and Implementation

The app was developed in collaboration with the university's technology division and was installed on an Apple sixth-generation iPod Touch device. Before the initiation of this study, the research team met with developers several times to provide feedback on the app features and interface design. The survey was administered through an app presented across multiple screens, with question completion automatically advancing the participant to the next screen. The questions allowed for different user inputs (eg, drop-down, radio button, and checkbox). Throughout the course of this study, the developers provided technical assistance related to software updates and other questions that arose. Developers also provided training to the research team on how to program the app and device for individual participants.

Given the sensitive nature of the data, security was a primary concern. We maximized the security of the data in several ways. First, the app was installed locally on laboratory-owned devices, which were provided to all participants for the duration of their reporting periods. Second, the data collected from the app were stored directly on the devices; the resulting data files were downloaded directly from the device to a secure server when participants returned to the laboratory (ie, never transmitted via the internet). Third, all data were recorded using a study identifier that contained no personally identifying information. Fourth, devices were also placed in restricted mode, which prevented participants from accessing a completed report or using other device functions. In addition, wireless transfer of data to and from the device was prevented. Finally, responses were all numeric and stored on the device with nondescript variable names. The list of variables was kept separate and accessible only by laboratory members. In the unlikely event someone gained access to the completed surveys, there would be no indication of what the data represented.

Consistent with numerous EMA protocols on addictions [30-33], both signal- and event-contingent assessments were administered [34]. Signal-contingent reporting involved a device prompt (ie, a notification) 4 times a day, once during each of the following periods: 8:00 AM to 11:30 AM, 11:30 AM to 3:00 PM, 3:00 PM to 7:00 PM, and 7:00 PM to 11:00 PM. The signal times within each window varied daily. Participants were also trained to self-initiate a report any time they intended to take a medication listed in any way a doctor did not direct them to use it (described under *Measures*). Signal- and event-based EMA report items were identical. Reports took approximately 2 min to complete. To reduce participant burden, a signal-contingent prompt was not sent within 2 h after a report had been completed. These EMA reports included items about intentions to misuse prescription drugs. If misuse intention of one or more of the medication classes was endorsed, participants were then sent a brief follow-up report 15 min later to assess misuse behaviors that might have occurred since the completion of the associated report; previous EMA protocols [35] have captured behaviors and contextual factors within this time frame. Acknowledging that participants might need to wait for a more

XSL•FO
RenderX

convenient or private time to respond, they were instructed to complete the follow-up within 15 min. The app automatically recorded the date and time of the start and completion of all reports and follow-ups.

During the first session, participants were thoroughly trained to complete the app reporting procedures. A comprehensive definition of the focal behavior was provided (ie, using a medication without a prescription of your own; using it in greater amounts, more often, or longer than you were told to take it, or using it in any other way a doctor did not direct you to use it); the term *prescription drug misuse* was not used. Participants were trained to provide both signal- and event-contingent reports and the follow-up reports. They were also instructed to contact the laboratory staff if they experienced any problems with the device. Participants completed a practice report during their laboratory session and were instructed to carry the device at all times.

Measures

Sociodemographics

During the first laboratory session, participants reported on their gender, race, and ethnicity. In addition, the participants reported fraternity or sorority affiliation and medication prescriptions they had received in their lifetime.

EMA and Follow-Up Reports

EMA reports captured prescription drug misuse intentions. Participants were asked "Are you about to take a medication listed here in any way a doctor did not direct you to use it?" (0=no; 1=yes). Four medication classes and examples were provided (pain relievers, tranquilizers, stimulants, and sedatives or barbiturates; 0=no; 1=yes). Reports also collected contextual information about the current location of participants and who they were with as well as hypothesized predictors of prescription drug misuse, including other substance use, mood, pain, fatigue, and stressful events. To strengthen the predictive design, reports assessed current feelings and situations; substance use covered behaviors in the past 15 min. Items were selected from brief, validated scales used in previous research with college-based populations when possible. The number of items on each report could vary because questions used conditional answer choices (eg, skip logic) whenever possible. Follow-up reports presented the following question for each of the 4 medication classes: "Have you recently taken a medication listed here, in any way a doctor did not direct you to use it?" (0=no; 1=yes). Additional questions about current mood states and recent physical activities were included in the follow-up reports.

Participant Feedback on the EMA Protocol (Acceptability)

During their second laboratory session, participants reported on their experiences with the app and device procedures. Participants rated the extent to which "the reporting occurred during a period that reflects my typical daily life" (0=not typical; 3=very typical) and the extent to which "the iPod touch device was user-friendly" (0=not friendly; 3=very friendly). Participants rated their broader study experiences with the following: "If I contacted the lab, my questions were addressed in a helpful and timely manner" (0=not helpful; 3=very helpful) and "Would you recommend this research opportunity to your friends?" (0=no; 1=yes).

Past-Year Misuse of Prescription Drugs

During the baseline visit, participants completed a series of questions about the use of prescription medications following the 2015 NSDUH [29]. For this study's 4 focal prescription medications, participants first indicated whether or not they had used any of the medications in the past 12 months. Following these initial questions, participants received a follow-up question for each endorsed medication asking if the use occurred in any way not directed by a doctor. Responses (0=no; 1=yes) were scored to reflect whether the participant engaged in any prescription drug misuse in the past year as well as any misuse of each of the 4 focal medication classes.

Analysis Plan

We employed independent samples *t* tests and chi-square analyses to compare demographic information, completion rates, and acceptability indicators of participants in the risk and nonrisk groups. Throughout this study, we report Fisher exact test when any single cell was based on <5 participants. Reactivity was tested by correlating indicators of time and length (ie, the number of days of reporting and the number of reports completed) with the outcomes of interest (ie, reports of prescription drug misuse intentions and behavior). Initial construct validity was examined by relating past-year survey reports of prescription drug misuse assessed during the study baseline to misuse behavior on the EMA. Given that signal- and event-based EMA report items were identical (and thus were indistinguishable in the resultant data files), both types of reports were included in the completion rates and construct validity results.

Results

Sample Characteristics

Participants in the risk and nonrisk groups did not differ with respect to the demographic characteristics of age or class standing (Table 1). The risk-group sample included proportionally more participants who self-identified as female; non-Hispanic, White (compared with other racial or ethnic backgrounds); and affiliated with a fraternity or sorority. Risk participants were also more likely to have a lifetime history of prescriptions (of any medication class) and of pain relievers, tranquilizers, and stimulant medications.

XSL•FO
RenderX

Table 1. Sample characteristics at baseline and tests of risk (n=300) and nonrisk (n=55) group differences.

Characteristics	Risk participants	Nonrisk participants	Statistical comparison		*P* value
			t test (*df*)	Chi-square (*df*)	
Age (years), mean (SD)	19.5 (0.71)	19.36 (0.68)	1.38 (353)	N/A[a]	.17
Class standing (freshman), n %	169 (56.3)	35 (64)	N/A	1.0 (1)	.31
Sex (female), n %	207 (69)	30 (55)	N/A	4.4 (1)	.04
Race and ethnicity (non-Hispanic, White), n %[b]	240 (80.3)	32 (58)	N/A	12.7 (1)	<.001
Fraternity/sorority member (affiliated), n %	107 (35.7)	6 (11)	N/A	13.1 (1)	<.001
Lifetime prescription history (endorsed any), n %	207 (69)	22 (40)	N/A	17.1 (1)	<.001
Pain reliever (endorsed)	151 (50.3)	18 (33)	N/A	5.8 (1)	.02
Tranquilizer (endorsed)	66 (22)	5 (9)	N/A	4.8 (1)	.03
Stimulant (endorsed)	53 (17.7)	2 (4)	N/A	N/A	.007[c]
Sedative or barbiturate (endorsed)	23 (7.7)	1 (2)	N/A	N/A	.15[c]

[a]N/A: not applicable.

[b]Response missing for 1 risk participant.

[c]Fisher exact test.

Data Selection and Descriptive Statistics

A total of 28,701 EMA reports were completed by 352 of the 355 participants; data were not available from 3 risk-group participants (2 did not return their data collection device and data from 1 participant's device was not retrievable due to a password error). An important feature of the EMA protocol was collecting self-initiated reports; therefore, the app did not restrict time between reporting and allowed participants to complete several surveys within a short time frame. Upon examination of the collected data, the study team deemed it reasonable for 2 reports to be completed within 5 min of each other. Any reports beyond the second that were completed within 5 min (ie, the third or greater report) were removed. This resulted in the removal of 1.1% (329/28,701) of the obtained reports that we did not plan to include in any analysis stemming from the broader project.

Some report submissions did not adhere to the EMA protocol. Given our goal of documenting the utility of the protocol, we focus only on the reports that were obtained during the scheduled reporting periods and only the follow-ups that were completed within the instructed time frame. Although the typical reporting period was scheduled for the 28 days following the first laboratory session, some participants continued reporting until they returned for the second laboratory session. We removed a total of 2147 reports that were completed outside of the participants' designated reporting days. The resulting 91.4% (26,225/28,701) of obtained reports were associated with 439 completed follow-ups; 64.0% (281/439) of the follow-ups were completed within 15 min of being sent (following the associated report) and were retained in this analysis.

In line with our sampling strategy, risk-group participants (143/297, 48.1%) were significantly more likely than the nonrisk counterparts (4/55, 7%) to endorse any prescription drug misuse intentions in daily life across the four medication classes (Fisher exact $P<.001$). Risk participants (108/297, 36.4%) were also significantly more likely than nonrisk participants (0/55, 0%) to complete one or more follow-up reports according to the study protocol (Fisher exact $P<.001$). In addition, risk participants (91/297, 30.6%) were more likely than nonrisk participants (0/55, 0%) to endorse any prescription drug misuse behavior in daily life on the follow-ups (Fisher exact $P<.001$).

Compliance

Overall, participants demonstrated consistent engagement with the EMA procedures and returned an average of 74.5 reports (SD 23.82; range 10-122). Table 2 shows report completion by group status. Notably, participants in the risk and nonrisk groups did not vary in the number of reports they completed or in the number of their reporting days. We also calculated the completion rate for each person, which was based on the number of their completed reports relative to the expected number of reports (assigned number of reporting days × 4 reports per day); participants in risk and nonrisk groups did not differ in their average completion rates (Table 2).

XSL•FO
RenderX

Table 2. Ecological momentary assessment completion by risk (n=297) and nonrisk (n=55) group status.

Variables	Risk participants	Nonrisk participants	Statistical comparison, *t* test (*df*)	*P* value
Reports, mean (SD)	73.65 (24.07)	79.11 (22.01)	–1.57 (350)	.12
Reporting days, mean (SD)	26.50 (4.05)	27.07 (3.05)	–0.99 (350)	.32
Completion rate[a]	0.69 (0.19)	0.73 (0.18)	–1.48 (350)	.14

[a]Calculated as the number of completed reports divided by the expected number of reports based on assigned reporting days.

Acceptability

Participants in the risk versus nonrisk groups reported similar experiences with procedures and uniformly positive experiences. As shown in Table 3, there were no reliable differences across groups in terms of evaluations of whether the reporting period was typical of their daily life or whether the device was user friendly. There was no difference in the likelihood of participants from the different groups contacting the research team for assistance; furthermore, the staff contacts that occurred were evaluated as equally helpful by participants across the groups. Nearly all participants who attended the second laboratory session (352/353, 99.7%) endorsed that they would recommend the research opportunity to their friends; 1 participant from the risk group left his or her response to this question blank.

Table 3. Ecological momentary assessment experiences by risk (n=298) and nonrisk (n=55) group status.

Variables	Risk participants	Nonrisk participants	Statistical comparison, *t* test (*df*)	*P* value
Reflecting typical daily life[a], mean (SD)	2.13 (0.74)	2.33 (0.70)	–1.83 (351)	.07
User-friendly device[a], mean (SD)	2.87 (0.37)	2.84 (0.37)	0.61 (351)	.54
Contacted the research team, n (%)	118 (39.6)	28 (51)	2.5 (1)[b]	.12
Helpful contact[a], mean (SD)	2.94 (0.27)	2.86 (0.36)	1.16 (34.76)	.25

[a]Rated on scales from 0 to 3.

[b]This result is from χ^2, not a *t* test.

Reactivity

We examined reactivity by testing whether prescription drug misuse intentions and behavior were more or less likely to be endorsed by participants over time, in terms of their reporting length (the number of days) and reporting amount (the number of reports). Specifically, reactivity was documented with prescription drug misuse intentions decreasing as a function of the number of reporting days (r=–0.01; P=.04) and number of reports completed (r=–0.03; P<.001). In contrast, when prescription drug misuse behavior occurred in daily life, reactivity in the opposite direction was found, with misuse behavior endorsed on the follow-up increasing as a function of reporting day (r=0.15; P=.01) and report number (r=0.18; P=.003), that is, we found evidence that prescription drug misuse intentions decreased over time (or were *less likely* to be endorsed as participants completed more days of reporting and more reports), whereas actual prescription misuse behavior increased over time (or was *more likely* to be endorsed as participants completed more reporting days and more reports). It should be noted that the magnitude of reactivity effect sizes for intentions was negligible and that for behavior was small [36].

Initial Construct Validity

There was a robust association between participants engaging in prescription drug misuse in the past year at the baseline assessment and during the daily life procedure: Reporting misuse behavior (of one or more medication classes) in daily life on the EMA follow-up was significantly more likely among participants who had indicated any past-year prescription drug misuse (90/255, 35.3%) than among those who did not (1/97, 1%; Fisher exact P<.001). Examination by medication class further revealed reliable consistency in participants reporting prescription misuse behavior across the methods. Specifically, the results indicated that the proportion of participants who endorsed misuse of a medication class in daily life according to the study protocol was significantly higher among those who had reported past-year misuse of that medication class than among those who did not for all 4 medication classes: pain medication (6/23, 26% vs 2/85, 2%; Fisher exact P=.001), tranquilizers (9/29, 31% vs 6/79, 8%; χ^2_1=9.8; P=.002), stimulants (76/95, 80% vs 1/13, 8%; Fisher exact P<.001), and sedatives or barbiturates (2/7, 29% vs 1/101, 1%; Fisher exact P=.01).

Discussion

Principal Findings

This study presented the development and implementation of an app installed on study-owned devices to collect participants' momentary experiences at the time of their prescription drug misuse intentions and behaviors. This study documented participants' EMA completion rates and provided evidence for compliance and acceptability among participants in both risk and nonrisk sampling groups. Low levels of reactivity have been reported (described later). In addition, we presented the initial construct validity of our EMA approach by linking reports of misuse behavior in daily life to prescription drug misuse reported on a standard survey. Together, the results of this study

XSL•FO
RenderX

provide early support for the feasibility and value of collecting young adults' reports of prescription drug misuse intentions and behavior in daily life.

Although EMA still relies on self-report, this method offers improvement over more traditional assessments by minimizing reliance on memory and collecting situational factors at the time of the behavior of interest. Accordingly, these designs can provide specific momentary information to identify *more particular* risk and protective factors associated with behaviors that hold important consequences for health and adjustment. This protocol is indeed highly novel for collecting hypothesized predictors and prescription drug misuse intentions *before* the follow-up reports of the behavior. The resulting responses can allow many contextually based and temporally precise investigations across levels (ie, broader calendar-based timing as well as characteristics of the individual or the moment) to identify direct and interactive situations that increase college students' likelihood of endorsing prescription drug misuse in daily life. Findings from our investigations, therefore, can provide highly translatable information for modifiable prevention and treatment targets.

Our sample was highly committed to the overall study and was particularly engaged with reporting in daily life. From the start of recruitment, interest in the research was high. We maintained continual interest from potential participants. Positive engagement was further reflected by nearly all participants returning to the second laboratory session with their device and saying they would recommend the experience to others. We obtained average completion rates that were similar to a pooled compliance rate on the basis of many EMA protocols that examined different substances [37]. We intentionally recruited and enrolled nonrisk participants into the study at proportionally the same rate that we enrolled the risk-group participants; therefore, academic rhythm and calendar timing were similar and broader societal and campus events that occurred were the same across these two groups. Notably, this study showed that participants from risk and nonrisk groups also had similar experiences with the reporting in daily life. Therefore, when testing hypothesized predictors of the focal substance use outcomes, we will be more confident in the findings.

The results indicated some evidence of reactivity to the momentary reporting procedure, as has been found in reporting other substances using EMA approaches (eg, alcohol) [32]. It is possible that participants were more attentive to the procedure at the beginning of their reporting periods and, therefore, were more likely to endorse misuse intentions in daily life earlier on during reporting. At the same time, it could be that reporting over time raised participants' likelihood of engaging in the behavior or that participants became less concerned about endorsing sensitive behaviors as they become more accustomed to the reporting procedures. We are not able to determine whether reactivity varied across signal-based versus event-based reports; however, collecting such information in future studies would shed light on participants' preferences for one type of reporting over the other. It is worth reiterating that the low magnitude of the reactivity documented for both misuse intentions and behavior suggests that the outcomes did not practically change over time as a function of having to report on behavior in real time.

In testing adherence to the method here, we conservatively focused on reports that matched protocol training time frames. The reports we plan to include in future studies with this data set will be preregistered on the Open Science Framework and will depend on particular research questions. When we investigate momentary experiences as predictors of prescription misuse behaviors in daily life, we plan to restrict analysis of the focal misuse behavior outcomes to those that were reported on follow-ups completed within 15 min of being sent (consistent with protocol training). Alternatively, studies that are designed to test background factors, or other person-level characteristics, associated with risk of engaging in prescription drug misuse in daily life would not need to be bound by these temporal restrictions. Indeed, for certain research aims, having more instances of the behavior available would result in more representative findings and more powerful statistical tests. These trade-offs should be considered and justified by the purpose of this study.

We note that our EMA reports were identical regardless of whether participants were responding to a prompt or self-initiating a report. Thus, we could not determine the completion rates for signal-based versus event-based reporting. We were also unable to establish how many participants completed event-based assessments (or the associated follow-ups). Another protocol development decision was that we did not have participants indicate retrospectively whether prescription drug misuse had occurred since their last report, that is, a coverage strategy [38]. There is a possibility that participants who become aware of the backward-looking option would be less likely to report the focal behavior in real time. This decision was particularly important for our protocol because the primary goal of the broader study was to collect momentary factors that predict real-time prescription drug misuse; any motivations or characteristics reported *after* the behavior occurred would no longer represent contextual triggers of misuse. Instead, we focused on thoroughly training participants to respond to the prompts and self-initiate as described in the protocol. Studies focused on other research questions and aims should consider whether the addition of retrospective endorsement of the behavior would bolster their design. Consistent with the development of any rigorous study, careful planning of an EMA design that matches the research question is critical.

Limitations

Our sample was representative of the college campus from which it was drawn but is not necessarily reflective of students at other types of higher education institutions or of nontraditional students. Additional work will be needed to understand the applicability of this method to individuals across different contexts and locations. A practical limitation of using a separate device for EMA data collection is that the approach may be more burdensome for participants and requires greater investment in study materials. We expect that these considerations were offset by the anonymity and protection of the participant data provided by a separate device. Although

XSL•FO
RenderX

reported behavior instances might be lower if the device was not readily available to the participant, the responses obtained may be more forthright due to security and privacy assurances. Nevertheless, in light of the recent successful apps on young adults' own smartphones to assess drinking in daily life [32,39], this study provides a foundation for the next steps in adapting the study of prescription drug misuse in daily life to different designs. Future work could examine such possibilities.

Conclusions

Establishing real-time mood states, pain symptoms, stressors, and other substance behaviors as predictors of college students' prescription drug misuse is critical for obtaining a more precise understanding of the behavior in daily life situations. In line with the social ecological perspective [8,40], our models will include various triggers to identify the most robust hypothesized factor. This study provided initial evidence that using an EMA design to obtain reports of prescription drug misuse intentions and behavior (along with contextual antecedents) in daily life is feasible and acceptable. Additional research will provide direct insights into educational and intervention prospects for reducing hazardous prescription drug behaviors during an important age period.

Acknowledgments

This work was supported by the National Institute on Drug Abuse of the National Institutes of Health under Award Number R01DA042093. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. McHugh RK, Nielsen S, Weiss RD. Prescription drug abuse: from epidemiology to public policy. J Subst Abuse Treat 2015 Jan;48(1):1-7 [FREE Full text] [doi: 10.1016/j.jsat.2014.08.004] [Medline: 25239857]
2. Sheridan J, Butler R. Prescription drug misuse in New Zealand: challenges for primary health care professionals. Res Social Adm Pharm 2011 Sep;7(3):281-293. [doi: 10.1016/j.sapharm.2010.06.005] [Medline: 21272526]
3. Paulozzi LJ. Prescription drug overdoses: a review. J Safety Res 2012 Sep;43(4):283-289. [doi: 10.1016/j.jsr.2012.08.009] [Medline: 23127678]
4. Linas BS, Genz A, Westergaard RP, Chang LW, Bollinger RC, Latkin C, et al. Ecological momentary assessment of illicit drug use compared to biological and self-reported methods. JMIR Mhealth Uhealth 2016 Mar 15;4(1):e27 [FREE Full text] [doi: 10.2196/mhealth.4470] [Medline: 26980400]
5. Serre F, Fatseas M, Debrabant R, Alexandre J, Auriacombe M, Swendsen J. Ecological momentary assessment in alcohol, tobacco, cannabis and opiate dependence: a comparison of feasibility and validity. Drug Alcohol Depend 2012 Nov 1;126(1-2):118-123. [doi: 10.1016/j.drugalcdep.2012.04.025] [Medline: 22647899]
6. Serre F, Fatseas M, Swendsen J, Auriacombe M. Ecological momentary assessment in the investigation of craving and substance use in daily life: a systematic review. Drug Alcohol Depend 2015 Mar 1;148:1-20. [doi: 10.1016/j.drugalcdep.2014.12.024] [Medline: 25637078]
7. Compton WM, Volkow ND. Abuse of prescription drugs and the risk of addiction. Drug Alcohol Depend 2006 Jun;83(Suppl 1):S4-S7. [doi: 10.1016/j.drugalcdep.2005.10.020] [Medline: 16563663]
8. Quintero G, Peterson J, Young B. An exploratory study of socio-cultural factors contributing to prescription drug misuse among college students. J Drug Issue 2016 Aug 3;36(4):903-931. [doi: 10.1177/002204260603600407]
9. Kelly BC, Wells BE, Pawson M, LeClair A, Parsons JT. Combinations of prescription drug misuse and illicit drugs among young adults. Addict Behav 2014 May;39(5):941-944 [FREE Full text] [doi: 10.1016/j.addbeh.2013.12.003] [Medline: 24462348]
10. McCabe SE, Knight JR, Teter CJ, Wechsler H. Non-medical use of prescription stimulants among US college students: prevalence and correlates from a national survey. Addiction 2005 Jan;100(1):96-106. [doi: 10.1111/j.1360-0443.2005.00944.x] [Medline: 15598197]
11. Ali MM, Dean D, Lipari R, Dowd WN, Aldridge AP, Novak SP. The mental health consequences of nonmedical prescription drug use among adolescents. J Ment Health Policy Econ 2015 Mar;18(1):3-15. [Medline: 25862204]
12. Arria AM, DuPont RL. Nonmedical prescription stimulant use among college students: why we need to do something and what we need to do. J Addict Dis 2010 Oct;29(4):417-426 [FREE Full text] [doi: 10.1080/10550887.2010.509273] [Medline: 20924877]
13. Benotsch EG, Koester S, Luckman D, Martin AM, Cejka A. Non-medical use of prescription drugs and sexual risk behavior in young adults. Addict Behav 2011;36(1-2):152-155. [doi: 10.1016/j.addbeh.2010.08.027] [Medline: 20863626]
14. McCabe SE, Cranford JA, Boyd CJ. The relationship between past-year drinking behaviors and nonmedical use of prescription drugs: prevalence of co-occurrence in a national sample. Drug Alcohol Depend 2006 Oct 1;84(3):281-288 [FREE Full text] [doi: 10.1016/j.drugalcdep.2006.03.006] [Medline: 16621337]

XSL•FO
RenderX

15. McLarnon ME, Barrett SP, Monaghan TL, Stewart SH. Prescription drug misuse across the lifespan: a developmental perspective. In: Verster JC, Brady K, Galander M, Conrod P, editors. Drug Abuse and Addiction in Medical Illness. New York, USA: Springer; 2012:213-230.
16. Whiteside LK, Cunningham RM, Bonar EE, Blow F, Ehrlich P, Walton MA. Nonmedical prescription stimulant use among youth in the emergency department: prevalence, severity and correlates. J Subst Abuse Treat 2015 Jan;48(1):21-27 [FREE Full text] [doi: 10.1016/j.jsat.2014.05.003] [Medline: 25012553]
17. Johnston LD, O'Malley PM, Bachman JG. Monitoring the Future: National Survey Results on Drug Use, 1975-2002, Volume II: College Students and Adults Ages 19-40. Bethesda, MD: National Institutes of Health; 2003. URL: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/137801/vol2_2002.pdf?sequence=1
18. Arria AM, Wilcox HC, Caldeira KM, Vincent KB, Garnier-Dykstra LM, O'Grady KE. Dispelling the myth of 'smart drugs': cannabis and alcohol use problems predict nonmedical use of prescription stimulants for studying. Addict Behav 2013 Mar;38(3):1643-1650 [FREE Full text] [doi: 10.1016/j.addbeh.2012.10.002] [Medline: 23254212]
19. Teter CJ, McCabe SE, LaGrange K, Cranford JA, Boyd CJ. Illicit use of specific prescription stimulants among college students: prevalence, motives, and routes of administration. Pharmacotherapy 2006 Oct;26(10):1501-1510 [FREE Full text] [doi: 10.1592/phco.26.10.1501] [Medline: 16999660]
20. White BP, Becker-Blease KA, Grace-Bishop K. Stimulant medication use, misuse, and abuse in an undergraduate and graduate student sample. J Am Coll Health 2006;54(5):261-268. [doi: 10.3200/JACH.54.5.261-268] [Medline: 16539218]
21. Schulenberg JE, Maggs JL. A developmental perspective on alcohol use and heavy drinking during adolescence and the transition to young adulthood. J Stud Alcohol Suppl 2002 Mar(14):54-70. [doi: 10.15288/jsas.2002.s14.54] [Medline: 12022730]
22. Brown SA, McGue M, Maggs J, Schulenberg J, Hingson R, Swartzwelder S, et al. A developmental perspective on alcohol and youths 16 to 20 years of age. Pediatrics 2008 Apr;121(Suppl 4):S290-S310 [FREE Full text] [doi: 10.1542/peds.2007-2243D] [Medline: 18381495]
23. McCabe SE, West BT, Teter CJ, Boyd CJ. Trends in medical use, diversion, and nonmedical use of prescription medications among college students from 2003 to 2013: connecting the dots. Addict Behav 2014 Jul;39(7):1176-1182 [FREE Full text] [doi: 10.1016/j.addbeh.2014.03.008] [Medline: 24727278]
24. Meisel MK, Goodie AS. Predicting prescription drug misuse in college students' social networks. Addict Behav 2015 Jun;45:110-112. [doi: 10.1016/j.addbeh.2015.01.025] [Medline: 25661988]
25. Snipes DJ, Jeffers AJ, Benotsch EG, McCauley J, Bannerman D, Granger C, et al. Religiosity in the non-medical use of prescription medication in college students. Am J Drug Alcohol Abuse 2015 Jan;41(1):93-99. [doi: 10.3109/00952990.2014.939755] [Medline: 25105761]
26. Wu L, Pilowsky DJ, Schlenger WE, Galvin DM. Misuse of methamphetamine and prescription stimulants among youths and young adults in the community. Drug Alcohol Depend 2007 Jul 10;89(2-3):195-205 [FREE Full text] [doi: 10.1016/j.drugalcdep.2006.12.020] [Medline: 17257780]
27. Papp LM, Blumenstock SM. Momentary affect and risky behavior correlates of prescription drug misuse among young adult dating couples: an experience sampling study. Addict Behav 2016 Feb;53:161-167. [doi: 10.1016/j.addbeh.2015.10.016] [Medline: 26540088]
28. Papp LM, Kouros CD, Curtin JJ. Real-time associations between young adults' momentary pain and prescription opioid misuse intentions in daily life. Am Psychol 2020 Sep;75(6):761-771. [doi: 10.1037/amp0000648] [Medline: 32915021]
29. Prescription Drug Use and Misuse in the United States: Results From the 2015 National Survey on Drug Use and Health. Substance Abuse and Mental Health Services Administration. 2017. URL: https://www.samhsa.gov/data/sites/default/files/NSDUH-FFR2-2015/NSDUH-FFR2-2015.htm [accessed 2020-06-18]
30. Buckner JD, Zvolensky MJ, Crosby RD, Wonderlich SA, Ecker AH, Richter A. Antecedents and consequences of cannabis use among racially diverse cannabis users: an analysis from ecological momentary assessment. Drug Alcohol Depend 2015 Feb 1;147:20-25 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.12.022] [Medline: 25578250]
31. Cooney NL, Litt MD, Cooney JL, Pilkey DT, Steinberg HR, Oncken CA. Alcohol and tobacco cessation in alcohol-dependent smokers: analysis of real-time reports. Psychol Addict Behav 2007 Sep;21(3):277-286 [FREE Full text] [doi: 10.1037/0893-164X.21.3.277] [Medline: 17874878]
32. Poulton A, Pan J, Bruns LR, Sinnott RO, Hester R. A smartphone app to assess alcohol consumption behavior: development, compliance, and reactivity. JMIR Mhealth Uhealth 2019 Mar 25;7(3):e11157 [FREE Full text] [doi: 10.2196/11157] [Medline: 30907738]
33. Preston KL, Epstein DH. Stress in the daily lives of cocaine and heroin users: relationship to mood, craving, relapse triggers, and cocaine use. Psychopharmacology (Berl) 2011 Nov;218(1):29-37 [FREE Full text] [doi: 10.1007/s00213-011-2183-x] [Medline: 21336579]
34. Wheeler L, Reis H. Self-recording of everyday life events: origins, types, and uses. J Pers 1991;59(3):339-354. [doi: 10.1111/j.1467-6494.1991.tb00252.x]
35. Thrul J, Bühler A, Ferguson SG. Situational and mood factors associated with smoking in young adult light and heavy smokers. Drug Alcohol Rev 2014 Jul;33(4):420-427. [doi: 10.1111/dar.12164] [Medline: 24893955]
36. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]

XSL•FO
RenderX

37. Jones A, Remmerswaal D, Verveer I, Robinson E, Franken IH, Wen CK, et al. Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. Addiction 2019 Apr;114(4):609-619 [FREE Full text] [doi: 10.1111/add.14503] [Medline: 30461120]
38. Shiffman S. Designing protocols for ecological momentary assessment. In: Stone AA, Shiffman S, Atienza AA, Nebeling L, editors. The Science of Real-time Data Capture: Self-reports in Health Research. New York, USA: Oxford University Press; 2007:27-53.
39. Fridberg DJ, Faria J, Cao D, King AC. Real-time mobile monitoring of drinking episodes in young adult heavy drinkers: development and comparative survey study. JMIR Mhealth Uhealth 2019 Nov 20;7(11):e13765 [FREE Full text] [doi: 10.2196/13765] [Medline: 31746774]
40. Nargiso JE, Ballard EL, Skeer MR. A systematic review of risk and protective factors associated with nonmedical use of prescription drugs among youth in the United States: a social ecological perspective. J Stud Alcohol Drugs 2015 Jan;76(1):5-20. [Medline: 25486389]

Abbreviations

EMA: ecological momentary assessment
NSDUH: National Survey on Drug Use and Health

Edited by G Eysenbach; submitted 21.06.20; peer-reviewed by F Serre, F Lamers; comments to author 15.07.20; revised version received 11.08.20; accepted 02.09.20; published 01.10.20.

Please cite as:
Papp LM, Barringer A, Blumenstock SM, Gu P, Blaydes M, Lam J, Kouros CD
Development and Acceptability of a Method to Investigate Prescription Drug Misuse in Daily Life: Ecological Momentary Assessment Study
JMIR Mhealth Uhealth 2020;8(10):e21676
URL: https://mhealth.jmir.org/2020/10/e21676
doi:10.2196/21676
PMID:32877351

©Lauren M Papp, Alexandra Barringer, Shari M Blumenstock, Pamela Gu, Madison Blaydes, Jaime Lam, Chrystyna D Kouros. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 01.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

App-Based Delivery of Clinical Emotional Freedom Techniques: Cross-Sectional Study of App User Self-Ratings

Dawson Church[1*], PhD; Peta Stapleton[2*], PhD; Debbie Sabot[2*], Grad Dip

[1]National Institute for Integrative Healthcare, Fulton, CA, United States

[2]School of Psychology, Faculty of Society & Design, Bond University, Robina, Australia

[*]all authors contributed equally

Corresponding Author:
Debbie Sabot, Grad Dip
School of Psychology
Faculty of Society & Design
Bond University
14 University Drive
Robina, 4226
Australia
Phone: 61 755952515
Email: dsabot@bond.edu.au

Abstract

Background: The burgeoning area of mobile health (mHealth) has experienced rapid growth in mobile apps designed to address mental health issues. Although abundant apps offer strategies for managing symptoms of anxiety and stress, information regarding their efficacy is scarce.

Objective: This study aimed to assess the effect of an mHealth app on user self-ratings of psychological distress in a sample of 270,461 app users. The Tapping Solution App guides users through the therapeutic protocols of Clinical Emotional Freedom Techniques (EFT), an evidence-based psychophysiological intervention that combines acupressure with elements of cognitive and exposure therapies.

Methods: App users provided self-ratings of emotional intensity before and after app sessions (termed "tapping meditations") using an 11-point Subjective Units of Distress scale. App user data for 23 tapping meditations, which addressed psychological symptoms of anxiety and stress, were gathered between October 2018 and October 2019, totaling 380,034 completed app sessions.

Results: Across 12 anxiety-tapping meditations, the difference in emotional intensity ratings from presession (mean 6.66, SD 0.25) to postsession (mean 3.75, SD 0.30) was statistically significant (P<.001; 95% CI −2.92 to −2.91). Across 11 stress-tapping meditations, a statistically significant difference was also found from presession (mean 6.91, SD 0.48) to postsession (mean 3.83, SD 0.54; P<.001; 95% CI −3.08 to −3.07). The results are consistent with the literature on the efficacy of Clinical EFT for anxiety and stress when offered in conventional therapeutic formats.

Conclusions: The findings provide preliminary support for the effectiveness of the mHealth app in the immediate reduction of self-rated psychological distress. As an adjunct to professional mental health care, the app promises accessible and convenient therapeutic benefits.

(JMIR Mhealth Uhealth 2020;8(10):e18545) doi:10.2196/18545

KEYWORDS

anxiety; stress; meditation; mobile health; Emotional Freedom Techniques (EFT); mobile phone

Introduction

Background

Symptoms of anxiety and psychological distress are highly prevalent in the adult population worldwide. Anxiety disorders are among the most common mental health disorders [1], with an estimated one-third of the global population affected by an anxiety disorder during their lifetime [2,3]. Furthermore, subclinical symptoms of anxiety are reported globally, which can significantly impair functioning and reduce quality of life [4]. Psychological stress is also a commonly reported mental health issue. More than 75% of American adults perceive

XSL•FO
RenderX

themselves as significantly stressed, with 42% expressing a desire to manage their stress [5].

Chronic exposure to stress is associated with enduring changes in the body's emotional, physiological, and behavioral responses [6]. These changes present an increased risk for mental health disorders and diseases such as depression, cardiovascular disease, autoimmune dysfunction, and cancer [7]. Several pathways link psychological distress to disease. Maladaptive behavioral changes in response to stress (eg, inadequate sleep, poor diet) and biological changes in the endocrine response system (eg, hypothalamic-pituitary-adrenocortical axis, sympathetic-adrenal-medullary) can increase individuals' risk of disease [8,9]. As a result, adverse psychological health imposes a substantial economic burden at individual and societal levels [10,11]. The very high levels of anxiety and psychological stress reported globally highlight the need for accessible psychological treatments with demonstrated efficacy to help reduce the behavioral and biological outcomes associated with poor mental health [7].

Mobile Health

Mobile technology innovation has significantly transformed aspects of everyday life. This technological platform has altered the way we consume entertainment, educate ourselves, and communicate with each other [12] by broadening access to services and increasing resource availability. Mobile technologies in the field of mental health have the potential to revolutionize traditional health care [13]. Mobile health (mHealth) is an emerging field in psychological health practice that uses wireless technologies supported by smartphones and mobile devices [14]. mHealth apps are considered as a new taxonomy of techniques that help to manage psychological distress [15] and enable users to work independently on aspects of self-improvement. As an adjunct to psychotherapy, mHealth apps offer increased access and availability than face-to-face health care [16].

The burgeoning growth in mental health apps has been largely attributed to rapid technological development, together with its convenience and ease of use [17,18]. Although mHealth apps (eg, medical, health, and fitness categories) currently account for 3.4% of the 3 million plus apps available via the Apple Store, an estimated 60% of smartphone users in the United States have at least one mHealth app installed on their smartphones [19]. There is an abundance of commercial apps offering users strategies and techniques for managing anxiety and stress. However, little information on the efficacy of these apps is available beyond self-rated reviews and star ratings [20]. The proliferation of mobile apps to address mental health calls for timely evaluation of their psychological benefits for app users.

Apps Designed for Anxiety

Several recent meta-analyses have comprehensively appraised the field of mHealth apps with particular focus on apps designed to manage symptoms of anxiety. Sucala et al [21] analyzed 52 apps available through European app stores (iTunes and Google Play), which targeted anxiety in general, worry, and/or panic attacks. In 63.5% of the apps analyzed, no information was given to users about the therapeutic method that informed its design. Of the apps that identified a therapeutic method, 26.9% were aligned with cognitive behavioral therapy, while 7.7% reported a combination of therapies (eg, meditation, mindfulness, neuro-linguistic programming). The majority of apps failed to disclose details regarding professional licensure and developer training. Of the 52 anxiety-based apps reviewed, Sucala et al [21] identified only two studies that presented feasibility and efficacy data [22,23], both of which suggested that the apps effectively reduced symptoms of anxiety. However, both the respective studies had notable limitations regarding their research design (eg, lack of empirically validated measures or high participant attrition). Accordingly, Sucala et al [21] recommended that the anxiety app design be grounded in psychotherapeutic techniques with demonstrated efficacy in face-to-face clinical settings. Furthermore, they cautioned that apps not grounded in empirical approaches can result in iatrogenic effects, thereby increasing symptoms of anxiety in app users.

The efficacy of anxiety-based smartphone-supported apps was also examined in a meta-analysis of 9 randomized controlled trials (RCTs) that met specific systematic review criteria [24]. Collectively, a significant reduction in anxiety symptoms was found in the anxiety app intervention groups compared with controls (N=1837), with the greatest benefits observed in the trials that compared smartphone interventions with wait list control conditions. Although significantly smaller effects were observed in studies that controlled for attention or user engagement, Firth et al [24] concluded that smartphone interventions appeared to reduce anxiety symptoms significantly more than controls. However, there was substantial variance in anxiety levels between and within study participants, indicating the need for research to identify specific user groups who may benefit most from anxiety interventions delivered via mobile device platforms.

Apps Designed for Stress

Apps designed to help users manage psychological stress have also been the subject of meta-analytic review. In 2016, Coulon et al [25] provided the first meta-analysis of evidence-based stress management apps. A total of 902 apps available on the Apple iOS platform were subject to a multilevel selection process, of which 32 apps met 3 specific criteria: domains related to evidence-based content, transparency of app developer details, and functional app interface. The most common therapeutic techniques among the apps were mindfulness, meditation, and diaphragmatic breathing. Several apps purported to deliver efficacious stress management techniques (eg, breathing techniques), despite providing inadequate guidance for users (eg, lack of instruction regarding the use of diaphragm muscles during breathing exercises). Therefore, Coulon et al [25] cautioned that apps delivering evidence-based methods require adequate behavioral skill instruction to avoid iatrogenic effects on app users.

Coulon et al [25] were the first to apply an established taxonomy of behavior change techniques [26] in the review of stress management apps. In extending the research of Coulon et al [25], Christmann et al [15] proposed an additional taxonomy of emotion-focused stress management strategies in their review

XSL•FO
RenderX

of free stress management apps available on Google Play. Of the 62 apps that met their inclusion criteria, 26 apps comprised behavioral change–based strategies and 15 apps presented emotion-focused stress management techniques. One app was common to both the analyses by Christmann et al [15] and Coulon et al [25]. In contrast with the review by Coulon et al [25], in which 48% of stress management apps drew on mindfulness or meditation techniques, only one-third (34%) of apps reviewed by Christmann et al [15] used empirically demonstrated approaches (eg, meditation, mindfulness, breathing, acupressure or EFT). The standardized taxonomy proposed by Christmann et al [15] was designed to enable greater comparability between different intervention types in stress management apps. Interestingly, although some apps reviewed by Christmann et al [15] offered users the opportunity to self-rate symptoms and stress levels, none used that information to address the pattern of self-rated symptoms within app content. They considered this as an important area for future health app design. This self-rating functional aspect of app design is addressed in this study.

Meta-analytic findings provide direction for future mHealth app assessment and development. Collectively, this body of work reinforces a crucial principle: mHealth apps must demonstrate positive outcomes for app users [27]. However, few studies have examined the effectiveness of mHealth apps in reducing symptoms of psychological distress [12]. The mHealth platform offers the potential for a range of self-management strategies to assist psychological symptoms of anxiety and stress, particularly for individuals who require psychological support but have limited access to regular health care [24,28]. As the mHealth app modality offers benefits such as increased flexibility, accessibility, convenience, and reduced cost [12,29], studies that examine the effectiveness of evidence-based apps are paramount to help inform and protect the growing population of app users.

The Tapping Solution App

The Tapping Solution App, developed by The Tapping Solution, LLC, is an Energy Psychology–based meditation app for use on smartphones and mobile devices. The app was designed to improve users' symptoms of psychological distress (eg, anxiety, stress, worry) and promote overall well-being using Emotional Freedom Techniques (EFT). EFT is a therapeutic approach in the field of Energy Psychology, which combines elements of exposure and cognitive therapy together with somatic stimulation. In the EFT therapeutic protocol, the individual taps with the fingertips on specific acupoints on the body (acupressure) while focusing on cognitions that produce emotional distress [30]. This focus on emotionally charged memories and beliefs draws from the field of exposure therapy. When paired with acupoint tapping, the emotional intensity of these memories is usually quickly reduced [31]. Since its inception in the 1990s, EFT has been a manualized therapy, leading to uniformity of application in research and training. The manualized form of the method is called Clinical EFT.

The EFT procedure begins with clients identifying an issue and rating their degree of distress. The EFT uses an 11-point Likert scale ranging from 0 (no emotional intensity) to 10 (maximum emotional intensity). This scale is called the Subjective Units of Distress (SUD) scale and originates in the work of Wolpe [32]. Clients provide a phrase that encapsulates their issue, such as "the car crash" or "the explosion." This "reminder phrase" is repeated throughout treatment to maintain and reinforce exposure to the issue while the acupoints are stimulated. A long form of EFT includes eye movements similar to those used in eye movement desensitization and reprocessing [33] and stimulates 14 acupoints. An abbreviated version stimulates 9 points (8 points are displayed in Figure 1 [34], excluding the side of the hand acupoint). The EFT short form is completed in less than 30 seconds and is referred to as a "round" of tapping. The procedure is repeated until the SUD levels drop, which may require several rounds.

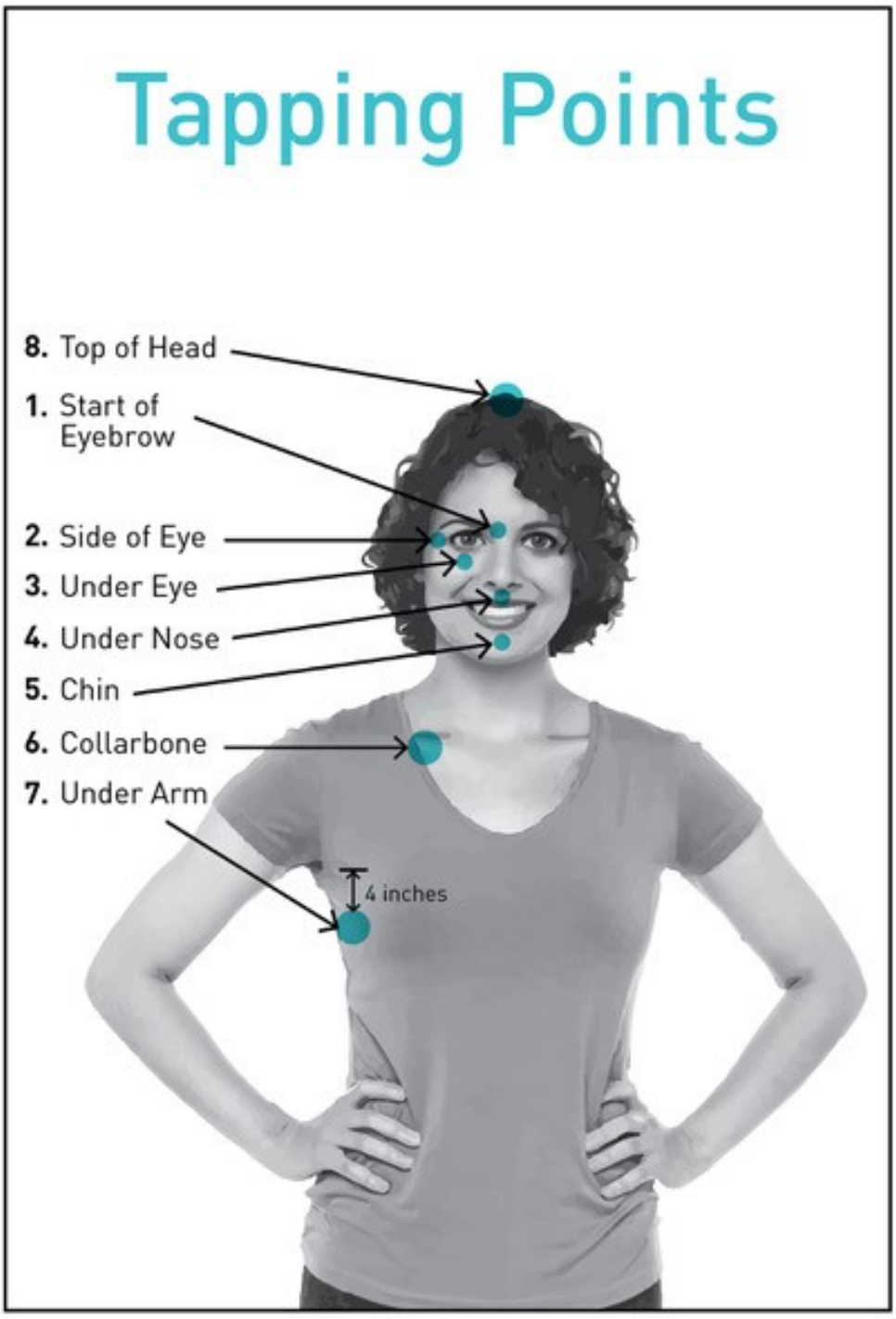
The psychological benefits of EFT intervention include improvements in symptoms of anxiety, posttraumatic stress disorder (PTSD), self-esteem, and pain [31,35-39]. Other studies have reported rapid improvements for a variety of additional psychological challenges such as performance blocks, social anxiety, excessive food cravings, and stress management [31,37,38,40-43]. A web-based research bibliography listing more than 100 clinical trials is publicly available [44].

In the 1990s, Division 12 (Clinical Psychology) of the American Psychological Association published standards for "empirically validated therapies" [45,46]. For the next two decades, the principles guided the design and reporting of EFT research [31]. Several studies have examined symptom levels before and after a single session of EFT (≤60 min in duration). These studies showed that EFT is effective for fear of public speaking [39,40,47], sports performance [48,49], anxiety and depression [50], phobic fear [51-53], and traumatic stress [54,55].

Although this study is the first app-based EFT study, it is important to note that several other studies have examined EFT delivered remotely. Hartung and Stein [56] compared the telephone delivery of EFT with in-person therapy. Although face-to-face delivery of EFT was significantly more efficacious than the telephone (91% vs 67% recovery rate), remote telephone sessions nonetheless remediated PTSD symptoms in 67% of the veterans treated. In a web-based EFT intervention of patients with fibromyalgia, Brattberg [57] found significant improvements in pain and other symptoms. Fibromyalgia was resolved in approximately one-third of the participants and another third reported partial pain relief. In addition, Church and Clond [58] compared participants in a web-based relationship class with a similarly sized sample taking the same class in-person. Although the relationship satisfaction outcomes were similar for both groups, they differed significantly on baseline measures of anxiety, depression, and relationship satisfaction. The authors suggested that the demographic and mental health characteristics of those seeking web-based treatment may differ substantially from those seeking in-person treatment.

XSL•FO
RenderX

Figure 1. Eight tapping points utilized in Emotional Freedom Techniques practice. The Tapping Solution App includes a point on the side of the hand. Copyright 2019 by Peta Stapleton. Reprinted with permission.



Support for the long-term efficacy of web-based EFT treatment has recently emerged in a 2-year follow-up of a web-based trial for food cravings [59]. The treatment group participants completed a self-paced web-based EFT treatment program comprising 7 modules throughout the 8-week intervention period, and a wait list group also completed the EFT web-based intervention following the end of the intervention period. From preintervention to immediately postintervention and 2-year follow-up, scores significantly improved for food cravings (−28.2%), power of food (−26.7%), restraint (+13.4%), depression (−12.3%), anxiety (−23.3%), and somatic symptoms (−10.6%). Significant improvements were also seen in carbohydrates and fast food cravings between 6 months and 2 years. Findings suggest that ongoing treatment for cravings for desirable food was not required following the 8-week web-based EFT intervention.

XSL•FO
RenderX

The physiological mechanisms of action underlying EFT have been elucidated in several studies. A triple blind RCT compared a single hour-long session of EFT with both talk therapy and rest [50]. Measures included psychological symptoms of anxiety and depression and biological assessment of cortisol, the stress hormone. The study found that psychological distress dropped by more than twice as much in the EFT group as it did in the other two groups. Cortisol levels declined significantly more in the EFT group. Another study examined both cortisol and immunoglobulin levels in participants receiving EFT over the course of a weeklong workshop [60]. A reduction was found in baseline cortisol of 37% and increased synthesis of immunoglobulins by 113% as well as improvements in a range of other physiological markers of general health. A study of pregnant women also found significant decreases in cortisol and increases in immunoglobulins after EFT treatment [61]. In addition, an RCT of veterans with clinical levels of PTSD found a significant reduction of 53% in symptoms such as flashbacks, nightmares, and hypervigilance. Participants received 10 EFT sessions, and their gene expression was measured before and after treatment. Significant upregulation was found in 6 genes related primarily to immunity and suppression of inflammation [62]. A similar study found regulated expression of 2 microRNAs associated with depression [63].

A criticism of EFT is that because it borrows elements from established therapies such as exposure and cognitive therapy, its acupressure component may be no more than placebo. A total of 6 dismantling or component studies rigorously tested this hypothesis and all found that acupoint tapping did indeed enhance treatment results. A review of these studies reported the same effect [64].

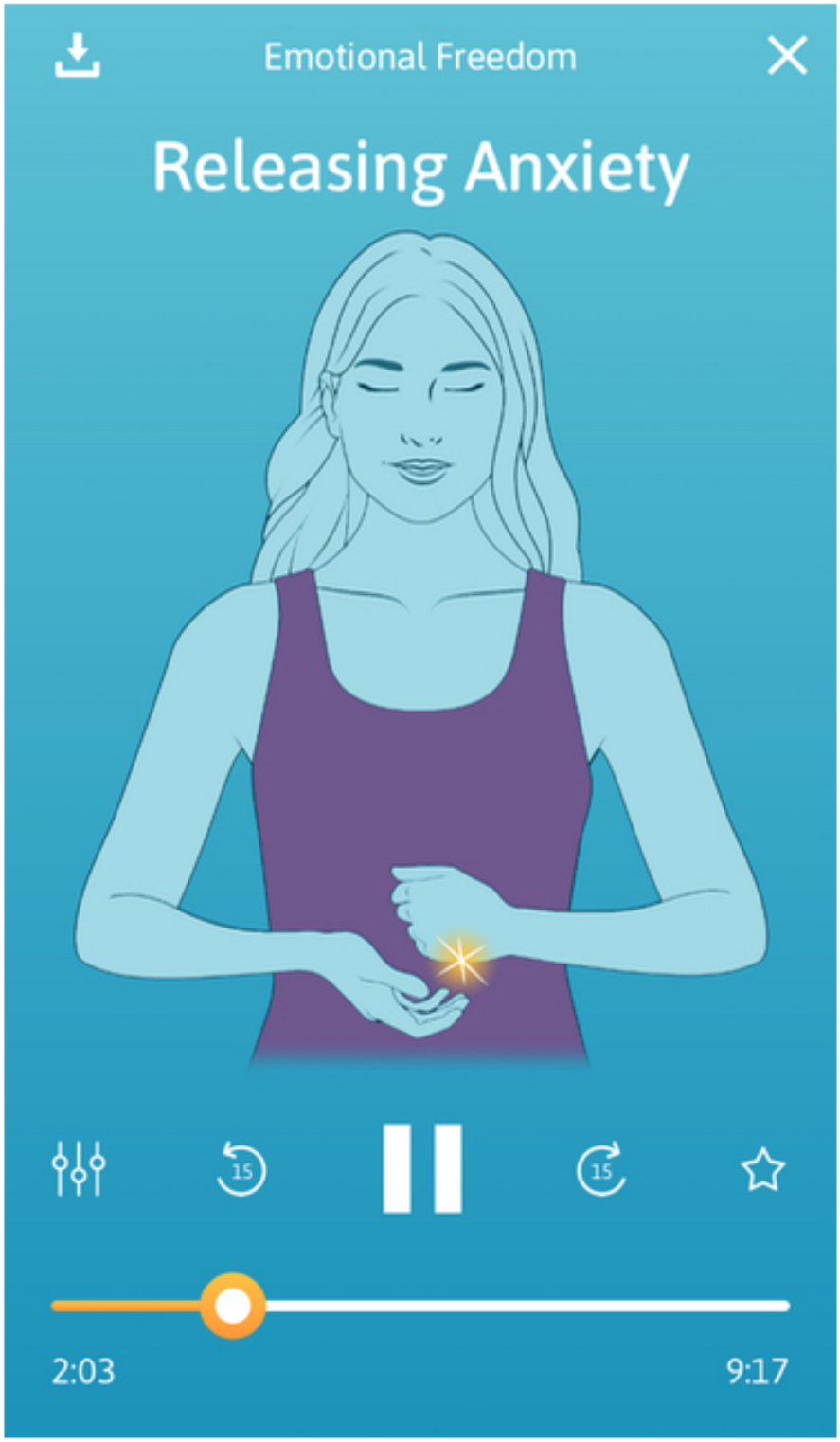
This Study

The Tapping Solution App recorded one million user sessions 12 months following its launch. The Tapping Solution App offers a suite of more than 220 guided tapping meditations, with category titles such as “Emotional Freedom,” “Fears and Phobias,” and “Sleep Support.” Each audio track (≤10 min in duration) guides users through a tapping sequence targeting a particular problem. The content was designed and recorded by a practitioner certified in Clinical EFT by EFT Universe, one of the largest EFT training organizations in the world. The scientific advisory committee for the app included researchers who had collectively published over 40 clinical trials, meta-analyses, and systematic reviews of EFT. The app is free to download from iOS (Apple) and Android (Google Play) platforms, with a paid version available that contains additional content and a lifetime subscription to updates. App features include a personalized dashboard with motivational quotes and a progress tracker, SUD self-assessments at the beginning and end of each tapping meditation, and a download option for offline activity. A visual acupoint graphic (Figure 2) is also presented in each session that highlights the acupoints on the face and body (as previously displayed in Figure 1).

This study sought to evaluate the impact of The Tapping Solution App on intensity self-ratings of anxiety and stress in a large sample of app users. Since studies have found large initial gains from Clinical EFT interventions as symptoms drop rapidly [31], it was hypothesized that a significant reduction in app users’ emotional intensity ratings would be found across app meditations.

XSL•FO
RenderX

Figure 2. Acupoints are highlighted on 9 acupoints on the face and body. The side of the hand is highlighted in this image.



Methods

Participants and Procedure

This study is a cross-sectional analysis of app user self-ratings. The sample comprised 270,461 app users aged between 18 and 65 years and above who had self-selected and downloaded The Tapping Solution App for use on a mobile device across 12 months from October 2018. The sample consisted of 81.9% (221,508/270,461) women and 18.1% (48,953/270,461) men. Demographics by age and sex are shown in Table 1. Approximately half the app users (135,771/270,461, 50.20%) were located in the United States, followed by the United Kingdom (36,512/270,461, 13.50%), Canada (34,078/270,461, 12.60%), and Australia (27,857/270,461, 10.30%). In terms of device use, 85.50% (324,929/380,034) of app sessions were completed on a smartphone and 14.50% (55,105/380,034) on a tablet device. Participants provided informed consent when

XSL•FO
RenderX

agreeing to a statement in the terms and conditions of The Tapping Solution App, which stated that their anonymized data would be used for research purposes. This study was not approved by an institutional review board or ethics committee because the data set that was analyzed consisted of existing third-party data that were deidentified [65].

Table 1. Demographics of app users by age and gender (N=270,461).

Demographics	Women, n (%)	Men, n (%)
Age (years)		
18-24	9737 (3.60)	3786 (1.40)
25-34	28,669 (10.60)	7843 (2.90)
35-44	45,708 (16.90)	9737 (3.60)
45-54	55,174 (20.40)	10,548 (3.90)
55-64	49,765 (18.40)	8655 (3.20)
≥65	32,726 (12.10)	8384 (3.10)
Total	221,508 (81.90)	48,953 (18.10)

Before each session, app users were advised to consult a doctor regarding any issue relating to a psychological or physiological symptom that required medical attention. Furthermore, the terms and conditions of app use stated that the content provided in the app did not substitute for advice, diagnosis, or treatment from a qualified health care professional. Participants received weekly email updates from one of the app developers who encouraged the use of various app session categories.

Anxiety App Sessions

The word search function was used within The Tapping Solution App to identify app meditations related to anxiety. In line with criteria from the anxiety app meta-review by Sucala et al [21], app meditation titles containing the words “anxiety,” “worry,” “panic attack,” “social anxiety,” and “fear” (ie, symptoms of generalized anxiety disorder) were used in the analyses. A total of 12 app meditations were identified, which included “Releasing Anxiety,” “Turn Your Day Around—Tapping for Anxiety, Tap and Breathe,” and “Releasing Anxiety in the Breath.”

Stress App Sessions

The word search function within the app was used to identify app meditations designed to target psychological stress symptoms. A total of 11 app meditation titles that referenced *stress* were identified, including “I’m Stressed About the World,” “Nervous Tension & Stress Release,” and “Releasing Evening Stress.”

Emotional Intensity Indicator

The SUD scale [32] provided a measure of emotional intensity. Psychological symptoms of anxiety and stress were self-rated by app users on a scale of 0 to 10 (0=*no distress at all* to 10=*worst distress imaginable*) before and after app sessions. This rating was represented on the app using a built-in Visual Analog Scale [66], in which users slid a dot along a visual scale to indicate their symptom intensity rating. Users provided 2 SUD scores: one at the start of the session and another on completion of the tapping meditation. The average time between pre– and post–app session ratings was approximately 10 min. Wolpe [67] developed the SUD scale for use with World War II veterans to measure the emotional impact of traumatic events. Increased SUD scores are associated with heightened arousal of the sympathetic nervous system [68]. SUD ratings are correlated with heart rate, respiratory rate, and galvanic skin response [69]. When interventions lower SUD levels, physiological signs of stress are reversed [70].

The App Intervention

An example of the app interface during a session is shown in Figure 2. Upon opening the app, users were provided with a short topic summary overview. For example, the *Releasing Anxiety* session description states the session purpose:

> *Anxiety is not just felt in our minds but with our whole body, which is why using a technique like tapping that uses the mind and body is so powerful. Anxiety often appears when we are worried about the future and feel disconnected from the present moment. Use this tapping meditation to begin to rewire your brain to release anxiety and stress and allow things to be easy.*

Session progress (time display in minutes and seconds) was visible for app users throughout each session (Figure 2).

Results

A total of 23 meditations available on The Tapping Solution App between October 2018 and October 2019 were identified. App session intensity reports for the 23 meditations, comprising 380,034 completed session plays, were uploaded from Google Analytics and imported to SPSS version 26 for analyses. The completed plays for 12 anxiety meditations ranged from 1025 for “Releasing Anxiety in the Mind” to 174,433 for “Releasing Anxiety.” Completed plays for 11 stress app meditations ranged from 2306 for “I’m Stressed About My Weight” to 10,659 for “Nervous Tension & Stress Release.” Figures 3 and 4 display the 12 anxiety meditations and 11 stress meditations by volume of completed plays and change in the net intensity rating.

XSL•FO
RenderX

Figure 3. Anxiety app meditation by title and total completed plays (n=316,323).

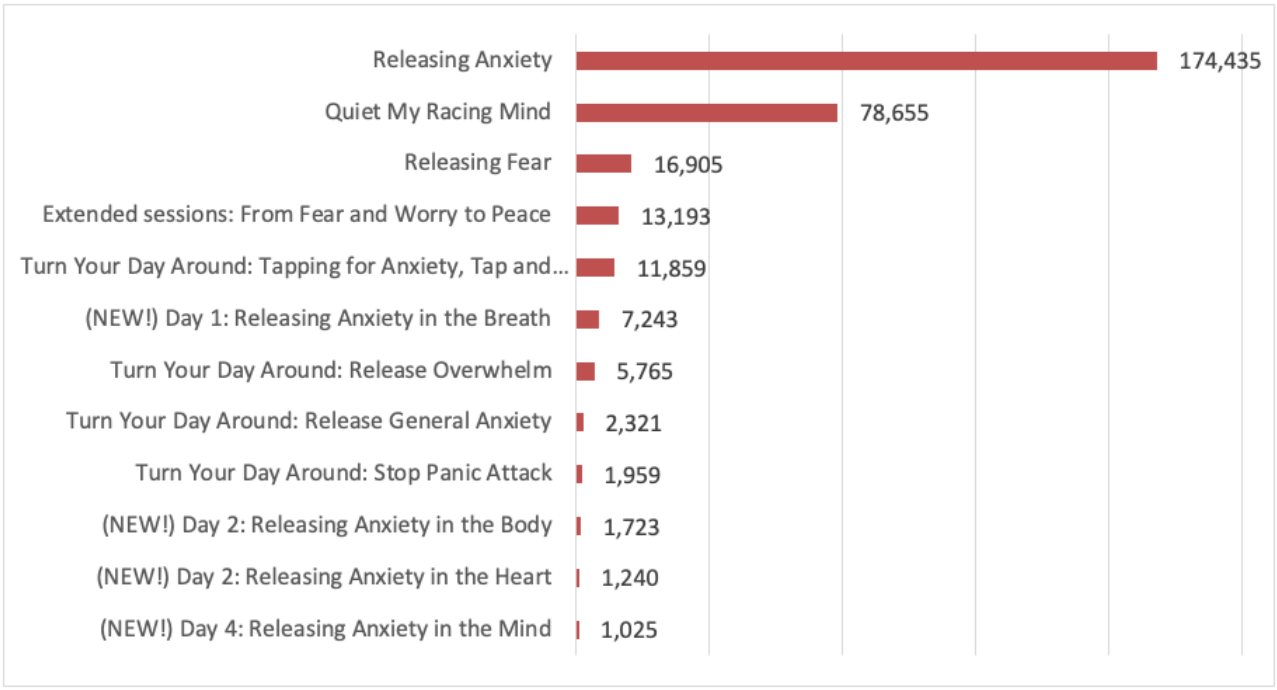
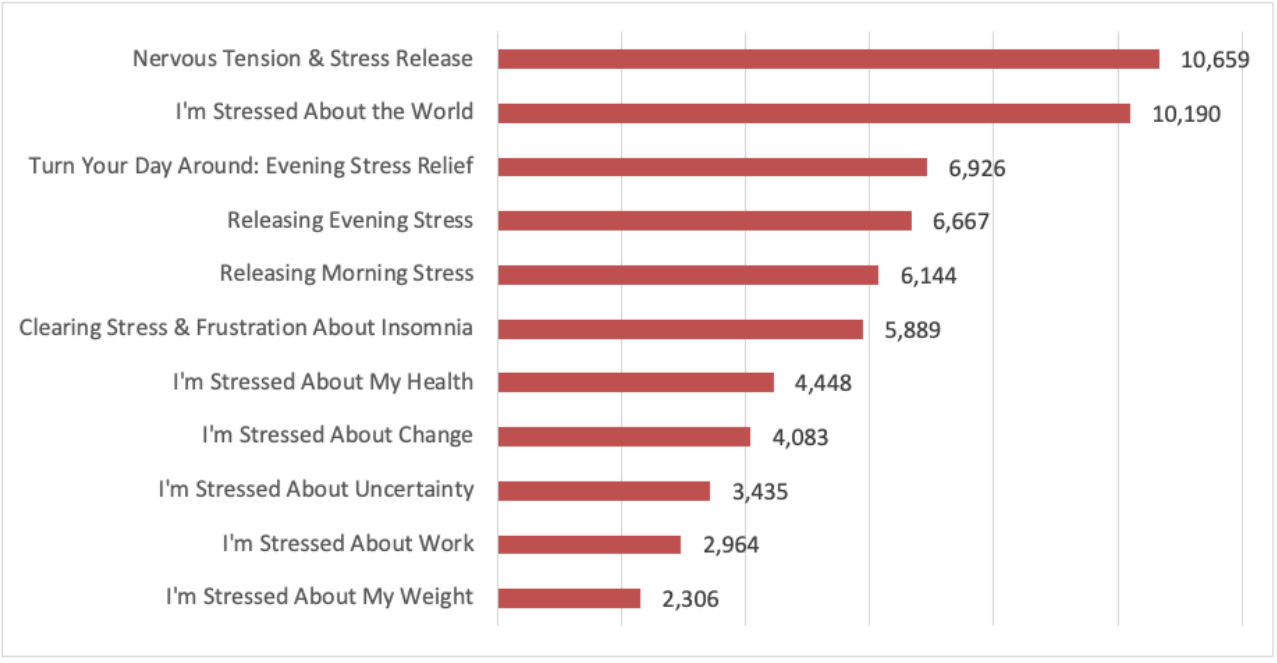


Figure 4. Stress app meditation by title and total completed plays (n=63,711).



Weighted means and SDs for user self-ratings across the anxiety and stress app meditations based on completed plays are displayed in Table 2. Across the 12 anxiety meditations, the difference in emotional intensity ratings from presession (mean 6.66, SD 0.25) to postsession (mean 3.75, SD 0.30) was statistically significant, $t_{316,322}$=−6009.64, P<.001, two-tailed, and this was a very large effect size (d=3.71). On average, emotional intensity ratings improved by −2.91 (95% CI −2.92 to −2.91) following app use. Across the 11 stress meditations, the difference in emotional intensity ratings for presession (mean 6.91, SD 0.48) and postsession (mean 3.83, SD 0.54) was statistically significant, $t_{63,710}$=−4455.81, P<.001, two-tailed, and a very large effect size (d=6.02). On average, emotional intensity ratings improved by −3.08 (95% CI −3.08 to −3.07) following app use.

XSL•FO
RenderX

Table 2. User self-ratings results for anxiety and stress app meditations based on completed app session plays (N=380,034).

App meditation	Total plays, n	Mean (SD)	Minimum score	Maximum score	Change (%)	*P* value
Anxiety (n=12)	316,323					
Presession		6.66 (0.25)	5.73	7.48	N/A[a]	N/A
Postsession		3.75 (0.30)	3.10	4.54	N/A	N/A
Net intensity change		2.91 (0.27)	2.47	3.35	−29.13	<.001
Stress (n=11)	63,711					
Presession		6.91 (0.48)	6.19	7.80	N/A	N/A
Postsession		3.83 (0.54)	2.94	4.60	N/A	N/A
Net intensity change		3.08 (0.17)	2.87	3.37	−30.80	<.001

[a]N/A: not applicable.

Discussion

Principal Findings

Given the staggering volume of mHealth apps available for download on smartphones or mobile devices, research examining the effectiveness of intervention-based mHealth apps is critical. This study aimed to provide a preliminary review of the impact of The Tapping Solution App on psychological distress ratings in a sample of 270,461 app users. Changes in emotional intensity ratings were assessed across 23 anxiety and stress-based app meditations using data from 380,034 completed app plays collated over 12 months. As hypothesized, a significant reduction in app users' emotional intensity ratings was found across app meditations. Presession to postsession results indicated that emotional intensity ratings dropped an average of 29.13% (P<.001) for the anxiety meditations and 30.80% (P<.001) for the stress meditations. The current results offer preliminary evidence to support the immediate and large effect of The Tapping Solution App in improving ratings of psychological distress in app users.

The results of this study are consistent with a large body of work that has found EFT to be efficacious in the reduction of symptoms of anxiety and psychological distress [35,47,71]. In the RCT of Church et al [50], statistically significant improvements in subjective reports of anxiety (−58.34%) were found following a brief 50-min EFT session. In this study, mean emotional intensity ratings improved between 29.1% and 30.8% following brief app-delivered tapping sessions. Although levels of psychological distress were measured using self-reported user ratings in this study, the results suggest evidence of statistically significant differences between presession and postsession for self-ratings of psychological distress following app use. The results are in line with electroencephalogram studies of EFT that have measured extensive changes in the activation of neural networks after treatment. These include the suppression of the brain-wave frequencies of anxiety and distress and expansion of those associated with healing and *flow* states [72-74]. Our findings also support previous research that has reported large initial gains from Clinical EFT intervention as symptoms of psychological distress drop rapidly and within highly compressed time frames [31].

Studies that assess brief single-session EFT interventions are more relevant to the study of an app than EFT delivered as traditional psychotherapy. Traditional ongoing psychotherapy has positive effects that may be attributed in part to therapy duration and other factors (eg, the supportive environment, face-to-face demand characteristics) [75]. However, brief single-session administration of EFT closely matches the short app session duration and the single-session use evident in the current user sample. In this study, 270,461 app users completed 380,034 app sessions across a 12-month period. This equates to an average of 1.4 completed plays per app user, which poses a significant question regarding app session repeat usage. It may be that the regular email update provided by the app developer led users' attention to alternative app sessions within their growing app session repertoire. However, this user aspect was outside the scope of the present preliminary review. Notwithstanding, the results of this study provide strong evidence that emotional intensity ratings immediately improved following a single time, or at least brief, EFT meditation app session of 10 min (or less) in duration.

Other potential issues should be considered in the evaluation of mHealth apps. Individuals who choose app intervention in place of professional health care may find their symptoms of anxiety and/or psychological stress intensify [76]. As a result, some app users may not seek additional therapeutic support, especially if they consider the app treatment to be ineffective [77]. mHealth apps as a therapeutic resource can also present challenges for treatment adherence. In the case of meditation-style techniques, the self-administration aspect of therapy may present difficulties for individuals with minimal meditation experience [12]. However, in the case of The Tapping Solution App, the verbal and visual guidance provided during each session adheres to the principles of Clinical EFT, which can assist even novice meditators. It is therefore recommended that mHealth app development be viewed as an adjunct to professional psychological services. Furthermore, although there is little evidence to suggest negative effects of meditation-based techniques [78], some studies have identified antisocial behavior, reduced emotional stability following meditation, and depersonalization following meditation therapy [79-81]. Emotional responses of fear, dread, and terror have also been reported following personal meditation practice [75]. Although adverse emotional responses to meditation-based apps

XSL•FO
RenderX

are unlikely, this research reinforces the importance of high-level examinations of mHealth app efficacy.

Methodological Issues

As with many web-based surveys, the current large convenience sample comprising app users was determined by self-selection rather than probability sampling, which can lead to biased estimates [82]. Current findings, therefore, remain specific to the self-selected users of The Tapping Solution App. Accordingly, it is important to note that participants may have presented higher levels of motivation than the general population and had previous meditation experience, which could have influenced the observed improvements. In addition, limited app user demographic variables were measured in this study. Future assessment of a range of demographic characteristics, such as socioeconomic factors and previous meditation experience, will help to delineate mHealth app user samples. In line with the recommendations by Firth et al [24], there is a need for research to examine specific populations (eg, anxiety disorders) to help delineate which user groups benefit the most from app-delivered interventions.

Furthermore, the assessment of psychological distress was based on self-reported emotional intensity ratings rather than empirical or clinician-rated psychological measures. Without diagnostic pretreatment and posttreatment assessments, it was not possible to determine the proportion of current users with clinical levels of anxiety disorders. This in turn has limited the generalizability of the findings to a nonclinical population. Finally, the current research did not control for user expectancy effects, the nonspecific effect of any treatment [50], or other potential treatment effects such as environment and frequency of app use. It is important to note that although some investigators were certified and proponents of the EFT method, the statistician and other investigators were not.

Future Directions

Since global levels of psychological stress are on the rise [1,3], an efficacious and convenient source of unlimited anxiety and stress management resources is needed. The burgeoning field of mHealth offers a dynamic platform for mental health management opportunities. mHealth apps can help facilitate the use of consumer personal data for academic research purposes. This is particularly important because novel data donation is largely supported by individuals when data collection is for research purposes that can benefit individual health [83].

Although the current results suggest that The Tapping Solution App reduced app users' self-rated emotional intensity relating to anxiety and stress, these findings are preliminary. Further examination of the app as an intervention tool using controls is required, including feasibility, efficacy, and longitudinal research data on app efficacy. In particular, future empirical assessment could align with the proposed frameworks to help investigate technology in health care. For example, Mohr et al [20] proposed the Continuous Evaluation of Evolving Behavioral Intervention Technologies framework as a timely and efficient alternative to RCTs. Their statistical evaluation of app efficacy can be implemented throughout clinical testing and can accommodate changing app versions [21].

As with any mental health intervention, mHealth apps may be more effective for some individuals. It is possible that the mental health characteristics of those seeking web-based treatment may differ substantially from those seeking in-person treatment [58]. Therefore, studies that focus on defining patient groups that benefit most from smartphone apps (ie, treatment for issues such as elevated anxiety) will contribute much needed insight into the field of mHealth app development. Future research could investigate the reasons for user attrition within app interventions. Such insight would help inform more engaging and effective app design in the future [24]. From a clinical perspective, the American Psychiatric Association has an App Evaluation Model available to members [84]. This resource provides a hierarchical rating system and rubric designed to assist the evaluation of mental health apps and guide app recommendations for clinical patients.

The assessment of physiological arousal is another promising area for mHealth technology. Smartphone technology offers the potential to combine stress reduction app interventions with biofeedback in mHealth psychological care [13,85]. Affective states can be assessed together with physiological measures, such as heart rate variability (HRV) and cortisol levels. Such technological assessment would help to define the benefits of EFT-based app intervention and extend previous research that has identified the effect of EFT on measures of heart rate and cortisol [50,60].

Future research could also explore the functional aspects of mHealth app design that increase user engagement and therefore app efficacy in treating psychological distress [24]. Several features of the mHealth platform are thought to increase engagement with therapeutic protocols, including the provision of visual aids and interactive rating functions [85,86]. The Tapping Solution App provided both functions, together with strong auditory features. In each app session, a female or male voice provided guided instruction to the sounds of waterfalls and rhythmic background music. Studies in meditation research have hypothesized that the therapeutic effect is a result of feeling relaxed, which decreases physiological arousal [87,88]. Therefore, efficacy studies examining app sensory features may help guide future app development with benefits for app user engagement.

Apps provide researchers with an opportunity to gather data quickly from large populations. This has the potential to shorten the *translational gap* between discovery and the availability of effective therapies to patients, estimated by several studies to an average of 17 years [89]. Apps also increase the pool of available subjects exponentially; most efficacy studies of psychotherapy rely on trials with fewer than 30 participants per group [46].

Finally, the role of mHealth apps to support clinician-administered evaluations and validated assessments (eg, Beck Anxiety Inventory [90]) is another important future area of study. Although apps could be recommended as part of an overall treatment plan, it is important to recognize that it may be unsafe for patients with severe psychopathology to use apps outside of the clinical setting because of the risk of unsupervised abreactions.

XSL•FO
RenderX

Acknowledgments

The authors acknowledge Nick Ortner, Joy Matthews, and Brad Rossman from The Tapping Solution App for providing data services that enabled this project. The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this paper: data analysis was made possible by donations from individuals to the National Institute for Integrative Healthcare.

Conflicts of Interest

The authors declare the following potential conflicts of interest with respect to the research, authorship, and/or publication of this paper: DC receives income from presentations and publications on the therapeutic approach described. PS receives income from presentations and publications on the therapeutic approach described. DS declares no conflicts of interest.

References

1. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DMS-IV disorders in the national comorbidity survey replication. Arch Gen Psychiatry 2005 Jun;62(6):617-627 [FREE Full text] [doi: 10.1001/archpsyc.62.6.617] [Medline: 15939839]
2. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. Dialogues Clin Neurosci 2015 Sep;17(3):327-335 [FREE Full text] [Medline: 26487813]
3. Baxter AJ, Scott KM, Vos T, Whiteford HA. Global prevalence of anxiety disorders: a systematic review and meta-regression. Psychol Med 2013 May;43(5):897-910. [doi: 10.1017/S003329171200147X] [Medline: 22781489]
4. Haller H, Cramer H, Lauche R, Gass F, Dobos GJ. The prevalence and burden of subthreshold generalized anxiety disorder: a systematic review. BMC Psychiatry 2014 May 1;14:128 [FREE Full text] [doi: 10.1186/1471-244X-14-128] [Medline: 24886240]
5. Stress in America. Paying With Our Health. American Psychological Association. URL: https://www.apa.org/news/press/releases/stress/2014/stress-report.pdf [accessed 2019-10-17]
6. Schneiderman N, Ironson G, Siegel SD. Stress and health: psychological, behavioral, and biological determinants. Annu Rev Clin Psychol 2005;1:607-628 [FREE Full text] [doi: 10.1146/annurev.clinpsy.1.102803.144141] [Medline: 17716101]
7. Cohen S, Janicki-Deverts D, Miller GE. Psychological stress and disease. J Am Med Assoc 2007 Oct 10;298(14):1685-1687. [doi: 10.1001/jama.298.14.1685] [Medline: 17925521]
8. Cohen S, Rodriquez MS. Pathways linking affective disturbances and physical disorders. Health Psychol 1995 Sep;14(5):374-380. [doi: 10.1037//0278-6133.14.5.374] [Medline: 7498107]
9. Kemeny ME, Schedlowski M. Understanding the interaction between psychosocial stress and immune-related diseases: a stepwise progression. Brain Behav Immun 2007 Nov;21(8):1009-1018. [doi: 10.1016/j.bbi.2007.07.010] [Medline: 17889502]
10. Hassard J, Teoh KR, Visockaite G, Dewe P, Cox T. The cost of work-related stress to society: a systematic review. J Occup Health Psychol 2018 Jan;23(1):1-17. [doi: 10.1037/ocp0000069] [Medline: 28358567]
11. Trautmann S, Rehm J, Wittchen H. The economic costs of mental disorders: do our societies react appropriately to the burden of mental disorders? EMBO Rep 2016 Sep;17(9):1245-1249 [FREE Full text] [doi: 10.15252/embr.201642951] [Medline: 27491723]
12. Carissoli C, Villani D, Riva G. Does a meditation protocol supported by a mobile application help people reduce stress? Suggestions from a controlled pragmatic trial. Cyberpsychol Behav Soc Netw 2015 Jan;18(1):46-53. [doi: 10.1089/cyber.2014.0062] [Medline: 25584730]
13. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3 [FREE Full text] [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
14. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
15. Christmann CA, Hoffmann A, Bleser G. Stress management apps with regard to emotion-focused coping and behavior change techniques: a content analysis. JMIR Mhealth Uhealth 2017 Feb 23;5(2):e22 [FREE Full text] [doi: 10.2196/mhealth.6471] [Medline: 28232299]
16. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 1;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
17. Clough BA, Casey LM. Technological adjuncts to enhance current psychotherapy practices: a review. Clin Psychol Rev 2011 Apr;31(3):279-292. [doi: 10.1016/j.cpr.2010.12.008] [Medline: 21382535]
18. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. Behav Modif 2015 Nov;39(6):785-804 [FREE Full text] [doi: 10.1177/0145445515595198] [Medline: 26187164]

XSL•FO
RenderX

19. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 4;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
20. Mohr DC, Cheung K, Schueller SM, Hendricks Brown C, Duan N. Continuous evaluation of evolving behavioral intervention technologies. Am J Prev Med 2013 Oct;45(4):517-523 [FREE Full text] [doi: 10.1016/j.amepre.2013.06.006] [Medline: 24050429]
21. Sucala M, Cuijpers P, Muench F, Cardo R, Soflau R, Dobrean A, et al. Anxiety: there is an app for that. A systematic review of anxiety apps. Depress Anxiety 2017 Jun;34(6):518-525. [doi: 10.1002/da.22654] [Medline: 28504859]
22. Enock PM, Hofmann SG, McNally RJ. Attention bias modification training via smartphone to reduce social anxiety: a randomized, controlled multi-session experiment. Cogn Ther Res 2014 Mar 4;38(2):200-216 [FREE Full text] [doi: 10.1007/s10608-014-9606-z]
23. Pham Q, Khatib Y, Stansfeld S, Fox S, Green T. Feasibility and efficacy of an mhealth game for managing anxiety: 'flowy' randomized controlled pilot trial and design evaluation. Games Health J 2016 Feb;5(1):50-67. [doi: 10.1089/g4h.2015.0033] [Medline: 26536488]
24. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
25. Coulon SM, Monroe CM, West DS. A systematic, multi-domain review of mobile smartphone apps for evidence-based stress management. Am J Prev Med 2016 Jul;51(1):95-105. [doi: 10.1016/j.amepre.2016.01.026] [Medline: 26993534]
26. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
27. Lewis TL, Wyatt JC. mHealth and mobile medical apps: a framework to assess risk and promote safer use. J Med Internet Res 2014 Sep 15;16(9):e210 [FREE Full text] [doi: 10.2196/jmir.3133] [Medline: 25223398]
28. Gunter RW, Whittal ML. Dissemination of cognitive-behavioral treatments for anxiety disorders: overcoming barriers and improving patient access. Clin Psychol Rev 2010 Mar;30(2):194-202. [doi: 10.1016/j.cpr.2009.11.001] [Medline: 19942331]
29. Luxton DD, Hansen RN, Stanfill K. Mobile app self-care versus in-office care for stress reduction: a cost minimization analysis. J Telemed Telecare 2014 Dec;20(8):431-435. [doi: 10.1177/1357633X14555616] [Medline: 25316037]
30. Church D. The EFT Manual. Fourth Edition. Fulton, CA: Energy Psychology Press; 2019.
31. Church D. Clinical EFT as an evidence-based practice for the treatment of psychological and physiological conditions. Psychol 2013;04(08):645-654 [FREE Full text] [doi: 10.4236/psych.2013.48092]
32. Wolpe J. Basic principles and practices of behavior therapy of neuroses. Am J Psychiatry 1969 Mar;125(9):1242-1247. [doi: 10.1176/ajp.125.9.1242] [Medline: 4886238]
33. Shapiro F. Eye movement desensitization and reprocessing (EMDR) and the anxiety disorders: clinical and research implications of an integrated psychotherapy treatment. J Anxiety Disord 1999;13(1-2):35-67. [doi: 10.1016/s0887-6185(98)00038-3] [Medline: 10225500]
34. About EFT. Peta Stapleton. 2019. URL: https://petastapleton.com/resources/the-ultimate-tapping-guide/ [accessed 2020-10-22]
35. Clond M. Emotional freedom techniques for anxiety: a systematic review with meta-analysis. J Nerv Ment Dis 2016 May;204(5):388-395. [doi: 10.1097/NMD.0000000000000483] [Medline: 26894319]
36. Gaesser AH, Karan OC. A randomized controlled comparison of emotional freedom technique and cognitive-behavioral therapy to reduce adolescent anxiety: a pilot study. J Altern Complement Med 2017 Feb;23(2):102-108. [doi: 10.1089/acm.2015.0316] [Medline: 27642676]
37. Gallo, Fred. Energy Psychology in Psychotherapy: A Comprehensive Sourcebook. New York, USA: W W Norton & Co; 2002.
38. Gallo, Fred P. Energy Psychology: Explorations at the Interface of Energy, Cognition, Behavior, and Health. Second Edition. Boca Raton, FL: CRC Press; 2004.
39. Stewart A, Boath E, Carryer A, Walton I, Hill L, Phillips D, et al. Can Matrix Reimprinting Be Effective in the Treatment of Emotional Conditions in a Public Health Setting? Results of a U.K. Pilot Study. Energy Psychology Journal 2013 May 15;5(1):13-18. [doi: 10.9769/EPJ.2013.5.1.AS]
40. Jones SJ, Thornton JA, Andrews HB. Efficacy of emotional freedom techniques (EFT) in reducing public speaking anxiety: a randomized controlled trial. Energy Psychology Journal 2011 May 1;3(1):- [FREE Full text] [doi: 10.9769/epj.2011.3.1.sjj]
41. Mollon P. Psychoanalytic Energy Psychotherapy. London, UK: Karnac; 2007.
42. Schulz KM. Integrating Energy Psychology into treatment for adult survivors of childhood sexual abuse: An exploratory clinical study from the therapist's perspective. Dissertation Abstracts International: Section B: The Sciences and Engineering 2007;68(1-B):635 [FREE Full text]
43. Stapleton P, Sheldon T, Porter B, Whitty J. A randomised clinical trial of a meridian-based intervention for food cravings with six-month follow-up. Behav Change 2012 Feb 22;28(1):1-16 [FREE Full text] [doi: 10.1375/bech.28.1.1]
44. EFT Research. 2020. URL: https://www.eftuniverse.com/research-studies/eft-research [accessed 2020-08-02]
45. Chambless D, Sanderson W, Shoham V. An update on empirically validated therapies. Clin Psychol 1996;49:5-18. [doi: 10.1037/e555332011-003]

XSL•FO
RenderX

46. Chambless DL, Hollon SD. Defining empirically supported therapies. J Consult Clin Psychol 1998;66(1):7-18. [doi: 10.1037/0022-006x.66.1.7]
47. Boath E, Stewart A, Carryer A. Tapping for success: A pilot study to explore if emotional freedom techniques (EFT) can reduce anxiety and enhance academic performance in university students. Innov Prac in Higher Ed 2013;1(3):- [FREE Full text]
48. Church D. The effect of EFT (emotional freedom techniques) on athletic performance: a randomized controlled blind trial. Open Sports Sci 2010 Jan 4;2(1):94-99 [FREE Full text] [doi: 10.2174/1875399x00902010094]
49. Church D, Downs D. Sports confidence critical incident intensity after a brief application of emotional freedom techniques: a pilot study. The Sport J 2012;15(1).
50. Church D, Yount G, Brooks AJ. The effect of emotional freedom techniques on stress biochemistry: a randomized controlled trial. J Nerv Ment Dis 2012 Oct;200(10):891-896. [doi: 10.1097/NMD.0b013e31826b9fc1] [Medline: 22986277]
51. Baker AH, Siegel LS. Energy Psychol J 2010 Nov 1;2(2):15-32 [FREE Full text] [doi: 10.9769/epj.2010.2.2.ahb.lss]
52. Salas MM, Brooks AJ, Rowe JE. The immediate effect of a brief energy psychology intervention (emotional freedom techniques) on specific phobias: a pilot study. Explore (NY) 2011;7(3):155-161. [doi: 10.1016/j.explore.2011.02.005] [Medline: 21571234]
53. Wells S, Polglase K, Andrews HB, Carrington P, Baker AH. Evaluation of a meridian-based intervention, emotional freedom techniques (EFT), for reducing specific phobias of small animals. J Clin Psychol 2003 Sep;59(9):943-966. [doi: 10.1002/jclp.10189] [Medline: 12945061]
54. Burk L. Single Session EFT (Emotional Freedom Techniques) for Stress-Related Symptoms After Motor Vehicle Accidents. Energy Psychology Journal 2010 May 01;2(1):-. [doi: 10.9769/epj.2010.2.1.lb]
55. Church D, Piña O, Reategui C, Brooks A. Single-session reduction of the intensity of traumatic memories in abused adolescents after EFT: a randomized controlled pilot study. Traumatology 2012;18(3):73-79 [FREE Full text] [doi: 10.1177/1534765611426788]
56. Hartung J, Stein P. Telephone Delivery of EFT (Emotional Freedom Techniques) Remediates PTSD Symptoms in Veterans. Energy Psychology Journal 2012 May 01;4(1):-. [doi: 10.9769/EPJ.2012.4.1.JH]
57. Brattberg G. Self-administered EFT (emotional freedom techniques) in individuals with fibromyalgia: a randomized trial. Integrat Med 2008;7(4):a [FREE Full text]
58. Church D, Clond M. Is online treatment as effective as in-person treatment?: psychological change in two relationship skills groups. J Nerv Ment Dis 2019 May;207(5):315-319. [doi: 10.1097/NMD.0000000000000975] [Medline: 30958423]
59. Stapleton P, Lilley-Hale E, Mackintosh G, Sparenburg E. Online delivery of emotional freedom techniques for food cravings and weight management: 2-year follow-up. J Altern Complement Med 2020 Feb;26(2):98-106. [doi: 10.1089/acm.2019.0309] [Medline: 31765223]
60. Bach D, Groesbeck G, Stapleton P, Sims R, Blickheuser K, Church D. Clinical EFT (emotional freedom techniques) improves multiple physiological markers of health. J Evid Based Integr Med 2019;24:2515690X18823691. [doi: 10.1177/2515690X18823691] [Medline: 30777453]
61. Yuniarti Y, Suwondo A, Runjati R, Maharani S, Ofiwijayanti H, Anastasia S. Influence of spiritual therapy emotional freedom technique (seft) on cortisol levels and immunoglobulin e (study of anxiety in pregnant women in independent practice midwives semarang). In: Asian Academic Society International Conference Proceeding Series. 2016 Presented at: ASEAN; May 12-13, 2016; Lubna Salsabila URL: http://aasic.org/proc/aasic/article/view/193
62. Church D, Yount G, Rachlin K, Fox L, Nelms J. Epigenetic Effects of PTSD Remediation in Veterans Using Clinical Emotional Freedom Techniques: A Randomized Controlled Pilot Study. Am J Health Promot 2018 Jan;32(1):112-122. [doi: 10.1177/0890117116661154] [Medline: 27520015]
63. Yount G, Church D, Rachlin K, Blickheuser K, Cardonna I. Do noncoding RNAs mediate the efficacy of energy psychology? Glob Adv Health Med 2019;8:2164956119832500. [doi: 10.1177/2164956119832500] [Medline: 30828482]
64. Church D, Feinstein D. The Manual Stimulation of Acupuncture Points in the Treatment of Post-traumatic Stress Disorder: a Review of Clinical Emotional Freedom Techniques. EFT Universe. 2019. URL: https://tinyurl.com/y3ae72gk
65. Research Involving the Secondary Use of Existing Data. Committee for Protection of Human Subjects. URL: https://cphs.berkeley.edu/engagement.pdf [accessed 2019-10-17]
66. Lesage F, Berjot S, Deschamps F. Clinical stress assessment using a visual analogue scale. Occup Med (Lond) 2012 Dec;62(8):600-605. [doi: 10.1093/occmed/kqs140] [Medline: 22965867]
67. Wolpe J, Brady JP, Serber M, Agras WS, Liberman RP. The current status of systematic densitization. Am J Psychiatry 1973 Sep;130(9):961-965. [doi: 10.1176/ajp.130.9.961] [Medline: 4580442]
68. Thyer BA, Papsdorf JD, Davis R, Vallecorsa S. Autonomic correlates of the subjective anxiety scale. J Behav Ther Exp Psychiatry 1984 Mar;15(1):3-7. [doi: 10.1016/0005-7916(84)90115-0] [Medline: 6470155]
69. Scheeringa MS, Zeanah CH, Myers L, Putnam F. Heart period and variability findings in preschool children with posttraumatic stress symptoms. Biol Psychiatry 2004 Apr 1;55(7):685-691. [doi: 10.1016/j.biopsych.2004.01.006] [Medline: 15065300]
70. Wilson DL, Silver SM, Covi WG, Foster S. Eye movement desensitization and reprocessing: effectiveness and autonomic correlates. J Behav Ther Exp Psychiatry 1996 Sep;27(3):219-229. [doi: 10.1016/s0005-7916(96)00026-2] [Medline: 8959423]

XSL•FO
RenderX

71. Boath, A Stewart, A Carryer. Tapping for PEAS emotional freedom technique (EFT) in reducing presentation expression anxiety syndrome (PEAS) in university students. Innovative Practice in Higher Education 2012;1(2) [FREE Full text]
72. Lambrou P, Pratt G, Chevalier G. Physiological and psychological effects of a mind/body therapy on claustrophobia. Subtle Energies and Energ Med 2003;14:239-251 [FREE Full text]
73. König N, Steber, Seebacher, von Prittwitz, Bliem, Rossi. How therapeutic tapping can alter neural correlates of emotional prosody processing in anxiety. Brain Sci 2019 Aug 19;9(8):- [FREE Full text] [doi: 10.3390/brainsci9080206] [Medline: 31430984]
74. Pennington J, Sabot D, Church D. EcoMeditation and emotional freedom techniques (EFT) produce elevated brain-wave patterns and states of consciousness. Energ Psychol: Theor, Res Treat 2019;11(1):13-40.
75. Vieten C, Wahbeh H, Cahn BR, MacLean K, Estrada M, Mills P, et al. Future directions in meditation research: recommendations for expanding the field of contemplative science. PLoS One 2018;13(11):e0205740 [FREE Full text] [doi: 10.1371/journal.pone.0205740] [Medline: 30403693]
76. Cuijpers P, Schuurmans J. Self-help interventions for anxiety disorders: an overview. Curr Psychiatry Rep 2007 Aug;9(4):284-290. [doi: 10.1007/s11920-007-0034-6] [Medline: 17880859]
77. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
78. Goyal M, Singh S, Sibinga EM, Gould NF, Rowland-Seymour A, Sharma R, et al. Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. JAMA Intern Med 2014 Mar;174(3):357-368 [FREE Full text] [doi: 10.1001/jamainternmed.2013.13018] [Medline: 24395196]
79. Castillo RJ. Depersonalization and meditation. Psychiatry 1990 May;53(2):158-168. [doi: 10.1080/00332747.1990.11024497] [Medline: 2191357]
80. Otis L. Adverse effects of transcendental meditation. In: Shapiro D, Walsh R, editors. Meditation: Classic and Contemporary Perspectives. New York, USA: Aldine; 1984:201-208.
81. Shapiro J. Adverse effects of meditation: a preliminary investigation of long-term meditators. Intern J of Psychosomat 1992;39:1-4 [FREE Full text]
82. Bethlehem J. Selection bias in web surveys. Intern Stat Rev 2010;78(2):161-188. [doi: 10.1111/j.1751-5823.2010.00112.x]
83. Skatova A, Goulding J. Psychology of personal data donation. PLoS One 2019;14(11):e0224240 [FREE Full text] [doi: 10.1371/journal.pone.0224240] [Medline: 31747408]
84. App Evaluation Model. American Psychiatric Association. 2019. URL: https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model [accessed 2019-10-17]
85. Dillon A, Kelly M, Robertson IH, Robertson DA. Smartphone applications utilizing biofeedback can aid stress reduction. Front Psychol 2016;7:832 [FREE Full text] [doi: 10.3389/fpsyg.2016.00832] [Medline: 27378963]
86. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. JMIR Mhealth Uhealth 2015 May 13;3(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]
87. Grossman P, Niemann L, Schmidt S, Walach H. Mindfulness-based stress reduction and health benefits. A meta-analysis. J Psychosom Res 2004 Jul;57(1):35-43. [doi: 10.1016/S0022-3999(03)00573-7] [Medline: 15256293]
88. Zeidan F, Johnson SK, Gordon NS, Goolkasian P. Effects of brief and sham mindfulness meditation on mood and cardiovascular variables. J Altern Complement Med 2010 Aug;16(8):867-873. [doi: 10.1089/acm.2009.0321] [Medline: 20666590]
89. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med 2011 Dec;104(12):510-520. [doi: 10.1258/jrsm.2011.110180] [Medline: 22179294]
90. Beck A, Epstein N, Brown G, Steer R. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56(6):893-897. [doi: 10.1037/0022-006X.56.6.893]

Abbreviations

EFT: Emotional Freedom Techniques
mHealth: mobile health
PTSD: posttraumatic stress disorder
RCT: randomized controlled trial
SUD: Subjective Units of Distress

XSL•FO
RenderX

Edited by G Eysenbach; submitted 03.03.20; peer-reviewed by E Leskowitz, J Manning; comments to author 04.05.20; revised version received 15.05.20; accepted 04.06.20; published 14.10.20.

Please cite as:
Church D, Stapleton P, Sabot D
App-Based Delivery of Clinical Emotional Freedom Techniques: Cross-Sectional Study of App User Self-Ratings
JMIR Mhealth Uhealth 2020;8(10):e18545
URL: http://mhealth.jmir.org/2020/10/e18545/
doi:10.2196/18545
PMID:32862128

©Dawson Church, Peta Stapleton, Debbie Sabot. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 14.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

African American Emerging Adult Perspectives on Unintended Pregnancy and Meeting Their Needs With Mobile Technology: Mixed Methods Qualitative Study

Lucy Annang Ingram[1*], PhD; Crystal Stafford[2*], DrPH; Quentin McCollum[3*], MPH; McKenzie Isreal[1*], BA

[1]Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[2]Department of Veteran Affairs, Durham VA Medical Center, Durham, NC, United States

[3]College of Social Work, University of South Carolina, Columbia, SC, United States

[*]all authors contributed equally

Corresponding Author:
Lucy Annang Ingram, PhD
Department of Health Promotion, Education, and Behavior
Arnold School of Public Health
University of South Carolina
915 Greene Street
Room 503G
Columbia, SC, 29202
United States
Phone: 1 803 777 4389
Email: lannang@sc.edu

Abstract

Background: In the United States, a disproportionate number of unintended pregnancies occur among African Americans, particularly those in their later teenage years and early 20s. Mobile technology is becoming more ubiquitous as a method for health promotion; however, relatively little research has been done with this population to determine their perspectives about unintended pregnancy, the potential of successfully using mobile technology to prevent unintended pregnancy, and the content of such programs.

Objective: The purpose of this study was to obtain the perspectives of African American emerging adults about unintended pregnancy and the use of mobile technology to reduce unintended pregnancy rates.

Methods: Focus groups and interviews were conducted with 83 African Americans, aged 18-21 years. Data were analyzed using an open coding process. Emergent codes were then added as needed, and themes and subthemes were identified.

Results: Participants cited the social environment and lack of education as primary reasons for disproportionate rates of unintended pregnancy. They noted that unintended pregnancy is an important issue and that they desire more sexual health information. They enthusiastically supported mobile technology as a means to communicate unintended pregnancy prevention programming and offered many suggestions for program content, look, and feel.

Conclusions: Young and emerging adult African Americans want and need sexual health resources, and a mobile-based platform could be widely accepted and address needs to lower disproportionate rates of unintended pregnancy. An essential next step is to use these findings to inform the development of a mobile-based unintended pregnancy prevention and sexual health program prototype to determine feasibility.

(JMIR Mhealth Uhealth 2020;8(10):e21454) doi:10.2196/21454

KEYWORDS

unintended pregnancy; emerging adult; African American; mobile technology; pregnancy; teenage years; health promotion; mobile; sexual health

XSL•FO
RenderX

Introduction

In the United States, unintended pregnancy is common. In 2011, the rate of unintended pregnancies was 45% [1,2]. Most notably, unintended pregnancy rates are highest among older teens and women in their early 20s [3,4]. Among these, a disproportionate number occurs among racial and ethnic minorities. Of all unintended pregnancies in the United States in 2011, 38% were among Whites, 50% were among Hispanics, and 64% were among African Americans [1]. Although unintended pregnancy rates have been on a steady decline nationally, the disparity among racial and ethnic minorities warrants further attention.

Unintended pregnancy, particularly among youth, can have health consequences for the mother and child, educational consequences such as reduced probability of obtaining a college degree, and economic consequences such as reduced lifetime earnings and taxpayer burden [5]. Other social consequences, such as shame and disconnection from friends and family, are prevalent among this population [6]. The reasons for these disproportionately high rates are varied and include structural inequalities, negative attitudes about contraception, limited partner communication, and inadequate knowledge about the effectiveness of different contraceptive options or about how to use specific methods [2,7-13].

When examining rates by region of the United States, racial disparities in unintended pregnancy rates are even more pronounced in the South. In particular, young adult African American women have some of the highest rates of unintended pregnancy in South Carolina [14]. South Carolina's birth rate is 21.7 births per 1000 among women aged 15-19 years [14]. The 2013 birthrate for 18–19-year-old African American women in South Carolina was 70.4 per 1000, whereas the rate among their white counterparts was 52.3 per 1000 [15]. These data reveal the significant need to address unintended pregnancy in South Carolina and among African American young adults in particular.

Young African Americans are just as likely to use the internet as young whites, and nearly all African Americans between the ages of 18 and 29 years have been found to use at least one social media app [16]. With the majority of the population having access to a smartphone, the use of mobile technology can play a significant role in promoting sexual health and ultimately reducing unintended pregnancies. In 2016, Mangone et al [17] published their research with analysis of over 6800 mobile apps related to pregnancy prevention. Overall, the investigators found that, while there were some innovative, creative, and educational apps in the market, very few used evidence-based practices, had information about modern contraception methods, and were tailored to a particular race or ethnicity. Moving forward, if mobile technology is to be a method of addressing unintended pregnancy, the products should be enhanced so that the most accurate and effective information is made available to the audiences that could greatly benefit from access to unintended pregnancy prevention and sexual health promotion information. To address this gap in the literature, we conducted a qualitative study of African Americans, aged 18-21 years, to obtain their perspectives about using mobile technology to prevent unintended pregnancy among young adults.

Methods

Study Participants

This study used focus group and interview methodology to determine perspectives on using mobile technology to reduce unintended pregnancy among young adult African Americans. Study participants consisted of African American young adults, aged 18-21 years. Additionally, efforts were made to recruit both parenting (those who were pregnant or had a child) and nonparenting young adults (those who were not pregnant and never had a child). Recruitment flyers were developed that incorporated images representative of the intended audience (eg, African American young adults, pregnant African American women with a male partner, teachers, health care professionals), a description of the study purpose, eligibility criteria, and staff contact information including a land line phone number, text-enabled phone number, and email address. Purposive sampling was employed whereby participants were recruited using a variety of methods, maximizing the locations where potential study participants were likely to be found. For example, flyers were posted in high-traffic areas for African American young adults and circulated among networks such as college professors and college student organizations (at both predominantly white institutions and historically Black colleges and universities), health educators, community groups, churches, public libraries, family planning clinics, social media websites (Facebook and twitter), and word of mouth. Community events where intended audience members congregate, such as sports tournaments and a local step show, were also venues for recruitment. Potential participants were screened for eligibility and asked their availability to attend a 90-minute focus group. We also employed the strategy of tentatively scheduling focus group meetings as we found that several potential participants were more amenable to having a scheduled "appointment" rather than having to be contacted later to determine their ability to attend a group on some future date. Additionally, spontaneous interviews were conducted in the event that the recruiter and eligible participant had availability at the time of the initial meeting. Eligibility criteria for the participants were: African American race, age 18-21 years, ability to attend a 90-minute focus group meeting, and English as the primary language. Potential participants were sent study reminders (via phone conversation, voice mail, email, or text message) at least 1 week, 3 days, and the day prior to, as well as the day of, the scheduled focus groups to help facilitate their attendance. Participants were enrolled January 2016 through June 2016. All study procedures were reviewed and approved by the University of South Carolina institutional review board.

Instrumentation

A brief, self-administered, demographic and sexual health questionnaire was developed by the lead researcher to obtain a profile of the study participants. Demographic measures included age, sex, race, ethnicity, marital status, educational level, employment status, and sexual orientation. Sexual health information included a range of questions to help describe

XSL•FO
RenderX

participants' sexual background as well as sexual health history. For example, questions were included about age of sexual debut, whether or not participants had children, as well as when or if they ever received formal sexual health instruction, the content of that instruction, and their perception of where most African American youth receive sexual health and teen pregnancy prevention information.

A focus group guide was developed to standardize the conduct of focus group sessions. Open-ended questions were developed by the first author based on the research aims of the study. The questions were guided by a review of extant literature in unintended pregnancy prevention, sexual health promotion, and mobile technology. Based on this, the focus group questions were organized into 4 categories: importance of teen pregnancy prevention, technology as a method to promote teen pregnancy prevention, content of a technology-based program to promote teen pregnancy prevention, and design and format of a technology-based program to promote teen pregnancy prevention. The lead author is a public health researcher and professor who has had over 20 years' experience in the development, implementation, and analysis of focus groups and qualitative instruments. Focus group questions were pretested with members of the intended audience to ensure content validity. Based on the feedback received during the pretest, the focus group questions were refined and finalized.

Procedures

Participants were invited to attend a focus group to share their thoughts about how young adult African Americans use mobile technology for sexual health information. The researcher and a research assistant obtained training about policies for the protection of human subjects in research, and at least one of the staff members was available at each focus group meeting to obtain informed consent. Additionally, the researcher trained the research assistant in best practices for conducting and moderating successful focus groups including information about note taking, encouraging participation, and navigating roadblocks. Participants were asked to review a copy of the informed consent document as a staff member reviewed it with them. Participants were encouraged to ask questions during and after the document review, were then asked to sign the consent form indicating their approval, and were given a copy of the consent document for their records. Once participants were enrolled into the study, they were asked to complete an anonymous self-administered paper-and-pencil survey. In order to limit feelings of anxiety regarding responses to sensitive questions, after the staff member provided participants their surveys, participants completed them, were asked not to attach their name to the survey, and were then asked to place the completed survey in a sealed envelope that would not be opened until the end of the recruitment day. All participants received a US $15 cash incentive for their participation. Focus groups were scheduled for days and times that were convenient to study participants and ranged from mid-day to late evening, any day of the week. In the event that the staff member and eligible participant had availability at the time of the initial meeting, spontaneous interviews were conducted. Meetings or interviews were conducted by a member of the research team who had been trained to moderate focus group discussions and conduct interviews.

Data Analysis

Focus group meetings were digitally audio-recorded. Audio recordings were reviewed and analyzed using the memoing technique [18], which focused on identifying repeated responses that emerged from participants as well as those that emerged and represented new ideas and notions. Memos allowed for coding and categorizing data as well as allowing the researchers to scrutinize the data to explore hypotheses, relationships between concepts, and explanations that emerged from the data. A draft of the codebook to categorize participant responses was established using an open coding process [19] using the discussion guide as a framework. Emergent codes were then added as needed. The data were triangulated using transcripts, field notes, and memos and analyzed by two independent researchers using inductive, semantic thematic analysis. Any discrepancies were resolved by consensus. Survey data were analyzed using SPSS version 23, and descriptive statistics were performed.

Results

Demographic Characteristics and Sexual Health Information

A total of 83 young adults participated in the study, and 13 focus groups and 7 interviews were conducted. Participants ranged in age from 18 to 21 years (mean age 19.5 years) and were primarily African American (82/83, 99%) and of non-Hispanic ethnicity (79/83, 95%). One participant identified as multiracial. Of the participants, 98% (81/83) reported never being married, and 52% (43/83) had some high school education or had earned at least a high school diploma. Of the participants, 93% (77/83) identified as heterosexual, with 7% (6/83) identifying as gay, lesbian, or bisexual.

Regarding sexual behaviors, 83% (69/83) of participants had ever had consensual sex, with 16 years as the average age of first sexual experience. Among those who were sexually experienced, 15% (11/83) had ever been or ever gotten someone pregnant. As it relates to formal sex education, the majority of participants (80/83, 96%) had received some form of sex education before age 18, with 75% (62/83) having their first instruction before entering high school. As part of their sex education, participants identified having learned about abstinence until marriage, birth control, condoms, and sexually transmitted infections (STIs), with the overwhelming majority having received information about condoms (76/83, 95%) and STIs (76/83, 95%). When asked about where they think most African American youth receive information about teen pregnancy prevention, over half of the sample identified family members (47/83, 57%) and school (45/83, 54%) as the most popular sources. The next most popular sources of information were friends and social media at 45% (37/83) and 41% (34/83) of participants, respectively. The sources identified least often were music (8/83, 10%), flyers/pamphlets (7/83, 8%), movies (5/83, 6%), newspapers/magazines (3/83, 4%), and church (1/83, 1%). When asked about whether they think that African

XSL•FO
RenderX

American youth receive enough accurate information about teen pregnancy prevention, most participants (60/83, 73%) responded "no" or "not sure" (see Table 1).

XSL•FO
RenderX

Table 1. Demographic and sexual health characteristics (N=83).

Characteristics	Results
Age (years), mean (range)	19.5 (18-21)
Sex, n (%)	
Male	34 (41)
Female	49 (59)
Race, n (%)	
African American	82 (99)
Multiracial	1 (1)
Ethnicity, n (%)	
Non-Hispanic	79 (99)
Hispanic	1 (1)
Marital status, n (%)	
Single/never married	81 (98)
Married	0 (0)
Divorced	0 (0)
Other	2 (2)
Highest level of education, n (%)	
Some high school	8 (10)
High school diploma	35 (42)
Some college	37 (45)
Associate's degree	1 (1)
Bachelor's degree	2 (2)
Some graduate school	0 (0)
Graduate degree	0 (0)
Sexual orientation, n (%)	
Heterosexual	77 (93)
Gay or lesbian	5 (6)
Bisexual	1 (1)
Sexual initiation, n (%)	
Yes	69 (83)
No	14 (17)
Age of sexual initiation (years), mean (range)	16 (12-18)
Ever been pregnant/gotten someone pregnant, n (%)	
Yes	11 (15)
No	60 (83)
Not sure	1 (7)
Formal sex education, n (%)	
Yes	80 (96)
No	3 (4)
Grade of earliest sex education, n (%)	
Elementary school (1st-5th grades)	19 (23)
Middle school (6th-8th grades)	43 (52)
High school (9th-12th grades)	17 (21)

XSL•FO
RenderX

Characteristics	Results
College	1 (1)
Content of sex education, n (%)	
Abstinence until marriage	58 (73)
Birth control	61 (77)
Condoms	76 (95)
Sexually transmitted infections	76 (95)
Common sources of African American youth teen pregnancy prevention information, n (%)	
Family member	47 (57)
School	45 (54)
Friends	37 (45)
Social media	34 (41)
Health care provider	33 (40)
Television	20 (24)
Websites	13 (19)
Music	8 (10)
Flyer/pamphlet	7 (8)
Movies	5 (6)
Newspaper/magazine	3 (4)
Other (church)	1 (1)
African American youth receive enough accurate teen pregnancy prevention information, n (%)	
Yes	22 (27)
No	45 (55)
Not sure	15 (18)

Perspectives on Teen Pregnancy and Prevention

Disproportionately High Rates of Pregnancy Among African American Older Adolescents are Due to Social Environment and Lack of Education

Participants did not seem to be surprised that, while rates of unintended pregnancy among youth and young adults have declined as a whole, we continue to see rates high among African Americans, particularly those aged 18-19 years. A variety of reasons for this statistic were proffered, with many of them centered around 2 notions: the social environment and lack of education.

As one participant stated, “I would think their social environment [is the reason why the rates are high among African Americans]. It depends who they are influenced by either at home or out in the community or even just what they see out in the community. They just think that it’s fine and it’s cool. They never talk about the right way to have sex…they only see about the wrong ways.” Another participant explicitly identified messages in music as a culprit. He stated that the messages in music promote “that type of behavior,” presumably unprotected sex. Other participants described the environs that are changing for that age group in that many may be going from high school to the “freedom” of college and the feeling that they are adults and can make decisions such as having sex without concern for parents’ judgment or oversight. As one participant noted, “18-19 is that transitional period, and when you come to college, you can do whatever you want to do.”

In addition to one’s environment as a causal factor, several participants suggested that the reason for the disproportionate rates was lack of education. As one participant stated, “Nobody is teaching about it, and there is little guidance. What little you do learn from school, you don’t always take it to heart, and teachers don’t answer all the questions, and you’re not comfortable talking to parents, so…” As stated by another participant, “There is definitely a lack of education about sex and preventing pregnancy,” so she was not surprised by the statistic.

Teen Pregnancy is an Important Issue

Overwhelmingly, participants noted that teen pregnancy is an important issue. Discussions about why it is important centered around the idea that young people should try to preserve their youth and with a child comes responsibilities that are vastly different and adult-like. According to one parenting participant, “Because now that I have a child, it is not easy, I wouldn't want anyone else who is not ready or prepared…I know I wasn’t…I had to rush and get prepared…and like in high school, you got school work, and if you have a job and all that…it’s a lot to

XSL•FO
RenderX

deal with." Another participant stated, "They haven't had time to develop yet, and they're still learning about themselves." Other comments included, "you haven't started your life so how can you start another life?" "you're still a child," teens should "wait until they are situated," and "babies shouldn't raise babies."

What I Didn't Learn About Teen Pregnancy Prevention

There was a variety of responses to the question about what participants did not learn about pregnancy prevention when they were in school, and in fact, responses more broadly reflected what they did not learn about sexual health in general. As it relates to pregnancy prevention, one parenting female participant noted that she did not learn about how hard it would be, and as she stated, "what all I had to go through." Others noted that they did not learn about how to put on a condom, abstinence, birth control, and the withdrawal method. More contextual responses were about practical things and less about clinical outcomes or biological processes. For example, one participant stated, "I wish that when I learned about pregnancy prevention and sex ed and stuff, they woulda told me that masturbation is okay. I mostly do it now to keep myself out the bull*$#!" indicating the need to learn about alternatives to seek pleasure. Another participant stated, "What they didn't say was while you're learning, your mindset will be totally different [than] when you're in that position, like when you're in that moment, your mind is not there". Responses that are applicable to what was not learned about sexual health in general included the consequences of having multiple sex partners, the high level of peer pressure in high school to engage in sexual activities, STIs, and sexual orientation (other than heterosexuality).

Preferred Format of a Pregnancy Prevention Program

Participants identified several formats for a pregnancy prevention program that reflected their preferred type of interaction and platform of interaction. Some participants preferred individual interaction while others preferred a group setting for interaction. Those preferring an individual format described the benefits of confidentiality and privacy. For example, one participant stated, "I would like an individual format so you can be more comfortable and have no judgment from others." For those who preferred a group format, reasons for this included modeling behavior and ability to glean from others' experience. One participant noted, "In a group, if your friends think that preventing pregnancy is cool, then you're gonna think it is too." Another participant stated, "If you have the group discussion, then it's like, if you have the right people in there, it's like the influence of it is just going to change their mind…" Yet another participant commented that they would prefer a group so that you "could hear other peoples' stories." In instances where the initial responses of a group were divergent, at the end of the discussion, each perspective could see the benefits of the other (even if they still preferred their own).

Regarding the use of mobile technology as a format, the majority of participants liked the idea of having a mobile platform to provide pregnancy prevention information. Reasons for this included "everyone is on their cell phone," "youth spend a lot of time on their social media," "I'm mostly on my phone…but not on a computer," and an app "would be good so you can express yourself privately…and I'm on my phone all the time." When further prompted about specific aspects of a mobile program, participants noted that it should be brief, interactive, have little text, be entertaining, and allow for access across multiple sittings (rather than needing to be completed at one point in time).

A Mobile Technology Program for Pregnancy Prevention

Program Content — Topics

When asked about the topics that should be included in a mobile technology program for pregnancy prevention among African American youth, participants offered a wide variety of responses. Similar to responses about what they didn't learn about pregnancy prevention, participants tended to respond about interest in information about sexual health content in general. Frequently suggested topics included condom use and birth control. Additional interests included information about relationship dynamics whereby participants mentioned consent, modeling ways to say "no," how to communicate with partners about birth control, and how to handle abuse. Several participants indicated wanting to know about sexual orientation and LGBTQ issues with several groups openly debating about whether pregnancy prevention would be applicable to sexual minorities. When prompted, participants also noted that substance use would be a topic of interest, particularly as it relates to its influence on decision making.

Program Content — Messages

Study participants had very insightful ideas about program messages. Some suggested that engaging in safer sex, having planned birth control, and waiting (until a certain age, until you're responsible, until you have a connection, or until marriage) should be clear messages. Others took a more inspired route providing suggestions such as "It goes down in the D.M." (a reference to the social media site Instagram's direct message feature…as well as a popular rap lyric), "Open books, not legs," "Be yourself," "A real man knows that she's worth the wait," and "Use the strap if you don't want the clap." Conversely, participants also offered suggestions for what not to include such as abstinence messages, information about the pull-out method, and the message, "don't have sex."

Gender Dynamics

When asked about whether men and women should receive the same program content, responses were mixed. Some participants stated that the same information should be given to both men and women. One participant's rationale was that, in fact, men should have information that is pertinent to women so that men can know how to "help her out or make her feel better." According to another participant, "the consequences are the same for both genders so, yeah, they should get the same information." Participants who suggested that men and women receive different information highlighted interesting perspectives on gender dynamics and gender roles. For instance, one male participant stated, men should know "never have unprotected sex even if she says she's on the pill," and another male participant stated that, "men should know how to think with the

XSL•FO
RenderX

head on their shoulders." According to one female participant, men should know "your body count doesn't make you a man." Another suggestion from a female participant was, "if you're serious about a partner, get tested together."

Parenting Dynamics

Participants provided different messages about whether parenting and nonparenting youth should receive the same program content. For those who thought parenting and nonparenting persons should receive the same information, many reflected on the fact that even though someone may have had a child already, they still may want to and need to know information about how to prevent a subsequent unintended pregnancy. As one participant stated, "I feel like, well the people who do already have kids and everything, it be some stuff they don't know that they didn't get or receive so they should all get the same information." For those whose perspective was that different information should be provided, they pointed out that parenting teens could serve as models and tell others "what it's about and how it is."

Program Design

Participants provided their ideas about the optimal design for a mobile technology program for pregnancy prevention for African American older adolescents. Participants provided feedback about colors ("it should bold," "there should be bright colors"), images ("there should be real pictures, but they shouldn't be too graphic" and "the people should be attractive"), music (if any music, soft music is preferred so as not to be distracting), and videos (preferably tutorials such as how to put on a condom). Some suggested alternatively making the look of the program an interactive feature whereby users could tailor it to their own style.

Participants also described different features that could enhance participation or interest. Suggestions included offering some type of incentive to download the program, connecting it to social media (to not only garner more followers and participants but also connect to modes of communication that they already use), integrating endorsements (these could be from celebrities, adult mentors, or even friends [or relatable others] who have used the program and liked it), and making it fun and attractive.

Discussion

Principal Findings

Based on national trends, unintended pregnancy is on the decline [1], which could suggest that the topic is of waning to little interest to young and emerging adults. However, we found the contrary to be true. Nearly 90 young people were recruited to participate in this study to provide their perspectives about unintended pregnancy. Our findings reveal that they are indeed concerned with the trends, particularly those highlighting disparities among African Americans, and have many ideas for how to promote unintended pregnancy prevention and overall sexual health in their communities.

Commonly, respondents found it difficult to divorce unintended pregnancy prevention from overall sexual health. Many suggestions that respondents provided about topics and messages of interest reflected conventional prevention content such as promoting condom and contraception use [20]. Additional suggestions highlighted contextual dynamics of engaging in sexual behavior such as relationship dynamics, emotional challenges involved with young adulthood and sexual relationships, and decision making (eg, whether or not to have sex, whether or not to use contraception). These findings reflect other studies wherein researchers have elicited perspectives of youth, young adult, or influential stakeholders about adolescent pregnancy prevention program content [21,22] that may be precursors to any actual behavior or complementary factors that co-occur with sexual behavior. It was also clear that, when asked about their own sexual health education, participants noted some clear deficiencies. The move to adopting comprehensive sex education programs is controversial in some states and communities [23-25], with South Carolina as no exception [26]. Participants' lack of exposure to consistent, medically accurate, sexual health information could be reflective of this context. These findings support developing programs that include comprehensive messages to promote sexual health and improve on sexual health knowledge among young and emerging adults.

The idea of using mobile technology to promote unintended pregnancy prevention and sexual health was widely supported. None of the participants suggested that such a platform would not work and in fact enthusiastically provided suggestions ranging from key features, interface look and feel, and compatibility with existing social media platforms. Other researchers have reported about the vast opportunity that mobile apps have for reaching youth and young adults with sexual health education content [17,27-32]; however, there remains a gap in the success of these programs among racial and ethnic minority populations that suffer disproportionately from poor health consequences. More work must be done to determine the best content, prevention strategies, and interactive functionality that these youth will be most receptive to on mobile platforms.

A significant finding that we want to highlight is the fact that study participants were not surprised by the fact that African American young adults outpace all others in unintended pregnancy rates and that this racial health disparity continues to persist. Either because this is a part of their reality or because they are accustomed to experiencing health outcomes different from majority populations, the researchers found this slightly concerning. Some respondents commented that there does not need to be different (read: tailored) programs because "we all experience these public health challenges"; however, a majority of respondents noted that it would be great to have programs and systems that they felt were "for them." Future research could do well to explore this notion further.

Limitations and Strengths

Our study is subject to several limitations. One limitation is the sampling technique used for the study. We used purposive sampling to obtain our study sample. In some regards, while this approach was successful in garnering eligible study participants, it could introduce bias and limit generalizability of study findings. Another limitation is the terminology used to discuss the focus of the project. Many of the study documents included the term "teen pregnancy"; however, due to the age

XSL•FO
RenderX

range of the study population exceeding the teen years we also used the term "unintended pregnancy." The vast majority of teen births are among women aged 18-19 years, and rates of unplanned pregnancy are highest among those in their late teens and early 20s [3,4]. While we are confident that the interchangeable use of the terms did not likely negatively impact our study findings, the discrepancy is noteworthy.

Strengths of the study should also be noted. Our ability to recruit a community sample of more than 80 18–21-year-old African American participants to conduct qualitative assessments is significant. This population is frequently characterized as "vulnerable" and "difficult-to-reach," which often translates to an excuse for a researcher's inability to obtain their active participation. Our recruitment efforts were successful because they were targeted to the spaces and places where the intended audience was likely to be found. Additionally, about the sample, 40% was comprised of men. Similar to the aforementioned characterization of African Americans, discussions with men about unintended pregnancy are limited. Most often, the topic is left the responsibility of women, with little consideration for the perspectives of male partners. More opportunities to elicit the perspectives of male partners are needed.

Conclusions

A qualitative approach proved to be useful for assessing young people's perspectives on unintended pregnancy as well as their thoughts about mobile technology as a viable platform for health promotion. While it was clear that many participants had received some form of sexual health education, many still expressed that young and emerging adults want and need additional sexual health resources and that a platform, such as one that is mobile-based, could be widely accepted, address needs to lower disproportionate rates of unintended pregnancy, and address barriers related to access to health information. An essential next step to build upon this study is to use these findings to inform the development of a mobile-based unintended pregnancy prevention and sexual health program prototype. Based on our findings, we offer the following recommendations to developers, researchers, and program planners working in this area:

1. Integrate sexual health content (highlighting unintended pregnancy) into a mobile app.
2. African American young adults recognize unintended pregnancy as an important issue. Develop content and messages that reflect this.
3. Allow for participants to navigate the app on their own as well as allow group interactive features.
4. Provide real-life scenarios such that it is clear that decision making can be challenging when one is in a compromised or heat-of-the-moment situation.
5. Provide authentic reasons why unintended pregnancy at a young age can be difficult. Highlight stories from teen parents who can share their experiences.
6. Provide information on a variety of topics including condoms, contraception, relationship dynamics, emotional dynamics of young adulthood, sexual decision making, and sexual orientation. Our participants did not support including messages about abstinence or less effective contraceptive methods.
7. Ensure that the design of the interface is modern and reflective of images and language that resonate with the demographic.
8. Ensure compatibility with social media.

Given our success of obtaining young and emerging adult perspectives for the current study, it would be advisable to include young people in the prototype development process as well. Once a prototype is developed, pilot testing for uptake and acceptability as well as outcomes such as improvements in knowledge and behavior would be appropriate before larger scale-up.

Conflicts of Interest

None declared.

References

1. Finer LB, Zolna MR. Declines in Unintended Pregnancy in the United States, 2008–2011. N Engl J Med 2016 Mar 03;374(9):843-852. [doi: 10.1056/nejmsa1506575]
2. Unintended pregnancy in the United States. Guttmacher Institute. 2019 Jan. URL: https://www.guttmacher.org/sites/default/files/factsheet/fb-unintended-pregnancy-us_0.pdf [accessed 2020-06-15]
3. Martin JA, Hamilton BE, Osterman MJK, Driscoll AK, Drake P. Births: Final Data for 2017. Natl Vital Stat Rep 2018 Nov;67(8):1-50. [Medline: 30707672]
4. Kost K, Maddow-Zimet I, Arpaia A. Pregnancies, Births and Abortions Among Adolescents and Young Women in the United States, 2013: National and State Trends by Age, Race and Ethnicity. Guttmacher Institute. 2017. URL: https://www.guttmacher.org/report/us-adolescent-pregnancy-trends-2013 [accessed 2020-10-05]
5. Bradburn EM, Carroll CD. Short-term enrollment in postsecondary education: Student background and institutional differences in reasons for early departure, 1996-98. National Center for Education Statistics. 2002 Nov. URL: https://nces.ed.gov/pubs2003/2003153.pdf [accessed 2020-10-05]
6. Smith W, Turan JM, White K, Stringer KL, Helova A, Simpson T, et al. Social Norms and Stigma Regarding Unintended Pregnancy and Pregnancy Decisions: A Qualitative Study of Young Women in Alabama. Perspect Sex Reprod Health 2016 Jun 11;48(2):73-81 [FREE Full text] [doi: 10.1363/48e9016] [Medline: 27166869]

XSL•FO
RenderX

7. Duncan C, Miller DM, Borskey EJ, Fomby B, Dawson P, Davis L. Barriers to safer sex practices among African American college students. J Natl Med Assoc 2002 Nov;94(11):944-951. [Medline: 12442997]
8. Annang L, Johnson SD, Pepper-Washington MA. The context of sexual risk among African-American female college students. Journal of Health Disparities Research and Practice 2009;3(1):99-114 [FREE Full text]
9. Borrero S, Farkas A, Dehlendorf C, Rocca CH. Racial and ethnic differences in men's knowledge and attitudes about contraception. Contraception 2013 Oct;88(4):532-538 [FREE Full text] [doi: 10.1016/j.contraception.2013.04.002] [Medline: 23697702]
10. Hodgson EJ, Collier C, Hayes L, Curry LA, Fraenkel L. Family planning and contraceptive decision-making by economically disadvantaged, African-American women. Contraception 2013 Aug;88(2):289-296 [FREE Full text] [doi: 10.1016/j.contraception.2012.10.011] [Medline: 23177266]
11. Rocca CH, Harper CC. Do racial and ethnic differences in contraceptive attitudes and knowledge explain disparities in method use? Perspect Sex Reprod Health 2012 Sep 08;44(3):150-158. [doi: 10.1363/4415012] [Medline: 22958659]
12. Carter MW, Bergdall AR, Henry-Moss D, Hatfield-Timajchy K, Hock-Long L. A qualitative study of contraceptive understanding among young adults. Contraception 2012 Nov;86(5):543-550. [doi: 10.1016/j.contraception.2012.02.017] [Medline: 22464411]
13. Jones J, Mosher W, Daniels K. Current contraceptive use in the United States, 2006-2010, and changes in patterns of use since 1995. Natl Health Stat Report 2012 Oct 18(60):1-25. [Medline: 24988814]
14. South Carolina Campaign to Prevent Teen Pregnancy: South Carolina 15-19 Year Old Birth Trends, 2002-2012. 2013. URL: http://www.teenpregnancysc.org/sites/default/files/uploads/Documents/State%20EPI%202011%20Teen%20Birth%20Data.pdf [accessed 2016-03-30]
15. Fact Forward. Accelerating progress: A road map for achieving further reductions in teen pregnancy. South Carolina Campaign to Prevent Teen Pregnancy. 2014 Nov 1. URL: http://www.teenpregnancysc.org/path-forward-accelerating-progress [accessed 2016-03-30]
16. Smith A. African Americans and Technology Use: A Demographic Portrait. Pew Research Center. 2014. URL: http://www.pewinternet.org/2014/01/06/african-americans-and-technology-use/ [accessed 2019-12-05]
17. Mangone ER, Lebrun V, Muessig KE. Mobile Phone Apps for the Prevention of Unintended Pregnancy: A Systematic Review and Content Analysis. JMIR Mhealth Uhealth 2016 Jan 19;4(1):e6 [FREE Full text] [doi: 10.2196/mhealth.4846] [Medline: 26787311]
18. Groenewald T. Memos and Memoing. In: Given L, editor. The SAGE encyclopedia of qualitative research methods. Los Angeles, CA: Janke RA; 2008:505-506.
19. Strauss A, Corbin J. Basics of qualitative research: techniques and procedures for developing grounded theory, 3rd ed. Thousand Oaks, CA: SAGE Publications, Inc; 1998.
20. Steiner RJ, Rasberry CN, Sales JM, Gaydos LM, Pazol K, Kramer MR, et al. Do health promotion messages integrate unintended pregnancy and STI prevention? A content analysis of online information for adolescents and young adults. Contraception 2018 Apr 21;98(2):163-170. [doi: 10.1016/j.contraception.2018.04.011] [Medline: 29684327]
21. Johnson-Motoyama M, Moses M, Kann TK, Mariscal ES, Levy M, Navarro C, et al. Parent, Teacher, and School Stakeholder Perspectives on Adolescent Pregnancy Prevention Programming for Latino Youth. J Prim Prev 2016 Dec 15;37(6):513-525. [doi: 10.1007/s10935-016-0447-2] [Medline: 27628931]
22. Woodhams E, Sipsma H, Hill BJ, Gilliam M. Perceived responsibility for pregnancy and sexually transmitted infection prevention among young African American men: An exploratory focus group study. Sex Reprod Healthc 2018 Jun;16:86-91. [doi: 10.1016/j.srhc.2018.02.002] [Medline: 29804783]
23. Ringwalt C, Vincus AA, Hanley S, Ennett ST, Bowling JM, Haws S. The prevalence of evidence-based drug use prevention curricula in U.S. middle schools in 2008. Prev Sci 2011 Mar 4;12(1):63-69 [FREE Full text] [doi: 10.1007/s11121-010-0184-3] [Medline: 20683664]
24. Demissie Z, Brener ND, McManus T, Shanklin SL, Hawkins J, Kann L. School health profiles 2014: Characteristics of health programs among secondary schools. Atlanta, GA: Centers for Disease Control and Prevention; 2015.
25. TFN Insider. Sex education in Texas public schools: Progress in the lone star state. Texas Freedom Network. 2011 Dec 11. URL: https://tfn.org/publication/sex-education-in-texas-public-schools-progress-in-the-lone-star-state/ [accessed 2020-10-05]
26. Kempner M. Sex education in South Carolina still failing 25 years after passage of comprehensive law. Rewire News. 2013 Jan 30. URL: https://rewirenewsgroup.com/article/2013/01/30/sex-education-in-south-carolina-25-years-after-state-adopts-mandate-classroom-ins/ [accessed 2020-10-05]
27. Kalke KM, Ginossar T, Shah SFA, West AJ. Sex Ed to Go: A Content Analysis of Comprehensive Sexual Education Apps. Health Educ Behav 2017 Dec 28;45(4):581-590. [doi: 10.1177/1090198117749259]
28. Bull S, Pratte K, Whitesell N, Rietmeijer C, McFarlane M. Effects of an Internet-based intervention for HIV prevention: the Youthnet trials. AIDS Behav 2009 Jun;13(3):474-487 [FREE Full text] [doi: 10.1007/s10461-008-9487-9] [Medline: 19037719]
29. Moreno MA, Vanderstoep A, Parks MR, Zimmerman FJ, Kurth A, Christakis DA. Reducing at-risk adolescents' display of risk behavior on a social networking web site: a randomized controlled pilot intervention trial. Arch Pediatr Adolesc Med 2009 Jan 05;163(1):35-41. [doi: 10.1001/archpediatrics.2008.502] [Medline: 19124701]

XSL•FO
RenderX

30. Read S, Miller L, Appleby P, Nwosu M, Reynaldo S, Lauren A, et al. Socially optimized learning in a virtual environment: Reducing risky sexual behavior among men who have sex with men. Human Communication Research 2006;32(1):1-34 [FREE Full text] [doi: 10.1111/j.1468-2958.2006.00001.x]
31. Chen E, Mangone ER. A Systematic Review of Apps using Mobile Criteria for Adolescent Pregnancy Prevention (mCAPP). JMIR Mhealth Uhealth 2016 Nov 10;4(4):e122 [FREE Full text] [doi: 10.2196/mhealth.6611] [Medline: 27833070]
32. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013 Jan 04;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]

Abbreviations

STI: sexually transmitted infection

Edited by G Eysenbach; submitted 15.06.20; peer-reviewed by T Winder, A Rung, A Nguyen; comments to author 28.09.20; revised version received 30.09.20; accepted 02.10.20; published 21.10.20.

Please cite as:
Ingram LA, Stafford C, McCollum Q, Isreal M
African American Emerging Adult Perspectives on Unintended Pregnancy and Meeting Their Needs With Mobile Technology: Mixed Methods Qualitative Study
JMIR Mhealth Uhealth 2020;8(10):e21454
URL: http://mhealth.jmir.org/2020/10/e21454/
doi:10.2196/21454
PMID:33084591

©Lucy Annang Ingram, Crystal Stafford, Quentin McCollum, McKenzie Isreal. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 21.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Creating a Smartphone App for Caregivers of Children With Atopic Dermatitis With Caregivers, Health Care Professionals, and Digital Health Experts: Participatory Co-Design

Xiaomeng Xu[1], MD, PhD; Konstadina Griva[1], PhD; Mark Koh[2], MBBS, MD; Elaine Lum[1], PhD; Woan Shin Tan[3,4], PhD; Steven Thng[5], MBBS; Josip Car[1,6], PhD, MD, DIC, MSc, FFPH, FRCP

[1]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[2]Dermatology Service, KK Women's and Children's Hospital, Singapore, Singapore

[3]Health Services and Outcomes Research Department, National Healthcare Group, Singapore, Singapore

[4]Geriatric Education and Research Institute, Singapore, Singapore

[5]Skin Research Institute of Singapore, Singapore, Singapore

[6]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Xiaomeng Xu, MD, PhD
Centre for Population Health Sciences
Lee Kong Chian School of Medicine
Nanyang Technological University
Clinical Sciences Building, 11 Mandalay Road
Singapore, 308232
Singapore
Phone: 65 81616659
Email: xxu020@e.ntu.edu.sg

Abstract

Background: Smartphone apps could support patients and caregivers in disease self-management. However, as patients' experiences and needs might not always align with clinical judgments, the eliciting and engaging of perspectives of all stakeholders in the smartphone app design process is of paramount importance.

Objective: The aims of this study are to better understand the needs of and challenges facing caregivers and health care professionals (HCPs) who care for children with atopic dermatitis (AD) and to explore the desirable features and content of a smartphone app that would support AD self-management.

Methods: This study adopted a qualitative participatory co-design methodology involving 3 focus group discussions: workshop one focused on caregivers; workshop two engaged with HCPs; and in the last workshop, caregivers and digital health experts were asked to design the wireframe prototype. The participants completed a sociodemographic questionnaire, a technology acceptance questionnaire, and a workshop evaluation form.

Results: Twelve caregivers participated in the first workshop, and 10 HCPs participated in the second workshop. Eight caregivers and 4 digital health experts attended the third workshop. Three superordinate themes that reflected caregivers' and HCPs' challenges and needs were identified: *empowerment by education*, *confusion over treatment*, and *emotional impact*. Workshop participants also raised a series of suggestions on the features and contents of the AD self-management app, which informed the last co-design workshop, and described their needs and challenges. In the last workshop, the participants developed a wireframe prototype of the app following the identified requirements and recommendations.

Conclusions: The co-design approach was found to be a successful way of engaging with the participants, as it allowed them to express their creativity and helped us to articulate the root of the clinical problems. The co-design workshop was successful in creating and generating new ideas and solutions for smartphone app development.

(JMIR Mhealth Uhealth 2020;8(10):e16898) doi:10.2196/16898

KEYWORDS

atopic dermatitis; eczema; mobile phone; telehealth

XSL•FO
RenderX

Introduction

Background

Atopic dermatitis (AD) is a chronic inflammatory skin condition that affects up to 25% of children worldwide [1,2]. It is characterized by periods of remission and relapse due to variable, often unknown, triggers [3]. Symptoms such as itchiness or soreness can strongly affect children's daily activities and cause sleep deprivation, which substantially undermine their quality of life (QoL) [4]. The burden of childhood AD extends to their caregivers as well [5,6]. Caregivers' sleep can be equally affected as parents have to manage their child's AD symptoms during the night, which can further influence their performance at home and work [7]. In addition, AD, especially childhood AD, requires skilled self-management from caregivers, including coping with a complex regime: caregivers are often required to perform extra duties, such as preparing special meals for their children, engaging in skincare regimens, and cleaning extensively to minimize outbreaks [5,8]. A child with AD can require approximately 2-3 hours of care from their caregivers daily [9]. To prevent relapses, numerous caregivers deem it necessary to cut down on leisure, defer social activities and events, and even sacrifice their jobs to take care of their children [10].

Lack of support leaves caregivers not feeling confident or even failing to manage and provide disease management support for their children's condition in between visits to the clinic. In clinical practice, it is difficult for health care professionals (HCPs) to educate their patients adequately. Doctors have insufficient time for patient education during the tightly scheduled consultations, and there are limited health educational materials provided in the hospital setting. Caregivers and patients also have difficulty in recalling the health information following a visit to the clinic [11]. The unmet management support needs encourage patients to seek additional information from the internet or friends, and although often most helpful, such apps may also not be trustworthy or easy to use. As such, caregivers must be equipped with supporting tools that are tailored to their needs and help them to properly manage patients' condition continuously [12].

Digital health tools are increasingly being leveraged to increase patient engagement and facilitate disease self-management for conditions such as diabetes, asthma, dementia, and cancer [13-17]. Smartphone app features can engage patients more in cultivating guidelines-recommended behaviors and providing evidence-based information and advice [18-21]. For example, the disease education feature can give patients insights into their disease condition and enable them to make informed clinical decisions [20,21]; the disease tracking feature can help users detect the possible allergens and help them eliminate aggravating factors and reduce recurrence [20]. It has also been demonstrated to improve clinical outcomes, and 80% of patients consider the service a viable alternative to in-person consultations [22]. Similar to digital management of AD, previous studies comparing AD digital management with traditional in-person care models have demonstrated no significant differences between them regarding clinical outcomes and QoL outcomes [23,24]. In addition, qualitative data showed that doctors found it helpful to use the website in consultations in a pilot randomized control trial [23]. However, when considering the substantial financial savings that digital health interventions in managing AD could reap, digital health could potentially be a promising substitute for in-person AD care [25].

Although many studies document the benefits of digital health tools in terms of supporting users to monitor, treat, and follow-up disease conditions, the quality of apps is a cause for concern, as in many cases, there is a lack of clinical evidence proving the effectiveness of apps. Many of them are not tailored to the user's needs; this may explain the low adoption and use of apps as part of routine care in clinical practice [26,27]. Working in partnership with stakeholders in the design process can better serve app development. By doing so, developers can more easily identify desirable features for an app as well as obstacles hindering long-term usage; together, app developers and stakeholders could improve the viability, usability, and effectiveness of apps in health care services [28-30].

Objectives

The co-design approach refers to patients and caregivers working in partnership with staff to improve the viability, usability, and effectiveness of health care services and is considered one of the best ways of improving patients' experience of services and of guiding developers into understanding the users' needs [28]. The co-design approach has been used to build self-management apps for chronic diseases such as heart failure [29] and dementia [17,29,31,32]. Co-design has been employed in AD app development before, but relevant web-based information is so sparse that previous experiments cannot easily be repeated by other researchers. Given the fact that self-management is crucial to tackling AD among children, an app involving all stakeholders with detailed procedures is needed to ensure that patients' needs and voices are sufficiently heard. The objectives of this study are to better understand the needs of and challenges confronting caregivers and HCPs who care for children with AD and to explore the desirable features and content that smartphone apps should offer in supporting AD self-management.

Methods

Participant Recruitment

Eligible participants were recruited among those of a baseline cohort in a prior study, who had provided consent to be recontacted for participation in this study [33,34]. Caregivers who met the following inclusion criteria were approached by making phone calls or sending emails: (1) adults providing unpaid care or assistance to children (age<16 years) with AD and (2) prior experience with using a smartphone app to manage health-related issues. The exclusion criteria were as follows: caregivers who did not own a smartphone or had no knowledge of mobile apps. HCPs who have experience in managing AD pediatric patients were recruited directly from the Dermatology Department at the KK Women's and Children's Hospital in Singapore. Digital health experts with experience in the field of digital health (app assessment or app development) were

XSL•FO
RenderX

recruited face to face from the Lee Kong Chian School of Medicine, Nanyang Technological University.

Study Design and Settings

Data were collected across three co-design workshops conducted in a specialist hospital in Singapore from September 2018 to March 2019 across two phases (caregivers' session was held at 7 pm and lasted for 2 hours; HCPs' session was held at Noon and lasted for 1.5 hours. The design session was held at 7 pm and lasted for 2 hours). The procedure of the co-design workshops is detailed in Textbox 1.

Textbox 1. Co-design workshops.

Phase 1

Content

- Share their problems and difficulties encountered in their daily life
- Propose their needs
- Discuss how the app can be used in their daily life and what features can be adopted by the smartphone app

Caregivers' session (session duration: 2 hours)

- Caregivers of pediatric patients with atopic dermatitis (n=14)

Health care professionals' session (session duration: 1.5 hours)

- Pediatric dermatologists (n=8)
- Pharmacists (n=2)

Phase 2

Content

- Briefing the need and proposed feature discussed in 4he previous caregivers' journey sessions
- Top needs were assigned to the participants
- Individual crazy eights exercise to sketch their prototype
- Team sketches
- Revisions
- Present the wireframe to the entire group

Design session (session duration: 2 hours)

- Caregivers (n=6)
- Digital health experts (n=4)

Workshops one and two adopted a focus group discussion format. Workshop one was a focus group discussion with caregivers: 2 researchers (KG and WS) facilitated the group discussion. Additional field notes were collated to capture relevant contextual information (XM and ZL). Semistructured questions were used to guide the focus group discussions. To better serve the research objectives, the semistructured questions used in the previous AD QoL study were designed to address both subjective feelings and objective findings regarding the most disturbing aspects of childhood AD. The topics raised during workshop one are shown in Textbox 2.

Workshop two was a focus group discussion involving HCPs; 2 researchers (JC and KG) facilitated the group discussion. Additional field notes were taken by XX to capture relevant contextual information. Semistructured questions were used to guide focus group discussions. The semistructured questions were designed based on feedback from caregivers in workshop one. They are shown in Textbox 3.

After completion of phase one, a final co-design workshop was conducted (facilitated by JC and XX). As the targeted user of the app is caregivers whose children have AD, the workshop only invited caregivers and digital health experts to design the wireframe prototype. HCPs were not invited in this phase, but we will seek their feedback after the app is developed. After the participants outlined various design requirements in workshops one and two, the facilitator asked participants to design the interface of the app themselves. The app had to address at least one of the top needs mentioned in the previous workshops. The workshop groups discussed and explored the ideas from the drawings and came to an agreement on the top features or ideas they wanted to include in one consolidated master sketch. Groups sketched a single layout that incorporates the top ideas. The teams regrouped and critiqued the wireframe prototype. After the feedback and revision session, the groups discussed and explored the ideas from the drawings and came to an agreement during the development of the consolidated master sketch.

XSL•FO
RenderX

Textbox 2. The topics raised during workshop one.

Probing questions

- Q1. What were your needs the first 6 months after the symptoms started, from 6 months to a year, or later? Did you feel well informed, especially with respect to (1) disease status, (2) treatment and management guidance, and (3) emotional support?
- Q2. Where do you get the information? Are you satisfied with your knowledge of taking care of your child's atopic dermatitis (AD)? What is the source of the information?
- Q3. Previous experience in app usage: Has anyone used a smartphone app to manage AD?
- Q4. If there is a smartphone app to help you manage your child's eczema, please list the top features that you want to have in the app.
- Q5. If there would be a smartphone app with all the information and features that you need, at what frequency do you think you will use it? What are the possible barriers that would keep you from consistent usage?

Textbox 3. The topics raised during workshop two.

Probing questions

- Q1. How do you educate your patients (their caregivers) during the clinic visit? Are you satisfied with the results?
- Q2. What are caregivers' need the first 6 months after the symptoms started, from 6 months to a year, or later? Does the education need to change according to (1) different severity, (2) duration of diseases, and (3) child age?
- Q3. Have you considered using smartphone apps to educate your patients (and their caregivers)?
- Q4. What features should an app have to address those problems?
- Q5. Would you mind providing a teleconsultation service to your patients?

Instruments

The following questionnaires were administered during the co-design workshops: Demographic questionnaire and technology acceptance questionnaire (Multimedia Appendix 1, [35]) and workshop evaluation form (Multimedia Appendix 2).

Demographic Questionnaire

Caregiver participants were asked to fill in the demographic questionnaire in workshops 1 and 3, including their date of birth, gender, residential status, highest qualification earned, and main daytime occupation and the child's date of birth, child's gender, date of the first diagnosis, and the biggest concern relating to the management of the child's AD.

Technology Acceptance Questionnaire

Caregiver participants were asked to fill in the technology acceptance questionnaire during workshops 1 and 3. The questionnaire was developed by the study team adopting the technology acceptance model [36,37]. The list of questions tested individuals' knowledge of health care app usage, acceptability, and usability, and it addressed perceived ease of use, perceived usefulness, user satisfaction, and usability [36,37].

Workshop Evaluation Form

At the end of each workshop, the participants were asked to fill in the evaluation form and provide feedback. Three questions were asked in the evaluation form: (1) a rating of the workshop, (2) the most helpful aspect of the workshop, and (3) any suggestions for improving the workshops.

Data Collection and Analysis

Workshops were audio-recorded, transcribed, and managed using NVivo software (QSR International, version 11) [38,39]. The transcript data were first broken down into concepts and reorganized to summarize key issues from each workshop. XX, KG, and JC reviewed and discussed concepts immediately after each workshop so that emerging issues could be explored in subsequent workshops. XM transcribed the audio recordings and processed the observation logs, photos, and written products to contextualize and gain more details about the discussion. Thematic analysis was chosen to analyze the transcriptions from the workshops [22]. An inductive strategy was employed, which allowed for themes to be generated from the data, which were then organized into higher-order themes. The NVivo software program was used to store and code the transcript. First cycle coding was performed on the raw data. Subsequently, researchers (XM and KG) met and compared each other's codes. An interrater reliability (IRR) test was performed between the researchers' findings (XM and KG); whenever an acceptable IRR was not achieved, the inconsistencies were jointly discussed until a consensus was reached.

Ethical Approval

All eligible participants provided informed consent before the workshops. Ethical approval was obtained from the Institutional Review Board of Nanyang Technological University before the commencement of the study (Nanyang Technological University Institutional Review Board: 2018-09-053).

Public and Patient Involvement

Patients, the public, or caregivers were not directly involved in the development of this study, but they were recruited as participants in this study.

XSL•FO
RenderX

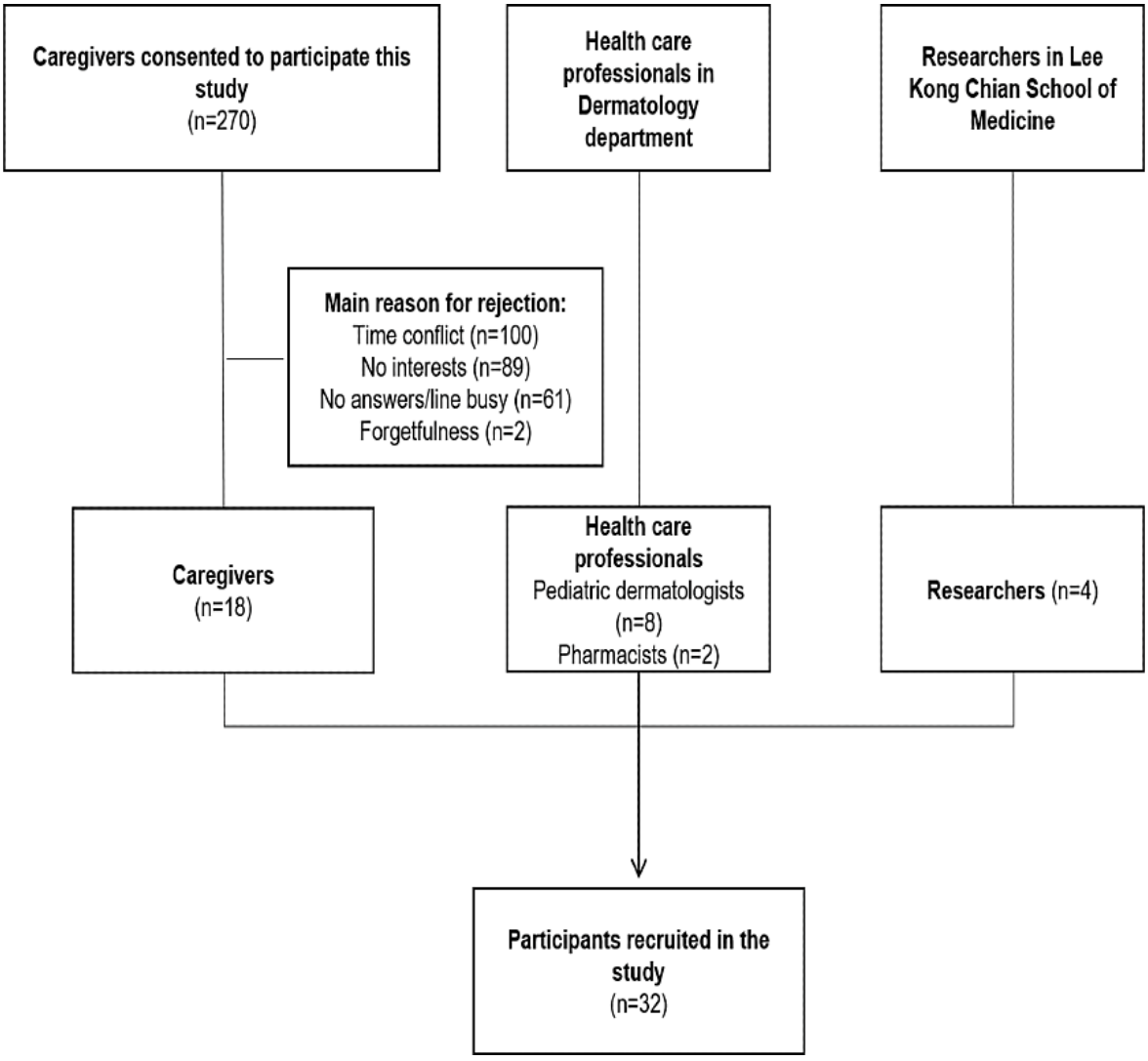
Results

Participants' Characteristics

The participant recruitment procedure and reasons for exclusion are listed in Figure 1. The characteristics of the caregivers are summarized in Table 1. Out of a total of 270 eligible caregivers, 18 caregivers participated (18/270, 6.7%). The caregivers' sample included a diverse range of demographics (Table 1): the majority of them were women (14/18, 78%), the mean age of the participants was 37.87 (SD 7.81) years, and the children's disease duration was 4.72 (SD 4.57) years (range: 1.66-10.83). Pediatric dermatologists (n=8) and pharmacists (n=2) were recruited for the second workshop: their mean age was 35.50 (SD 6.69) years and their mean practice time was 9.76 (SD 6.57) years. Four digital health experts with a mean age of 37.80 (SD 8.12) years attended the third workshop.

The caregivers' perceptions of using self-management smartphone apps were assessed to test perceived usefulness, ease of use, and user satisfaction (Table 2). A total of 89% (16/18) of the participants found it easy to get mobile tracking on children's disease, and 89% (16/18) of them felt confident in using a smartphone app to manage children's AD. A total of 78% (14/18) of the participants believed that using a smartphone app to manage AD would increase the caregiver's QoL.

Figure 1. Participants' recruitment process. Caregivers (n=2) were recruited and participated in 2 workshops (workshop one and workshop three).



XSL•FO
RenderX

Table 1. Demographic characteristics of caregiver participants.

Variables and characteristics	Value
Caregiver	
Gender, n (%)	
Female	14 (78)
Male	4 (22)
Age (years), mean (SD)	37.87 (7.81)
Qualification, n (%)	
Bachelor's degree and above	10 (56)
Diploma	4 (22)
O level and below	4 (22)
Children's disease duration (years), mean (SD)	4.72 (4.57)
Health care professional	
Gender, n (%)	
Female	8 (80)
Male	2 (20)
Age (years), mean (SD)	35.50 (6.69)
Specialization, n (%)	
Pediatrics dermatology	8 (80)
Pharmacy	2 (20)
Practice duration (years), mean (SD)	9.76 (6.57)
Researcher	
Gender, n (%)	
Female	2 (50)
Male	2 (50)
Age (years), mean (SD)	37.80 (8.12)
Specialization, n (%)	
Public health	4 (100)

XSL•FO
RenderX

Table 2. Perceived usefulness, ease of use, and user satisfaction of smartphone app among caregivers (strongly agree, agree, neutral, disagree, and strongly disagree).

Perceptions	Strongly agree, number of participants; n (%)	Agree, number of participants; n (%)	Neutral, number of participants; n (%)	Disagree, number of participants; n (%)	Strongly disagree, number of participants; n (%)
Are you aware of the availability of health apps for smartphones	12 (67)	6 (33)	0 (0)	0 (0)	0 (0)
A smartphone self-management app will enable me to get information quickly	5 (28)	10 (56)	3 (17)	0 (0)	0 (0)
A smartphone self-management app will allow my doctor to follow-up the disease from outside of the hospital	4 (24)	10 (56)	4 (24)	0 (0)	0 (0)
A smartphone self-management app will save the time of doctors and nurses	1 (5)	6 (33)	6 (33)	3 (17)	2 (11)
I would find it easy to get mobile tracking on my child's disease	4 (22)	12 (67)	2 (11)	0 (0)	0 (0)
Using a health app will be clear and understandable	4 (22)	10 (56)	2 (11)	1 (6)	1 (6)
It would be easy for me to become familiar with using a smartphone self-management app	4 (22)	10 (56)	2 (11)	1 (6)	1 (6)
I feel confident in using a smartphone app to manage my child's AD[a]	2 (11)	14 (78)	2 (11)	0 (0)	0 (0)
I believe using a smartphone app to manage AD will increase my quality of life	3 (17)	11 (61)	4 (22)	0 (0)	0 (0)
I think that I would like to use a smartphone self-management app on a daily basis	3 (17)	12 (67)	3 (17)	0 (0)	0 (0)

[a]AD: atopic dermatitis.

Contextual Themes

Overview

A total of 8 primary themes were identified that reflected caregivers' and HCPs' needs and concerns for AD self-management support: (1) understanding of causes, triggers, and symptoms of AD; (2) understanding the disease course and the complications related to AD; (3) which medication to choose and when to use it; (4) generic instruction versus specific instruction; (5) alternative treatments preferable to steroids; (6) guilt and frustration; (7) self-esteem; and (8) the duty to make their child feel better. We subsequently grouped the primary themes into 3 superordinate themes: *empowerment by education*, *confusion over treatment*, and *emotional impact* are foundational issues for support caregivers when helping them to manage a child's AD. These themes were common to all participant groups and characterized by subthemes to capture the diversity of participant perspectives. Subthemes described by HCPs focused predominantly on steroid usage and understanding of causes and triggers, whereas subthemes described by caregivers additionally touched on their feelings such as guilt, frustration, and self-esteem. These nuanced perspectives on subthemes are further explained below.

Empowerment by Education

Both the caregivers and HCPs believed that education enables patients and their caregivers to handle AD and live with AD. The caregiver's perception of diseases often shapes their attitude about self-management and affects their adherence and capability of engaging in self-management following a medical consultation. Two aspects were highlighted during the discussion: (1) understanding the causes, triggers, and symptoms of AD and (2) understanding disease course and the complications related to AD.

Understanding the Causes, Triggers, and Symptoms of AD

The caregiver participants expressed a desire to understand more about AD, especially those caregivers whose children were newly diagnosed with AD. They experienced much confusion during the first few months after the first diagnosis, especially in relation to *cause and triggers* and *symptoms* aspects. They wanted to know what environments or temperatures would prevent them from triggering flare-ups and were concerned as to whether existing symptoms were *normal* (quotes 1 and 2 in Multimedia Appendix 3).

Understanding the Disease Course and the Complications Related to AD

Caregivers also described their confusion about the *disease course* and the *complications* related to AD. They claimed that

XSL•FO
RenderX

existing support and education received during the consultation might be insufficient and sometimes may not suit their needs. According to one caregiver, "No one told them what happened and what happened next." Some caregivers did not realize the importance of disease course until they discovered that their child was following another child's disease pattern (quote 3 in Multimedia Appendix 3.). They also reported that lack of awareness and understanding of the complications related to AD could throw them into uncertainty. Therefore, relevant knowledge should be provided in the first place so as to prepare the caregivers mentally and to help them successfully and smoothly pass the most helpless stage (quote 4 in Multimedia Appendix 3).

HCPs also highlighted the need for a greater understanding of the causes and triggers for patients and their caregivers (quote 5 in Multimedia Appendix 3). However, they also underlined that the caregivers should lower their expectations and accept the truth that their children need to live with the disease. Otherwise, they could develop an irrational attitude, which could hinder them from treating AD properly, particularly if they spend most of the time trying to find a *cure* for AD (quote 6 in Multimedia Appendix 3).

Confusion Over Treatment

There was also a consensus that confusion over AD treatment could be another barrier to patients and their caregiver's proper self-management, one that could have a cascading effect on the outcomes of AD. Caregivers and patients can feel confused over (1) which medication to choose and when to use it, (2) receiving general instructions from a doctor, and (3) treatments that might be preferable to steroids.

Which Medication to Choose and When to Use It

Caregivers, especially first-time parents, can feel overwhelmed with all the information available to them when they are selecting medications. There were too many brands for them to choose in the market, and they did not know precisely which medication to choose and when to use it (quote 7 in Multimedia Appendix 3). Confusion in selecting medication was also observed among doctors. Caregivers received different prescriptions from doctors when they switched doctors (quote 8 in Multimedia Appendix 3).

Generic Instruction Versus Specific Instruction

Having to administer multiple creams also contributes to the treatment-related burden and confusion experienced by caregivers. Although the administering of lotions itself was not found to be a major issue (as caregivers have usually worked out a routine), they still had pending questions relating to the treatment timing, sequence, and amount, that is, of *applying the creams*. They reported that the doctors always give them guidance in general, but that self-management in daily life is more detailed and difficult. For example, even with the same patient, the severity of AD lesions can vary widely, and so confusion over which cream should be applied to which specific lesion can arise (quote 9 in Multimedia Appendix 3).

Alternative Treatments Preferable to Steroids

The option of using less aggressive treatments was discussed during both the caregivers' session and HCPs' session. Some parents worried about the side effects of steroids in children and mentioned that they preferred to seek alternative forms of treatment (eg, traditional Chinese medicine and ayurvedic medicine), which are deemed safer (quote 10 in Multimedia Appendix 3.). HCPs, on the other hand, expressed concerns over so-called alternative medications. They claimed that they may contain stronger steroids and that some patients may miss the best treatment opportunities either because they are avoiding or over-applying steroids (quote 11 in Multimedia Appendix 3).

Emotional Impact

Peeling and erythema of the skin affect children's appearance and can be detrimental to their social life. When the skin condition is relapsed, the children's mood becomes irascible, and their irritable emotions will induce irritation of the skin and worsen their skin condition. Three subthemes relating to the emotional impact of AD were identified during the workshops: (1) guilt and frustration, (2) self-esteem, and (3) the duty to make their child feel better.

Guilt and Frustration

The feeling of guilt was repeatedly mentioned by caregivers whenever they talked about their child's illness. It was difficult for some of the parents to express feelings of guilt, especially the subtler form of guilt, which was expressed in the form of questions (quote 12 in Multimedia Appendix 3). When the parents reflected on their situation, they sometimes regretted things they had done at an earlier stage in life (quote 13 in Multimedia Appendix 3). Guilt may also stem from applying strong medicines to young children. Some felt worried over *putting their kid through something that was not necessary* and the risk that *there is a side effect after the child grows up* (quote 14 in Multimedia Appendix 3). The problem of unhelpful *helpful* advice by others was also raised up repeatedly by the caregivers during the workshop, which could be another cause of feelings of guilt among them (quote 15 in Multimedia Appendix 3). Such advice includes doctors and peers prescribing overuse of steroids as well as unspoken blame and judgment imposed by families, friends, and HCPs.

Self-Esteem

Young teenagers are more self-conscious and care more about their body image. When the skin lesions reoccur, children might find it frustrating and sometimes show anger. This is worrying as the emotions may affect the child's social activities in school and eventually affect their confidence and self-esteem. During the workshop, parents expressed the same worries that when their children got older, AD could affect their children's self-esteem (quote 16 in Multimedia Appendix 3).

The Duty to Make Their Child Feel Better

In addition to experiencing guilt and anger, the parents also expressed more ambiguous concerns that the AD affects their children's self-esteem. They also felt a strong responsibility to make their children feel better and felt concerned that their children would become more self-aware as they grow older and

XSL•FO
RenderX

that AD would affect their emotional health in the long run (quote 17 in Multimedia Appendix 3).

Ideas and Features of the AD Self-Management App

During the workshops, various functionality requirements were discussed among participants. The most prevalent requirement across the groups was knowledge of the symptoms and on managing the symptoms, medication usage, and triggers. They also expressed a desire to brief caregivers and children about the issues that they might face in the future following diagnosis. This could potentially guide people in transitioning from raising awareness to making concrete actions and plans and developing coping strategies that could help them master everyday challenges.

The HCPs also emphasized that such information should be provided in the right amount in the local language. Excessive information provided during medical consultations may overwhelm both patients and their caregivers. However, inadequate information may also leave them feeling directionless, a feeling which induces them into "going online for relevant information which could lead doctors to spend more consultation time in correcting the wrong online [information]."

The goal is to ensure that users fully understand the requirements and thus reduce confusion over information. To accomplish this, features such as podcasts, chatbots, or web-based lectures were proposed as the potential modes for delivering such information. HCPs also highlighted that the apps should provide other patients' stories or feedback to debunk the myths concerning or against steroids.

The other requirement discussed during the workshop was the need for patients and caregivers to communicate with HCPs. In this context, participants proposed that technology could be used throughout the whole process of the health care journey, from preparation for the clinical consultation, facilitation of clinical consultation, disease status follow-up, to real-time management after medical consultation. Here, the main goal of the technology would be to help caregivers identify and manage flare-ups; to mitigate confusion; and, at the same time, facilitate HCPs to gain a better understanding of patients' needs and limitations.

Participants also opined that the technology should enable the user to provide support to others, in this case, by sharing knowledge and experiences with their peers in Singapore. They suggested that apps should be designed to provide social support and facilitate connectedness with others. For example, when other caregivers encounter difficulties when managing flare-ups, their peers could support them and detail their own experiences. Besides, peers can also comfort caregivers emotionally by recognizing their efforts and mitigating their guilt. This might be accomplished using a web-based forum with messaging, chat, or call functions. These were proposed as modes through which other caregivers can provide support for each other.

In addition, both caregivers and HCPs suggested that the app should be personalized. The personalization should embody 2 aspects: (1) features and content should be tailored to each individual's needs and (2) features and content should suit the local context. In this way, the app could help them with prioritizing and making new choices and with defining their own goals and the small steps necessary for reaching these goals. During the workshops, the participants advised that the app should be personalized to address every patient's unique circumstances. They recommended that the app should provide all the information needed as soon as a client first uses the app (quote 18 in Multimedia Appendix 3).

Image analysis was also mentioned during the workshop. It was recommended that caregivers could take photographs and use the app to analyze each photograph to measure the severity. They also advised that the app should give advice or suggestions corresponding to the extent and severity of the skin condition, so that clients can get *immediate help without seeing the doctor*.

Wireframe Prototype Development

The wireframe prototype contained features designed to address the challenges and confusions mentioned in the previous workshops. The *must-have* features and *nice-to-have* features recommended by participants, challenges addressed by the features, and how the final wireframe prototype reflects the features are summarized in Table 3.

XSL•FO
RenderX

Table 3. Ideal features for wireframe prototype.

Features	Challenges	Possible solutions
Must-have features		
Knowledge on medication use and their side effects	Which medication to choose and when to use it	Chatbot (Figure 2)
Knowledge on medication use and their side effects	Generic instruction versus specific instruction	Chatbot (Figure 2)
Knowledge on medication use and their side effects	Alternative treatments preferable to steroids	Chatbot (Figure 2)
Knowledge on the symptoms and appropriate management	Understanding the disease course and the complications related to AD[a]	Chatbot (Figure 2)
Knowledge on the external triggers leading to AD[a] flares	Understanding the causes, triggers, and symptoms of AD	Chatbot (Figure 2)
Monitoring of skin conditions	Understanding the causes, triggers, and symptoms of AD	Disease diary (Figure 2)
Monitoring of skin conditions	Understanding the disease course and the complications related to AD	Journal (Figure 2)
Communication channels with health care professionals	Understanding the causes, triggers, and symptoms of AD	Teleconsultation (Figure 2)
Communication channels with health care professionals	Understanding the disease course and the complications related to AD	Teleconsultation (Figure 2)
Communication channels with health care professionals	Which medication to choose and when to use it	Teleconsultation (Figure 2)
Communication channels with health care professionals	Generic instruction versus specific instruction	Teleconsultation (Figure 2)
Communication channels with health care professionals	Alternative treatments preferable to steroids	Teleconsultation (Figure 2)
Monitoring of external triggers that may lead to AD flares	Understanding the causes, triggers, and symptoms of AD	Disease diary (Figure 2) and disease monitor (Figure 2)
Nice-to-have features		
Debunking the myths about AD	Alternative treatments preferable to steroids	Chatbot (Figure 2)
Knowledge on the natural history and development of AD	Understanding the causes, triggers, and symptoms of AD	Chatbot (Figure 2)
Reminders for topical application and medication taking	Which medication to choose and when to use it	Disease monitor (Figure 2)
Peer support from other caregivers or patients	Understanding the causes, triggers, and symptoms of AD	Forum (Figure 2)
Peer support from other caregivers or patients	Understanding the disease course and the complications related to AD	Forum (Figure 2)
Peer support from other caregivers or patients	Which medication to choose and when to use it	Forum (Figure 2)
Peer support from other caregivers or patients	Generic instruction versus specific instruction	Forum (Figure 2)
Peer support from other caregivers or patients	Alternative treatments preferable to steroids	Forum (Figure 2)
Peer support from other caregivers or patients	Guilt and frustration	Forum (Figure 2)
Peer support from other caregivers or patients	Self-esteem	Forum (Figure 2)
Peer support from other caregivers or patients	Duty to make their child feel better	Forum (Figure 2)
Monitoring of caregivers' and children's emotions	Guilt and frustration	Disease diary (Figure 2)
Monitoring of caregivers' and children's emotions	Self-esteem	Journal (Figure 2)
Monitoring of caregivers' and children's emotions	Duty to make their child feel better	Journal (Figure 2)
Emotional support with interactive app features	Guilt and frustration	Forum (Figure 2)
Emotional support with interactive app features	Self-esteem	Teleconsultation (Figure 2)
Emotional support with interactive app features	Duty to make their child feel better	Teleconsultation (Figure 2)

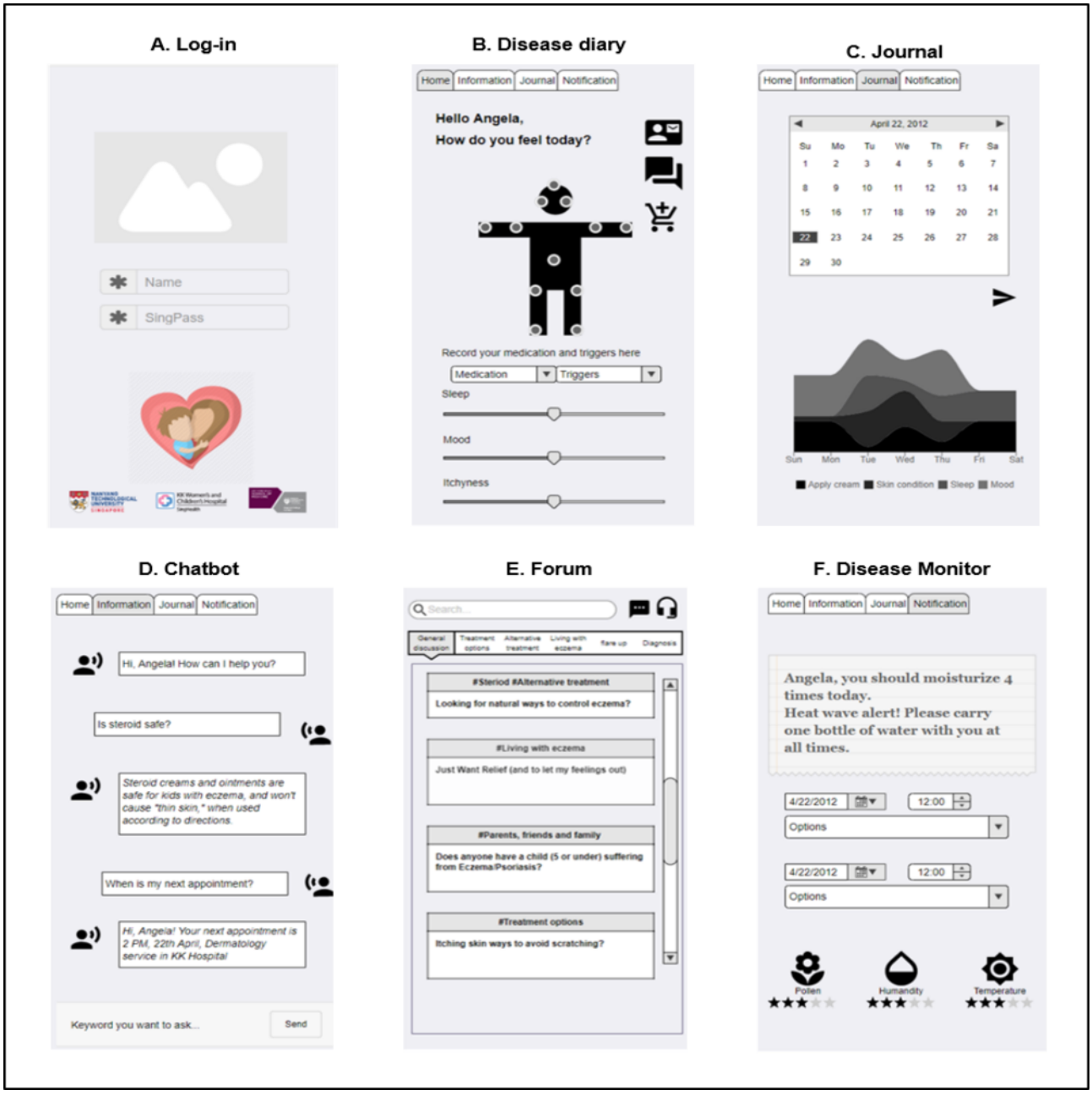
[a]AD: atopic dermatitis.

As shown in the master sketch of the prototype, the prototype allows users to log in with their account (Figure 2) and record lesions' photos of AD in different body parts and provides instructions relating to treatment, symptoms, triggers, and people's emotions (Figure 2). The *Journal* section enables caregivers to track the patient's disease condition and could

XSL•FO
RenderX

prove useful to their doctor during a visit to the clinic (Figure 2).

Figure 2. The master sketch of wireframe prototype.



Another core feature of the prototype is the education element. Participants designed both the chatbot feature and forum feature, both of which are for educational purposes (Figure 2). The chatbot allows users to access relevant information quickly and easily, which is presented in a clear, understandable format. The forum feature builds a bridge between caregivers and their peers by letting them exchange local information and, to some extent, allowing peers to provide informational and emotional support to users (Figure 2). Participants also designed a notification feature, through which the prototype could remind users to avoid the possible triggers around them and decrease AD relapse (Figure 2).

Workshop Evaluation Feedback

The workshop evaluation comments indicated that the co-design workshop was successful in creating and generating new ideas and content for smartphone app development. All the participants agreed that the workshop was organized, that its aims were clear, and its activities had been carefully prepared. All the participants felt that the workshop methods used were appropriate for the audience and said that they enjoyed the workshop. However, some issues were also identified in the evaluation feedback. For example, 21% (7/32) of participants highlighted the need to improve workshop time management and felt that the time should be better utilized; they even suggested a longer session.

XSL•FO
RenderX

Discussion

Summary

For this study, we organized a series of co-design workshops, which were attended by caregivers living with children with AD, doctors, pharmacists, and digital health experts. During the proceedings, participants jointly explored their needs, preferences, and perceptions in relation to the proposed smartphone self-management app. Phase 1 identified themes that reflected caregivers' and HCPs' needs and concerns for AD self-management support, along with suggestions concerning the necessary features and contents of an AD self-management app; the data collected during phase one, therefore, informed phase two, which in turn comprised a smartphone app designing activity. The purpose of the interactive designing exercise was for participants to develop a wireframe prototype. The co-design workshops were designed to encourage social interaction, stimulate creativity, and thus help us to articulate the root of the clinical problems. The findings from the workshops align with those in previous studies, in that co-design interaction can foster a co-creative space and lead to considerable contributions by and involvement of people with chronic disease conditions [40,41].

The wireframe prototype has the potential to solve current AD management challenges. First, information gathered by disease diary and disease monitoring features (Figure 2) in prototype are quite comprehensive with both visual AD symptoms (by taking photos), personal feelings (sleep disturbance, itchiness), psychological impact (mood), and suspected allergens, which could be very helpful for doctors in making precise clinical decisions. In addition, involvement of HCPs in the app development process could ensure that the information provided is reliable. Third, the forum feature (Figure 2) could potentially become a peer support platform. By communicating with each other, caregivers can support each other technically and emotionally and, thus, enable a smooth transition following a consultation and empower caregivers to manage their children's disease by making them feel competent and capable.

Research Implications

Some issues were also identified during the workshop. First, it is important to ensure that all the participants fully understand and can keep pace with workshop information and activities—or, to use the colloquialism, *read off the same page*. The participants may interpret questions or tasks differently, and there is a risk of digressing the topic or narrating the same topic repeatedly, which can lead to the workshop being less productive in terms of producing the ideas or thoughts anticipated. Second, at the start of third workshop, we delivered a briefing presentation of the current research and technology used for AD management, to encourage the participants to allow their own needs and challenges to guide app co-design. However, it should be mentioned that the briefing presentation might have limited their thinking, and they may have been inclined to follow the information shared during the briefing session.

Strengths and Limitations

This study can inform future research studies on co-designing smartphone apps for patients affected by chronic diseases. Participants with a heterogeneous background were invited to co-design the app. This could be considered as a strength of this study as the app design by them will be understandable and beneficial for the wider population regardless of user background. Previous studies have explored treatment needs widely: Batchelor et al [42] tried to identify the uncertainties in eczema treatment that are important to stakeholders. Teasdale et al [43] and Swallow [44] discussed the caregivers' views of steroid usage [43,44]. However, this chapter goes further in exploring personal experiences, emotional needs, and feelings among both caregivers and HCPs and considering how their points of view could be linked to features of smartphone apps.

However, the scope of the study is limited as the target audience of the app consists entirely of caregivers whose children have AD. However, that limitation alone does not preclude us from being able to satisfy the aims of this study because the content and features caregivers mentioned are to some extent interchangeable. Some of the issues and features caregivers mentioned can also be favored for both children and adults. Second, the co-design process is not rigorously following a participatory co-design framework as continuous evaluation of the prototype was not performed [45]. However, considering that this study has gone through the main process of the framework, this study is still aligned with the overall goal of co-design methodology [29].

Another limitation is that participants were observed to discuss some topics repetitively. Thematic saturation of the workshops was probably less than satisfactory as other themes were barely discussed in phase one, and it might have emerged that we conducted more workshops. However, this limitation can be partly explained as those viewpoints raised repetitively were valued more than other topics from the caregiver's perspective and, therefore, should be prioritized when designing an AD self-management app. In addition, as the final wireframe prototype is designed by both digital health experts and caregivers, it is difficult for us to differentiate caregivers' perspectives and digital health experts' opinions in phase two.

In summary, this study employed a co-design approach to jointly create an AD self-management app wireframe prototype, together with caregivers, HCPs, and digital health experts. The participants' perceptions and preferences as well as the co-design approach used in this study form a novel contribution to the canon of literature on smartphone app design and the use of digital health tools in an engaging way. Future studies should work on app development as well as invite more stakeholders with separate sessions to test its functionality and usability.

XSL•FO
RenderX

Acknowledgments

This research was funded by a principal investigator start-up grant (Lee Kong Chian School of Medicine, Nanyang Technological University). JC's post at Imperial College London is supported by the National Institute for Health Research NW London Applied Research Collaboration.

Authors' Contributions

XX and JC were involved in the study conception and design and in obtaining ethical approval. XX, MK, KG, JC, and WS were involved in the data collection. XX and KG were involved in thematic data analysis. The manuscript was written by XX. All authors commented on the draft and agreed with the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire for eczema smartphone app co-design workshop.
[DOCX File , 177 KB - mhealth_v8i10e16898_app1.docx]

Multimedia Appendix 2
Workshop evaluation form.
[DOCX File , 77 KB - mhealth_v8i10e16898_app2.docx]

Multimedia Appendix 3
Illustrative quotes.
[DOCX File , 19 KB - mhealth_v8i10e16898_app3.docx]

References

1. Lewis-Jones MS, Finlay AY. The children's dermatology life quality index (CDLQI): initial validation and practical use. Br J Dermatol 1995 Jun;132(6):942-949. [doi: 10.1111/j.1365-2133.1995.tb16953.x] [Medline: 7662573]
2. Mancini AJ, Kaulback K, Chamlin SL. The socioeconomic impact of atopic dermatitis in the United States: a systematic review. Pediatr Dermatol 2008;25(1):1-6. [doi: 10.1111/j.1525-1470.2007.00572.x] [Medline: 18304144]
3. Thomsen SF. Atopic dermatitis: natural history, diagnosis, and treatment. ISRN Allergy 2014;2014:354250 [FREE Full text] [doi: 10.1155/2014/354250] [Medline: 25006501]
4. Lifschitz C. The impact of atopic dermatitis on quality of life. Ann Nutr Metab 2015;66(Suppl 1):34-40 [FREE Full text] [doi: 10.1159/000370226] [Medline: 25925339]
5. Cheung R, Nelson W, Advincula L, Young Cureton V, Canham DL. Understanding the culture of Chinese children and families. J Sch Nurs 2005 Feb;21(1):3-9. [doi: 10.1622/1059-8405(2005)021[0003:utcocc]2.0.co;2] [Medline: 15660489]
6. Beattie PE, Lewis-Jones MS. An audit of the impact of a consultation with a paediatric dermatology team on quality of life in infants with atopic eczema and their families: further validation of the Infants' Dermatitis Quality of Life Index and Dermatitis Family Impact score. Br J Dermatol 2006 Dec;155(6):1249-1255. [doi: 10.1111/j.1365-2133.2006.07525.x] [Medline: 17107397]
7. Moore K, David TJ, Murray CS, Child F, Arkwright PD. Effect of childhood eczema and asthma on parental sleep and well-being: a prospective comparative study. Br J Dermatol 2006 Mar;154(3):514-518. [doi: 10.1111/j.1365-2133.2005.07082.x] [Medline: 16445784]
8. Cheung WK, Lee RL. Children and adolescents living with atopic eczema: an interpretive phenomenological study with Chinese mothers. J Adv Nurs 2012 Oct;68(10):2247-2255. [doi: 10.1111/j.1365-2648.2011.05915.x] [Medline: 22221066]
9. Su JC, Kemp AS, Varigos GA, Nolan TM. Atopic eczema: its impact on the family and financial cost. Arch Dis Child 1997 Feb;76(2):159-162. [doi: 10.1136/adc.76.2.159] [Medline: 9068310]
10. Lee BW, Detzel PR. Treatment of childhood atopic dermatitis and economic burden of illness in Asia Pacific countries. Ann Nutr Metab 2015;66(Suppl 1):18-24 [FREE Full text] [doi: 10.1159/000370221] [Medline: 25925337]
11. Kessels RP. Patients' memory for medical information. J R Soc Med 2003 May;96(5):219-222 [FREE Full text] [doi: 10.1258/jrsm.96.5.219] [Medline: 12724430]
12. Hanifin JM, Tofte SJ. Patient education in the long-term management of atopic dermatitis. Dermatol Nurs 1999 Aug;11(4):284-289. [Medline: 10670359]
13. Morrison D, Wyke S, Agur K, Cameron EJ, Docking RI, Mackenzie AM, et al. Digital asthma self-management interventions: a systematic review. J Med Internet Res 2014 Feb 18;16(2):e51 [FREE Full text] [doi: 10.2196/jmir.2814] [Medline: 24550161]

XSL•FO
RenderX

14. Sheehy S, Cohen G, Owen KR. Self-management of diabetes in children and young adults using technology and smartphone applications. Curr Diabetes Rev 2014;10(5):298-301. [doi: 10.2174/1573399810666141010113050] [Medline: 25311195]
15. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
16. 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 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
17. Sun F. Caregiving stress and coping: a thematic analysis of Chinese family caregivers of persons with dementia. Dementia (London) 2014 Nov;13(6):803-818. [doi: 10.1177/1471301213485593] [Medline: 24339084]
18. Lewis-Jones S, Mugglestone MA, Guideline Development Group. Management of atopic eczema in children aged up to 12 years: summary of NICE guidance. Br Med J 2007 Dec 15;335(7632):1263-1264 [FREE Full text] [doi: 10.1136/bmj.39405.503773.AD] [Medline: 18079551]
19. Eichenfield LF, Tom WL, Chamlin SL, Feldman SR, Hanifin JM, Simpson EL, et al. Guidelines of care for the management of atopic dermatitis: section 1. Diagnosis and assessment of atopic dermatitis. J Am Acad Dermatol 2014 Feb;70(2):338-351 [FREE Full text] [doi: 10.1016/j.jaad.2013.10.010] [Medline: 24290431]
20. Eichenfield LF, Tom WL, Berger TG, Krol A, Paller AS, Schwarzenberger K, et al. Guidelines of care for the management of atopic dermatitis: section 2. Management and treatment of atopic dermatitis with topical therapies. J Am Acad Dermatol 2014 Jul;71(1):116-132 [FREE Full text] [doi: 10.1016/j.jaad.2014.03.023] [Medline: 24813302]
21. Tay YK, Chan YC, Chandran NS, Ho MS, Koh MJ, Lim YL, et al. Guidelines for the management of atopic dermatitis in Singapore. Ann Acad Med Singap 2016 Oct;45(10):439-450 [FREE Full text] [Medline: 27832218]
22. Hersh W, Helfand M, Wallace J, Kraemer D, Patterson P, Shapiro S, et al. A systematic review of the efficacy of telemedicine for making diagnostic and management decisions. J Telemed Telecare 2002;8(4):197-209. [doi: 10.1258/135763302320272167] [Medline: 12217102]
23. Santer M, Muller I, Yardley L, Burgess H, Selinger H, Stuart BL, et al. Supporting self-care for families of children with eczema with a web-based intervention plus health care professional support: pilot randomized controlled trial. J Med Internet Res 2014 Mar 4;16(3):e70 [FREE Full text] [doi: 10.2196/jmir.3035] [Medline: 24594972]
24. Armstrong AW, Johnson MA, Lin S, Maverakis E, Fazel N, Liu F. Patient-centered, direct-access online care for management of atopic dermatitis: a randomized clinical trial. JAMA Dermatol 2015 Feb;151(2):154-160. [doi: 10.1001/jamadermatol.2014.2299] [Medline: 25338198]
25. van Os-Medendorp H, Koffijberg H, Eland-de Kok PC, van der Zalm A, de Bruin-Weller MS, Pasmans SG, et al. E-health in caring for patients with atopic dermatitis: a randomized controlled cost-effectiveness study of internet-guided monitoring and online self-management training. Br J Dermatol 2012 May;166(5):1060-1068. [doi: 10.1111/j.1365-2133.2012.10829.x] [Medline: 22268960]
26. Flaten HK, St Claire C, Schlager E, Dunnick CA, Dellavalle RP. Growth of mobile applications in dermatology - 2017 update. Dermatol Online J 2018 Feb 15;24(2):- [FREE Full text] [Medline: 29630159]
27. van Galen LS, Xu X, Koh MJ, Thng S, Car J. Eczema apps conformance with clinical guidelines: a systematic assessment of functions, tools and content. Br J Dermatol 2020 Feb;182(2):444-453. [doi: 10.1111/bjd.18152] [Medline: 31179535]
28. Scott J, Heavey E, Waring J, Jones D, Dawson P. Healthcare professional and patient codesign and validation of a mechanism for service users to feedback patient safety experiences following a care transfer: a qualitative study. BMJ Open 2016 Jul 12;6(7):e011222. [doi: 10.1136/bmjopen-2016-011222] [Medline: 27406641]
29. Woods L, Cummings E, Duff J, Walker K. Design thinking for mhealth application co-design to support heart failure self-management. Stud Health Technol Inform 2017;241:97-102. [Medline: 28809190]
30. Verbiest ME, Corrigan C, Dalhousie S, Firestone R, Funaki T, Goodwin D, et al. Using codesign to develop a culturally tailored, behavior change mHealth intervention for indigenous and other priority communities: A case study in New Zealand. Transl Behav Med 2019 Jul 16;9(4):720-736. [doi: 10.1093/tbm/iby093] [Medline: 30388262]
31. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mhealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]
32. Nutten S. Atopic dermatitis: global epidemiology and risk factors. Ann Nutr Metab 2015;66(Suppl 1):8-16 [FREE Full text] [doi: 10.1159/000370220] [Medline: 25925336]
33. Xu X, van Galen LS, Koh MJ, Bajpai R, Thng S, Yew YW, et al. Factors influencing quality of life in children with atopic dermatitis and their caregivers: a cross-sectional study. Sci Rep 2019 Nov 5;9(1):15990 [FREE Full text] [doi: 10.1038/s41598-019-51129-5] [Medline: 31690745]
34. Xu X, Galen LS, Bajpai R, Koh M, Thng S, Yew Y, et al. Factors influencing family burden of paediatric patients with atopic dermatitis: a cross-sectional study. Acta Derm Venereol 2020 Jul 28;100(14):adv00212 [FREE Full text] [doi: 10.2340/00015555-3458] [Medline: 32179930]
35. Abu-Dalbouh. A questionnaire approach based on the technology acceptance model for mobile tracking on patient progress applications. Journal of Computer Science 2013 Jun 01;9(6):763-770. [doi: 10.3844/jcssp.2013.763.770]

XSL•FO
RenderX

36. Azza AJ. How is Technology Accepted by Users? A Review of Technology Acceptance Models and Theories. In: Proceedings of The IRES 17th International Conference. 2015 Presented at: IRES'15; December 16-17, 2015; London, United Kingdom.
37. Holahan PJ, Lesselroth BJ, Adams K, Wang K, Church V. Beyond technology acceptance to effective technology use: a parsimonious and actionable model. J Am Med Inform Assoc 2015 May;22(3):718-729. [doi: 10.1093/jamia/ocu043] [Medline: 25773129]
38. Houghton C, Murphy K, Shaw D, Casey D. Qualitative case study data analysis: an example from practice. Nurse Res 2015 May;22(5):8-12. [doi: 10.7748/nr.22.5.8.e1307] [Medline: 25976531]
39. Richards L. Handling Qualitative Data: A Practical Guide. Thousand Oaks, CA: Sage Publications; 2005.
40. Clemensen J, Rothmann MJ, Smith AC, Caffery LJ, Danbjorg DB. Participatory design methods in telemedicine research. J Telemed Telecare 2017 Oct;23(9):780-785. [doi: 10.1177/1357633X16686747] [Medline: 28027678]
41. Lupton D. Digital health now and in the future: findings from a participatory design stakeholder workshop. Digit Health 2017;3:2055207617740018. [doi: 10.1177/2055207617740018] [Medline: 29942616]
42. Batchelor JM, Ridd MJ, Clarke T, Ahmed A, Cox M, Crowe S, et al. The Eczema Priority Setting Partnership: a collaboration between patients, carers, clinicians and researchers to identify and prioritize important research questions for the treatment of eczema. Br J Dermatol 2013 Mar;168(3):577-582. [doi: 10.1111/bjd.12040] [Medline: 22963149]
43. Teasdale EJ, Muller I, Santer M. Carers' views of topical corticosteroid use in childhood eczema: a qualitative study of online discussion forums. Br J Dermatol 2017 Jun;176(6):1500-1507. [doi: 10.1111/bjd.15130] [Medline: 27753076]
44. Swallow V. Parents managing childhood long-term conditions: longitudinal research is needed on online forums as sources of peer-to-peer information. Br J Dermatol 2017 Jun;176(6):1430-1431. [doi: 10.1111/bjd.15563] [Medline: 28581231]
45. Snijders T, Bosker R. The Worship and Principles of the Church of England: a Sermon, Preached at the Opening of Christ's Church, in Great-barrington, on Christ-mass Day, M.dcc.lxiv. London, UK: William Goddard; 1999.

Abbreviations

AD: atopic dermatitis
HCP: health care professional
IRR: interrater reliability
QoL: quality of life

Edited by G Eysenbach; submitted 04.11.19; peer-reviewed by T Loncar-Turukalo, G Humphrey; comments to author 23.12.19; revised version received 26.03.20; accepted 03.06.20; published 29.10.20.

Please cite as:
Xu X, Griva K, Koh M, Lum E, Tan WS, Thng S, Car J
Creating a Smartphone App for Caregivers of Children With Atopic Dermatitis With Caregivers, Health Care Professionals, and Digital Health Experts: Participatory Co-Design
JMIR Mhealth Uhealth 2020;8(10):e16898
URL: http://mhealth.jmir.org/2020/10/e16898/
doi:10.2196/16898
PMID:33118949

©Xiaomeng Xu, Konstadina Griva, Mark Koh, Elaine Lum, Woan Shin Tan, Steven Thng, Josip Car. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 29.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of an Intervention Targeting Multiple Health Behaviors Among High School Students: Participatory Design Study Using Heuristic Evaluation and Usability Testing

Ulrika Müssener[1], PhD; Kristin Thomas[1], PhD; Catharina Linderoth[1], MSc; Marie Löf[1,2], PhD; Katarina Åsberg[1], MSc; Pontus Henriksson[1], PhD; Marcus Bendtsen[1], PhD

[1]Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

[2]Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden

Corresponding Author:
Ulrika Müssener, PhD
Department of Health, Medicine and Caring Sciences
Linköping University
Sandbäcksgatan 1
Linköping, 58183
Sweden
Phone: 46 732702426
Email: ulrika.mussener@liu.se

Abstract

Background: Mobile electronic platforms provide exciting possibilities for health behavior promotion. For instance, they can promote smoking cessation, moderate alcohol consumption, healthy eating, and physical activity. Young adults in Sweden are proficient in the use of technology, having been exposed to computers, smartphones, and the internet from an early age. However, with the high availability of mobile health (mHealth) interventions of varying quality, it is critical to optimize the usability of mHealth interventions to ensure long-term use of these health promotion interventions.

Objective: This study aims to investigate the usability of an mHealth intervention (LIFE4YOUth) targeting health behaviors among high school students through heuristic evaluation and usability testing.

Methods: A preliminary version of the LIFE4YOUth mHealth intervention, which was aimed at promoting healthy eating, physical activity, smoking cessation, and nonrisky drinking among high school students, was developed in early 2019. We completed a total of 15 heuristic evaluations and 5 usability tests to evaluate the usability of the mHealth intervention prototype to improve its functioning, content, and design.

Results: Heuristic evaluation from a total of 15 experts (10 employees and 5 university students, both women and men, aged 18-25 years) revealed that the major usability problems and the worst ratings, a total of 17 problems termed *usability catastrophes*, concerned shortcomings in displaying easy-to-understand information to the users or technical errors. The results of the usability testing including 5 high school students (both girls and boys, aged 15-18 years) showed that the design, quality, and quantity of content in the intervention may impact the users' level of engagement. Poor functionality was considered a major barrier to usability. Of the 5 participants, one rated the LIFE4YOUth intervention as poor, 2 rated as average, and 2 assessed it as good, according to the System Usability Scale.

Conclusions: High school students have high expectations of digital products. If an mHealth intervention does not offer optimal functions, they may cease to use it. Optimizing the usability of mHealth interventions is a critical step in the development process. Heuristic evaluation and usability testing in this study provided valuable knowledge about the prototype from a user's perspective. The findings may lead to the development of similar interventions targeting the high school population.

(JMIR Mhealth Uhealth 2020;8(10):e17999) doi:10.2196/17999

KEYWORDS

mHealth intervention; health behavior; high school students; participatory design; heuristic evaluation; usability testing; mobile phone

XSL•FO
RenderX

Introduction

Health Behaviors Among Young People

Chronic diseases are the leading cause of death and disability worldwide. Globally, up to an estimated 80% of cases of heart disease, stroke, and type 2 diabetes and more than 30% of cancers can be prevented by reducing smoking, harmful alcohol use, improving diet, and engaging in regular physical activity [1,2]. Previous research has shown that multiple risk behaviors increase the risk of chronic disease and all-cause mortality, more so than the combined effects of single behaviors [3-5]. Health behaviors typically emerge during adolescence, track into adulthood, and commonly co-occur [6-8]. Therefore, adolescence is a critical age to intervene and interrupt a trajectory toward poor adult health [9-11]. Evidently, effective and evidence-based health behavior promotion interventions are needed.

Health Behavior Promotion Among Youths Through Multiple Mobile Health Interventions

Mobile platforms provide exciting possibilities for the promotion of health behaviors through mobile health (mHealth) interventions. Previous research has shown that interventions targeting multiple health behaviors at the same time might be effective in improving the general lifestyle among adults [5,12], with less evidence among adolescents [13]. A meta-analysis [14] examined the effectiveness of text message–based interventions for tobacco and alcohol cessation in a young adult population. Only 5 of the 14 studies reported significant differences between groups of substance use behavior outcomes, and the included randomized controlled trials (RCTs) lacked details regarding the intervention content. A more recent systematic review and meta-analysis [15] investigated the effectiveness of school-based eHealth interventions, defined as interventions delivered via the internet, computers, mobile technology, or telehealth, to prevent multiple health risk behaviors among adolescents. A total of 22 publications were included, assessing 16 interventions that targeted 2 or more of the following behaviors: alcohol use, smoking, diet, physical activity, screen time and sitting, or sleep. Only short-term effects were found for improving physical activity, screen time, and fruit and vegetable intake, and all trials were considered to be of low quality. There was limited evidence of their effect on alcohol and smoking. Although the selection criteria in the meta-analysis included studies with intervention components delivered via mobile technology, no studies of mHealth-only interventions were identified [15].

The Development of mHealth Interventions

Young adults are referred to as digital natives; many are proficient in the use of technology, having been exposed to computers, smartphones, and the internet from an early age. Indeed, young adults have the highest level of smartphone ownership among all age groups [16]. It is critical to optimize the usability of interventions targeting this age group; if they do not enjoy the program, they may disengage. Several reviews have shown the feasibility, acceptability, and efficacy of digital interventions for behavior change [17,18] and other health interventions, such as disease self-management [19-21] and medication adherence [22,23]. Previous systematic reviews on texting and mobile app interventions emphasized the urgent need to examine the development processes of mHealth interventions [23,24].

In a participatory design approach, research is undertaken with, rather than on, people, allowing researchers to gain an understanding of context-specific requirements and challenges [25]. Heuristic evaluation and usability testing are commonly used to support the development and refinement of the effectiveness of mHealth interventions. During heuristic evaluation, trained evaluators review an intervention to find usability problems, assign them to a specific category of heuristics, and ascribe a severity rating to provide distinct usability information [26,27]. Heuristics are often used to identify usability issues, such as problems with unclear functions, confusing navigation, and consistency issues [27-29]. Usability tests consist of a human-computer interaction and refer to evaluating an intervention by testing it with potential end users with the goal of identifying understandability, ease of learning, and attractiveness and to determine participants' satisfaction with the intervention. Usability testing provides developers with feedback about what does and does not work in the intervention and determines whether the features in the interventions are acceptable to and feasible for users and also determines what can be improved [30-33].

Objectives

This study aims to investigate the usability of an mHealth intervention (LIFE4YOUth) targeting health risk behaviors among high school students through heuristic evaluation and usability testing.

Methods

Procedures

A preliminary version of the LIFE4YOUth mHealth intervention targeting health risk behaviors among high school students was developed in early 2019. LIFE4YOUth is one of 7 multiple mHealth interventions in the MoBILE (mHealth–Multiple Lifestyle Behaviors) research program (funded by Forte 2018-01410, principal investigator: ML) aimed at promoting healthy eating, physical activity, smoking cessation, and nonrisky drinking in 7 different populations in the health care system. The intervention includes information about health behaviors, tips on behavior change strategies, and activities. The formative research process of developing a novel multiple mHealth intervention is described in detail in a study protocol elsewhere [34]. This study reports on findings from the first stage of the formative research process: heuristic evaluation and usability testing. The aim is to investigate the usability of a prototype app in terms of function, content, and design.

Setting, Participants, and Recruitment

Heuristic Evaluation

Recruitment of participants for the heuristic evaluation was undertaken by members of the research team through paper advertisements (posters) in public areas at Linköping University, Sweden. The inclusion criteria for the heuristic evaluation were

XSL•FO
RenderX

university students and employees, both women and men, aged 18 to 25 years, at the Faculty of Medicine and Health Sciences at Linköping University who were willing to participate and owned a mobile phone. Participants showed their interest by contacting the research leader via email or telephone.

Usability Tests

School staff at 5 high schools selected for convenience in Östergötland, Sweden, were contacted via email and informed about the research project. Students from all schools were invited to participate in the usability testing. The recruitment of participants was performed by school staff through paper advertisements (posters and leaflets), digital advertisements (student email and school website), and information disseminated in the classrooms. The inclusion criteria for the usability tests were high school students, both female and male, aged 15 to 18 years, willing to participate and owning a mobile phone. High school students showed their interest by contacting the research leader via email or telephone.

Data Collection

Informed Consent

All participants provided written informed consent before participation in all study procedures (heuristic evaluation and usability tests).

Heuristic Evaluations

A total of 15 experts (10 employees and 5 university students) were recruited. For the heuristic evaluation, a set of 10 heuristics published by Nielsen [26] was used to evaluate the LIFE4YOUth prototype. The heuristics for usability evaluation according to Nielsen are listed in Textbox 1.

We selected Nielsen's 10 heuristics because they have been thoroughly tested and are quick and easy to apply. When applying heuristic evaluation, participants evaluate an app to find usability problems, assign them to a specific category of heuristics, and assign a severity rating [26,27]. All participants were invited to a brief training session (approximately 45 min) conducted by a research assistant (CL), to receive instructions on the main principles of heuristic evaluation, and to learn to use the heuristics to evaluate the intervention [27]. The participants were sent a link to a high-fidelity prototype [35] of the intervention, including the actual software start page, menu page, and the 4 intervention modules (alcohol, smoking, physical activity, and diet). Each participant was asked to identify usability problems independently in a given protocol (Multimedia Appendix 1). Participants were asked to identify a problem, describe it, identify the relevant heuristic for the problem (eg, visibility of system status or match between the system and the real world), and give it a severity rating [26,27,29].

The procedure itself was a two-part process: the participants first familiarized themselves with the system and its usage with reference to the materials and training provided by the assistant. Then, they independently applied the 10 heuristics, as given in Textbox 1. The participants detected a usability problem, assigned each problem to a violation of a heuristic, and described the problem in their own words. Then, participants assigned severity scores to each problem by using the severity rating factors of impact presented in Textbox 2.

The introduction took place in the beginning of May 2019 in a conference room at Linköping University, Sweden. The evaluations were performed wherever the participants preferred and were sent to the research assistant in a prepaid envelope. After 10 days, a reminder was sent by a text message. Heuristic evaluations from all participants (n=15) were gathered during the last week of May 2019.

Textbox 1. Heuristics for usability evaluation according to Nielsen.

1. Visibility of system status: The system should always keep users informed about what is going on through appropriate feedback within reasonable time
2. Match between system and the real world: The system should speak the users' language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order
3. User control and freedom: Users often choose system functions by mistake and will need a clearly marked *emergency exit* to leave the unwanted state without having to go through an extended dialog. Support undo and redo
4. Consistency and standards: Users should not have to wonder whether different words, situations, or actions mean the same thing
5. Error prevention: Even better than good error messages is a careful design, which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action
6. Recognition rather than recall: Minimize users' memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialog to another. Instructions for the use of the system should be visible or easily retrievable whenever appropriate
7. Flexibility and efficiency of use: Accelerators—unseen by the novice user—may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions
8. Aesthetic and minimalist design: Dialogs should not contain information that is irrelevant or rarely needed. Every extra unit of information in a dialog competes with the relevant units of information and diminishes their relative visibility
9. Help users recognize, diagnose, and recover from errors: Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution
10. Help and documentation: Although it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large

XSL•FO
RenderX

Textbox 2. Scale for severity rating according to Nielsen.

0: Not a usability problem
1: Cosmetic problem only—need not to be fixed unless extra time is available
2: Minor usability problem—fixing this should be given low priority
3: Major usability problem—important to fix, should be given high priority
4: Usability catastrophe—imperative to fix this before product can be released

Usability Tests

The usability test consisted of a human-computer interaction evaluation, which focused on the perceptions and performance of users in a laboratory setting. It entailed the completion of 30 goal-oriented tasks by targeted end users [30,32]. In total, 5 usability tests were completed with high school students. The tests were carried out in June 2019 in a medical informatics laboratory room and were run on an iPhone. The participants went through a 60-min session, during which all interactions with the intervention were recorded using video. The same prototype, as in the heuristic evaluation, was used. The test manager gave the participants printouts with the 30 tasks. Examples of tasks used in the usability testing are described in Textbox 3.

The participants were asked to complete the 30 tasks, and the test manager asked the participants to explain their actions as they performed them using a think-aloud method [36,37]. An observer (the research leader) noted potential issues as the given tasks were performed by the participants. After completing the session, the participants completed a paper version of the System Usability Scale (SUS), providing a global view of their subjective assessment of usability based on 10 questions [38]. SUS has been shown to be reliable and valid with a variety of different technologies and users.

Textbox 3. Examples of tasks in the usability testing.

- Go to the start page and explain what you think of the layout
- Where can you find support to drink less?
- Log what you drink. Explain how you interpret the graphs for the alcohol consumption you logged
- You know that people feel healthier after exercising and want to find out more about getting started. How would you look for this information?
- If you want to plan a physical activity, how do you find that activity, and specify your level of involvement?
- You can set personal goals for your eating habits under the “Timeline” tab. Explain how to set your goal and what the graph shows you
- What information is there under the “Risks” tab in the Diet module? Does it give you a good overview?
- Go to the Smoking module. What do you think about the scope of the information provided on the first page?
- If you want to know more about the benefits of quitting smoking, how would you go about this?

Data Analyses

Heuristic Evaluation

The 10 heuristics were pooled, and the identified problems were categorized as major issues [26-29]. The focus of the analysis was to identify usability problems and critical issues and to explore whether any functions in the intervention performed better than others. A master list was compiled to collect all the described usability problems, duplicate problems were removed to enable analyses, and the list was verified by 2 of the authors (UM and CL) for accuracy and to ensure validity and prevent bias in the analysis process. Descriptive statistics were used to summarize the heuristic violations and associated severity scores. The severity rating scale from 0 (not a problem) to 4 (usability catastrophe) is shown in Textbox 2.

Usability Tests

The usability tests were recorded and transcribed verbatim by a professional transcription company. All transcripts were checked and validated by the first author (UM). Analysis of the video recordings was inspired by program theory development using an inductive approach, taking both verbal and visible conduct into consideration [39]. Transcripts were analyzed thematically using an iterative coding procedure. The focus of the analysis was on the features of the intervention that needed to be redesigned with regard to function, content, and design. Overall, 2 authors (UM and CL) individually read the transcriptions and viewed the video recordings to acquire a comprehensive understanding. The categories were identified using an iterative process of reading and rereading the transcripts. Patterns were searched for, and usability issues were coded into categories. The first coding was initiated by the first author. Next, the coding was presented and discussed between the 2 authors (UM and CL), and boundaries for coding were established jointly [39].

The analysis of the SUS score was conducted according to the scoring strategy of Brooke [40], with the score for each item ranging from 0 to 4. The score from positively worded items (1, 3, 5, 7, and 9) is calculated as the scale position minus 1. For the negatively worded items (2, 4, 6, 8, and 10), the score is calculated as 5 minus the scale position. The sum of the scores is then multiplied by 2.5 to obtain the overall value of SUS ranging from 0 to 100. The average SUS scores were used to

XSL•FO
RenderX

identify average satisfaction [40]. According to Bangor et al [38], the average SUS score of approximately 70 can be interpreted as good or acceptable [38].

Results

Heuristic Evaluation

The heuristic evaluation resulted in a total of 121 usability problems and 131 heuristic violations reported by 15 participants. The usability problems identified through heuristic evaluation are summarized and presented by place of occurrence (eg, alcohol, diet, physical activity, and smoking modules as well as the start page), number of heuristic violations, and mean severity ratings in Figure 1. The alcohol module generated the maximum number of usability problems and problem descriptions (n=49), followed by the diet module (n=39), the physical activity module (n=13), and the smoking cessation module (n=8). The start page also had usability problems (n=12). The average severity ratings ranged from 2.1 to 2.8 (on a scale of 0-4). The start page module had the highest severity rating, that is, 2.8.

The line in Figure 1 shows the mean severity ratings for usability problems on a scale of 0 to 4, where 0=none, 1=cosmetic problems, 2=minor problems, 3=major problems, and 4=usability catastrophe.

Of the 10 types of heuristic violations depicted in Table 1, categories 2 (*Match between system and the real world*) and 4 (*Consistency and standards*) dominated at 33 and 32, respectively. The heuristic categories 1 (*Visibility of system status*; n=4) and 10 (*Help and documentation*; n=3) had the lowest violations across all views.

Figure 1. Number of usability problems identified presented by place of occurrence, counts of heuristic violations identified by the 15 participants.

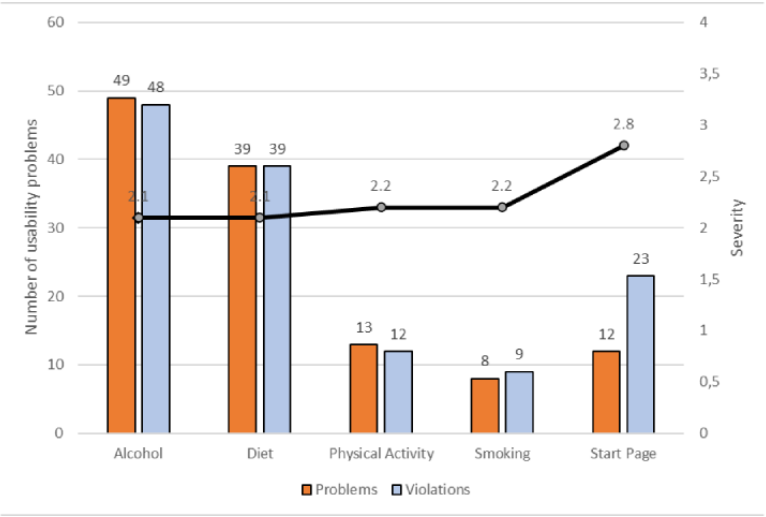


Table 1. Presentation of the frequencies of 10 heuristic violations sorted by Nielsen's heuristics reported by the 15 participants.

Violations	Number of participants
Visibility	4
Match	33
User control	6
Consistency	32
Error	8
Recognition	8
Flexibility	21
Aesthetics	10
Recover	5
Help	3

Most severity ratings were in the major severity category, indicating that fixing the problem should be given high priority (severity rating 3). The problems categorized with the highest severity rating were in the alcohol and diet modules. Most of the minor usability problems, that is, fixing those problems should be given low priority (severity rating 2), were also in the alcohol and diet modules. There were also a total of 17 problems reported as usability catastrophes across all modules, indicating the imperative to fix the problem before the product can be released (severity rating 4). Figure 2 shows severity ratings across system views.

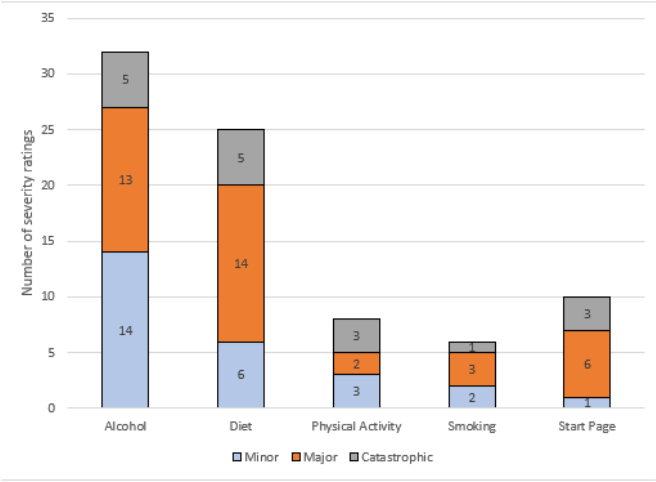
When analyzing the nature of usability problems, the heuristic evaluation revealed that major usability problems and catastrophic ratings concerned shortcomings in displaying

XSL•FO
RenderX

easy-to-understand information to the users or technical errors. Examples of these types of usability problems provided by participants are as follows:

> *Navigation unclear. There's a constant mixture of headings [in the various modules], unclear headlining, different selectable functions, sometimes with hidden text. You don't know where you are* [in the intervention]
>
> *If the diagram's x- and y-axes don't have any labels, you don't understand because the table headings are unclear and don't stand alone.*
>
> *In the figures it's unclear which direction the scales go in, you don't understand the colors, it's confusing.*
>
> *Several choices of wording, and the concepts are difficult and complicated. Think about having simple, consistent wording and simpler language.*
>
> *It's not possible to save all the choices you've entered, you can't change your choices. The planning disappears when you browse through the tabs.*
>
> *You can't go back easily. In other words, it's not possible to find your way "home" easily, there's no home button.*

Figure 2. Number of severity ratings (minor, major, and catastrophic) for each module as reported by the 15 participants.



Usability Tests

A total of 5 high school students, 3 adolescent girls and 2 adolescent boys, participated in the usability testing. The analysis of the data from the transcripts revealed 3 main categories of barriers that limited usability: (1) design, (2) content, (3) and functionality.

Design

Design refers to how the intervention is perceived visually. The data showed that the design of the prototype influenced engagement and interest among users, thereby motivating or hindering user engagement with the intervention. Design comprises both aesthetics, that is, how attractive the intervention is perceived to be, and clarity, that is, how the intervention is structured and presented. For example, participants described how clear headlines and appealing infographics were integral to usability. Moreover, participants highlighted the use of clearly presented text, figures, and tables as a way to improve the usability of the intervention. In addition, characteristics of the intervention such as the use of attractive pictures, symbols, and videos were described as further improving usability:

> *And then I wonder if you could make it a bit more fun or something /.../ to have some background color other than white.* [Adolescent 2]
>
> *Perhaps it would be good to have images? I don't think it captures the user's interest much [without images].* [Adolescent 1]

Regarding the font size, participants preferred a larger font size and suggested clearly defined headlines:

> */.../ an intervention that looks a bit half-hearted and where things aren't, like, centered or whatever. Perhaps it looks less professional than it actually is, and so you don't trust it as much /.../ because you've put so much effort into it being scientifically correct, so you could put a bit of effort into it looking good or, like, fixing it.* [Adolescent 4]
>
> */.../ and bigger, clearer headings. I think bullet points are good, easy to read /.../ I don't think you get particularly excited using it.* [Adolescent 1]

Content

The second category refers to both the quality and quantity of intervention *content*. Problems relating to quantity were described as too much text, with large blocks of text making the information difficult to process, which subsequently limits the usability:

> */.../ there's a lot of text. It gets a bit, like, too much text to manage to read through everything. /.../ you don't want there to be too much text. It should be, like, quite quick and easy.* [Adolescent 5]
>
> *You can quite easily get tired with a lot of text.* [Adolescent 3]
>
> *Sometimes less is more /.../ if it's simpler, it's easier.* [Adolescent 4]

XSL•FO
RenderX

The quality of the content referred to the terminology whereby language was perceived as too heavy, too complicated, and difficult to comprehend. Indeed, words, phrases, and concepts were perceived as unfamiliar and not tailored to the target age group. Thus, engaging with the content was too taxing and limited the usability:

> *What on earth is "moderate level" and what on earth is "strenuous level"?* [Adolescent 4]
>
> *"Dietary index"? What does "dietary index" mean?* [Adolescent 1]
>
> *Very strange words. People don't use them at all!* [Adolescent 3]

Functionality

The third category *functionality* referred to the need for the app to be effortless to navigate, to be quick to use, and to have a logical flow, according to participants who wanted an easy-to-use interface. Poor functionality was considered a major usability barrier. Participants described that usability was limited when navigation was complex and included multiple modules or functions. For example, participants stated that scrolling to find information required too much effort and the intervention needed to be easier to grasp:

> *I think fewer stages would be good.* [Adolescent 1]
>
> *Yes, because otherwise perhaps it's a bit so-so, that you go in and first you can click there, and then click there /.../ in other words, it gets too much.* [Adolescent 2]

Participants stated that time was precious and they did not want to spend time navigating unnecessarily, such as entering what they had been eating or how many activity minutes they had participated in. Thus, engagement with the app must be effective and targeted to facilitate usability:

> *You should be able to get an overview extremely quickly.* [Adolescent 4]
>
> *Like, how many times I ate fruit or berries last week? Yes, I might have eaten fruit once. I might have eaten it ten times, I might not have eaten it at all. I don't know. It was a bit difficult.* [Adolescent 5]

Participants also described the importance of a logical flow between different components of the app. In addition, the features that guided or prompted navigation of the intervention could facilitate usability:

> *You need a bit more help orientating yourself, where you are. Or some kind of main menu that comes up, and then you can tap on alcohol and after you've tapped on that subheadings appear. Then you can choose between them.* [Adolescent 4]
>
> *Maybe [it would help if] everything is in categories instead, and you tap, and maybe then it appears. Not showing everything there from the start ... instead, you can go into what you're interested in.* [Adolescent 1]

Furthermore, relating to the logical flow within the intervention, there was a desire for consistency between the different modules to improve the usability of the intervention:

> *Because maybe you can't have the same subheadings for everything, but that they're still very consistent, that there are reminders and text messages, then they should be in all [modules] so you can get familiar with it and find things.* [Adolescent 1]

The SUS Score

The results of the SUS scores are given in Table 2. As shown in this table, the analysis of the SUS score identified that the intervention was rated with an average score of 66.6. According to Bangor et al [38], an average SUS score below 70 indicates that the system has shortcomings that need to be addressed. Of the 5 participants, one rated the LIFE4YOUth app as poor, 2 as average, and 2 assessed it as good.

Table 2. Result of the System Usability Scale.

Questions	P1[a]	P2	P3	P4	P5	Average
1. I think that I would like to use this app	4	4	3	2	3	3.2
2. I found the app unnecessarily complex	1	0	2	1	1	1
3. I found the app easy to use	1	3	3	2	2	2.2
4. I think I would need support from a technical person to use this app	4	4	3	3	4	3.6
5. I found the various functions in this app were well integrated	2	3	4	2	2	2.6
6. I thought there was too much inconsistency in this app	1	2	3	2	2	2
7. I would imagine that most people would learn to use this app very quickly	0	3	4	3	4	2.8
8. I found the app very cumbersome to use	1	3	3	2	4	2.6
9. I felt very confident using the app	1	2	3	1	4	2.2
10. I needed to learn a lot of things before I could start using this app	4	4	3	4	4	3.8
SUS[b] score (sum×2.5) maximum 100	47.5	70	77.5	55	75	66.6

[a]P1-5: person 1-5.

[b]SUS: System Usability Scale.

XSL•FO
RenderX

Discussion

Principal Findings

As described in our previous protocol study for a participatory design [34], the prompt expansion of device capability presents many challenges for developers of mHealth interventions, especially when designing interventions that aim to affect multiple individual lifestyle behaviors [25]. This study investigated the usability of the LIFE4YOUth intervention targeting health behaviors among high school students. Previous studies have suggested using a combination of different usability methods to provide insights to developers about potential usability problems [41]. This paper reports data from the first stage of a formative research process [42], which included heuristic evaluation and usability testing.

The heuristic evaluation revealed that the major usability problems and the catastrophic ratings concerned shortcomings such as information display and comprehensive information, meaning that the intervention needed to speak the users' language with consistent information that appears in a natural and logical order for the users. The results from the usability testing showed that design (aesthetics and clarity), content (quality and quantity), and functionality (effortless and logic flow) enabled usability. The findings of this study are consistent with those of previous research, which found that participants wanted features that reduced the amount of time and effort required from them [43]. The amount of data participants are expected to enter into self-monitoring apps should be carefully considered in future intervention development [44] because the results showed that frustration with a large quantity of data is one of the most common complaints of users and results in apps being deleted entirely [45].

Engagement was an issue closely related to usability. For instance, participants explained that when the intervention did not appear to have a logical flow, they would quickly cease to use it. Previous research has stressed that engagement refers both to how a user interacts with technology and their emotional response to it [46]. In a study investigating usability barriers and enablers for interventions targeting harmful drinking in young adults, participants stated that positive experiences of usability made them engage more with the intervention and made them more likely to keep using it [43]. In a think-aloud study among adult smokers and drinkers, users revealed their choice of smoking cessation or alcohol reduction apps to be influenced by the apps immediate look and feel, social proof, and titles judged to be realistic and relevant [47,48]. Individuals seem more motivated to engage with and process information more thoroughly if the message is personally relevant and meaningful [49]. Theoretical models of user engagement propose that an individual's characteristics and personal circumstances may influence their user experiences of digital interventions [50]. Engagement is an ongoing issue for mHealth intervention development. Low login rates and limited use of interventions are issues consistently reported in the literature, and higher engagement through logins and repeated use is associated with better effects of the intervention [51].

According to previous research, the most important factors during the design process are flexibility and responsiveness to the input and feedback of the target audience [52]. Optimizing usability for mHealth interventions is a critical step in the development process. Young adults, the age group with the highest use of mHealth interventions, have high expectations of digital products. If an intervention does not have optimal usability, engagement will be compromised [46]. We agree with previous researchers who conclude that it is not good enough to have an evidence-based intervention per se, but rather, aspects such as delivery, design, aesthetics, usability, and functionality need to be carefully considered [43]. As there is a high demand for interventions that target skilled technology users, such as high school students, the LIFE4YOUth mHealth intervention needs to be iteratively refined and improved and will, thereafter, be pilot tested. An RCT will be conducted to determine the efficacy of the intervention at a later stage. If effective in the subsequent RCT, the program has the potential to be implemented nationally through school health services. According to the World Health Organization, education can play an important role in health promotion of the youth [53]. School multidisciplinary teams provide good accessibility for high school students, and therefore, a natural setting for attempting to endorse health behaviors for as many high school students as possible [15].

Strengths and Limitations

A limitation of the study is the small and partly nonrepresentative sample that highlights the need for caution when interpreting the results. The results cannot be used for far-reaching conclusions.

Combining heuristic evaluation and usability testing is a strength of this study. Heuristic evaluation provided insights to developers about potential usability problems, particularly in terms of identifying problems with user interface usability. The results from the heuristic evaluation were also used as inspiration to create tasks applied in the usability tests. Usability tests provided knowledge regarding whether specific tasks could be performed in the sequences of actions they were designed for to give direct input into how real users used the system.

The heuristic evaluations were performed wherever the participants preferred, for example, in the participant's home. This was done to facilitate participation and to optimize that the participants felt no time pressure. Hence, the validity of the data could not be controlled for. This study was not conducted to identify every usability problem with the mHealth intervention but instead to show how heuristic evaluation and usability testing with a small number of users could identify a large proportion of usability problems and assist in making significant improvements to an mHealth intervention targeting multiple health behaviors. The methods used were valuable in identifying not only major areas and themes that needed modification but also smaller, easily fixed problems that users encountered.

Conclusions

Through participatory design using heuristic evaluation and usability testing, this study resulted in in-depth knowledge

XSL•FO
RenderX

regarding the aspects of intervention content and structure that end users (eg, high school students) considered important. This knowledge can be used when designing an mHealth intervention targeting multiple health behaviors. In summary, heuristic evaluation showed that the major usability problems and the catastrophic ratings concerned information display and language use. Usability testing showed that design (aesthetics and clarity), content (quality and quantity), and functionality (effortless and logic flow) enabled usability. This knowledge is valuable in guiding further development of a final version of the novel multiple mHealth intervention program LIFE4YOUth.

Acknowledgments

This study was funded by a grant from the Swedish Research Council for Health, Working Life and Welfare (Forte 2018-01410, principal investigator: ML).

Conflicts of Interest

MB owns a private company that develops and distributes lifestyle interventions for use in health care settings.

Multimedia Appendix 1
Protocol used for participants to report usability problems tied to the place of occurrence, usability problem, problem description, heuristics violated, and severity rating scoring.
[DOCX File , 13 KB - mhealth_v8i10e17999_app1.docx]

References

1. Gore FM, Bloem PJ, Patton G, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10-24 years: a systematic analysis. Lancet 2011 Jun 18;377(9783):2093-2102. [doi: 10.1016/S0140-6736(11)60512-6] [Medline: 21652063]
2. Mokdad AH, Forouzanfar MH, Daoud F, Mokdad AA, El Bcheraoui C, Moradi-Lakeh M, et al. Global burden of diseases, injuries, and risk factors for young people's health during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2016 Jun 11;387(10036):2383-2401. [doi: 10.1016/S0140-6736(16)00648-6] [Medline: 27174305]
3. Meader N, King K, Wright K, Graham HM, Petticrew M, Power C, et al. Multiple risk behavior interventions: meta-analyses of RCTs. Am J Prev Med 2017 Jul;53(1):e19-e30 [FREE Full text] [doi: 10.1016/j.amepre.2017.01.032] [Medline: 28258777]
4. Schuit A, van Loon AM, Tijhuis M, Ocké M. Clustering of lifestyle risk factors in a general adult population. Prev Med 2002 Sep;35(3):219-224. [doi: 10.1006/pmed.2002.1064] [Medline: 12202063]
5. de Vries H, Kremers S, Smeets T, Reubsaet A. Clustering of diet, physical activity and smoking and a general willingness to change. Psychol Health 2008;23(3):265-278. [doi: 10.1080/14768320701349107] [Medline: 25160478]
6. Simmonds M, Llewellyn A, Owen CG, Woolacott N. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. Obes Rev 2016 Feb;17(2):95-107. [doi: 10.1111/obr.12334] [Medline: 26696565]
7. Hall WD, Patton G, Stockings E, Weier M, Lynskey M, Morley KI, et al. Why young people's substance use matters for global health. Lancet Psychiatry 2016 Mar;3(3):265-279. [doi: 10.1016/S2215-0366(16)00013-4] [Medline: 26905482]
8. Stockings E, Hall WD, Lynskey M, Morley KI, Reavley N, Strang J, et al. Prevention, early intervention, harm reduction, and treatment of substance use in young people. Lancet Psychiatry 2016 Mar;3(3):280-296. [doi: 10.1016/S2215-0366(16)00002-X] [Medline: 26905481]
9. Patton GC, Coffey C, Sawyer SM, Viner RM, Haller DM, Bose K, et al. Global patterns of mortality in young people: a systematic analysis of population health data. Lancet 2009 Sep 12;374(9693):881-892 [FREE Full text] [doi: 10.1016/S0140-6736(09)60741-8] [Medline: 19748397]
10. Catalano RF, Fagan AA, Gavin LE, Greenberg MT, Irwin CE, Ross DA, et al. Worldwide application of prevention science in adolescent health. Lancet 2012 Apr 28;379(9826):1653-1664 [FREE Full text] [doi: 10.1016/S0140-6736(12)60238-4] [Medline: 22538180]
11. Global Action Plan for the Prevention and Control of Noncommunicable Diseases-2020. World Health Organization. 2013. URL: https://apps.who.int/iris/bitstream/handle/10665/94384/9789241506236_eng.pdf?sequence=1 [accessed 2020-09-09]
12. Prochaska JJ, Spring B, Nigg CR. Multiple health behavior change research: an introduction and overview. Prev Med 2008 Mar;46(3):181-188 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.001] [Medline: 18319098]
13. Hale DR, Fitzgerald-Yau N, Viner RM. A systematic review of effective interventions for reducing multiple health risk behaviors in adolescence. Am J Public Health 2014 May;104(5):e19-e41. [doi: 10.2105/AJPH.2014.301874] [Medline: 24625172]
14. Mason M, Ola B, Zaharakis N, Zhang J. Text messaging interventions for adolescent and young adult substance use: a meta-analysis. Prev Sci 2015 Feb;16(2):181-188. [doi: 10.1007/s11121-014-0498-7] [Medline: 24930386]
15. Kraag G, Zeegers MP, Kok G, Hosman C, Abu-Saad HH. School programs targeting stress management in children and adolescents: a meta-analysis. J School Psychol 2006 Dec;44(6):449-472. [doi: 10.1016/j.jsp.2006.07.001]

XSL•FO
RenderX

16. Suffoletto B. Text message behavioral interventions: from here to where? Curr Opin Psychol 2016 Feb 1;9:16-21 [FREE Full text] [doi: 10.1016/j.copsyc.2015.09.012] [Medline: 26665157]
17. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
18. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
19. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-centered ehealth interventions for children, adolescents, and adults with sickle cell disease: systematic review. J Med Internet Res 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]
20. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
21. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: a systematic review of interventions across the cancer continuum. Psychooncology 2020 Jan;29(1):17-37. [doi: 10.1002/pon.5280] [Medline: 31692183]
22. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
23. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text messaging and mobile phone apps as interventions to improve adherence in adolescents with chronic health conditions: a systematic review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]
24. Ricci-Cabello I, Bobrow K, Islam SM, Chow CK, Maddison R, Whittaker R, et al. Examining development processes for text messaging interventions to prevent cardiovascular disease: systematic literature review. JMIR Mhealth Uhealth 2019 Mar 29;7(3):e12191 [FREE Full text] [doi: 10.2196/12191] [Medline: 30924790]
25. Maar MA, Yeates K, Toth Z, Barron M, Boesch L, Hua-Stewart D, et al. Unpacking the black box: a formative research approach to the development of theory-driven, evidence-based, and culturally safe text messages in mobile health interventions. JMIR Mhealth Uhealth 2016 Jan 22;4(1):e10 [FREE Full text] [doi: 10.2196/mhealth.4994] [Medline: 26800712]
26. Heuristic evaluation. In: Usability Inspection Methods. New York, USA: John Wiley & Sons; 1994.
27. Heuristic Evaluations and Expert Reviews. Usability. 2019. URL: https://www.usability.gov/howto-and-tools/methods/heuristic-evaluation.html [accessed 2020-09-09]
28. Georgsson M, Staggers N, Weir C. A modified user-oriented heuristic evaluation of a mobile health system for diabetes self-management support. Comput Inform Nurs 2016 Feb;34(2):77-84 [FREE Full text] [doi: 10.1097/CIN.0000000000000209] [Medline: 26657618]
29. Baumel A, Muench F. Heuristic evaluation of ehealth interventions: establishing standards that relate to the therapeutic process perspective. JMIR Ment Health 2016 Jan 13;3(1):e5 [FREE Full text] [doi: 10.2196/mental.4563] [Medline: 26764209]
30. Usability Engineering. In: Usability Engineering. San Francisco, CA: Academic Press; 1993.
31. Hinchliffe A, Mummery WK. Applying usability testing techniques to improve a health promotion website. Health Promot J Austr 2008 Apr;19(1):29-35. [doi: 10.1071/he08029] [Medline: 18481929]
32. Running a Usability Test. Usability. 2019. URL: https://www.usability.gov/how-to-andtools/methods/running-usability-tests.html [accessed 2020-09-09]
33. Ehrler F, Weinhold T, Joe J, Lovis C, Blondon K. A mobile app (bedside mobility) to support nurses' tasks at the patient's bedside: usability study. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e57 [FREE Full text] [doi: 10.2196/mhealth.9079] [Medline: 29563074]
34. Müssener U, Löf M, Bendtsen P, Bendtsen M. Using mobile devices to deliver lifestyle interventions targeting at-risk high school students: protocol for a participatory design study. JMIR Res Protoc 2020 Jan 6;9(1):e14588 [FREE Full text] [doi: 10.2196/14588] [Medline: 31904576]
35. Walker M, Takayama L, Landay JA. High-Fidelity or Low-Fidelity, Paper or Computer? Choosing Attributes when Testing Web Prototypes. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2016 Presented at: HFES'16; September 19-23, 2016; Hilton, Washington. [doi: 10.1177/154193120204600513]
36. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
37. Charters E. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. Brock Educ J 2003 Jul 1;12(2):-. [doi: 10.26522/brocked.v12i2.38]
38. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
39. Video in Qualitative Research. New York, USA: Sage Publishing; 2010.

XSL•FO
RenderX

40. John B. SUS: A quick and dirty usability scale. In: Jordan P, editor. Usability Evaluation In Industry. London, UK: Taylor & Francis Group; 2016.
41. Thyvalikakath TP, Monaco V, Thambuganipalle H, Schleyer T. Comparative study of heuristic evaluation and usability testing methods. Stud Health Technol Inform 2009;143:322-327 [FREE Full text] [Medline: 19380955]
42. Brown KG, Gerhardt MW. Formative evaluation: an integrative practice model and case study. Personnel Psychol 2002 Dec;55(4):951-983. [doi: 10.1111/j.1744-6570.2002.tb00137.x]
43. Milward J, Deluca P, Drummond C, Watson R, Dunne J, Kimergård A. Usability testing of the branch smartphone app designed to reduce harmful drinking in young adults. JMIR Mhealth Uhealth 2017 Aug 8;5(8):e109 [FREE Full text] [doi: 10.2196/mhealth.7836] [Medline: 28790022]
44. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
45. Brauer C. AppDynamics. 2014. URL: https://www.appdynamics.com/media/uploaded-files/1425406960/app-attention-span-research-report-1.pdf [accessed 2019-11-18]
46. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]
47. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
48. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
49. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 8;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]
50. Short CE, James EL, Rebar AL, Duncan MJ, Courneya KS, Plotnikoff RC, et al. Designing more engaging computer-tailored physical activity behaviour change interventions for breast cancer survivors: lessons from the iMove More for Life study. Support Care Cancer 2017 Nov;25(11):3569-3585. [doi: 10.1007/s00520-017-3786-5] [Medline: 28624949]
51. Hutton HE, Wilson LM, Apelberg BJ, Tang EA, Odelola O, Bass EB, et al. A systematic review of randomized controlled trials: web-based interventions for smoking cessation among adolescents, college students, and adults. Nicotine Tob Res 2011 Apr;13(4):227-238. [doi: 10.1093/ntr/ntq252] [Medline: 21350042]
52. Abroms LC, Whittaker R, Free C, van Alstyne J, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
53. New Horizons for Health Through Mobile Technologies. World Health Organization. 2011. URL: https://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf [accessed 2019-11-14]

Abbreviations

mHealth: mobile health
RCT: randomized controlled trial
SUS: System Usability Scale

Edited by G Eysenbach; submitted 28.01.20; peer-reviewed by S Badawy, S Dick, M Nitsch; comments to author 12.06.20; revised version received 09.09.20; accepted 13.09.20; published 29.10.20.

Please cite as:
Müssener U, Thomas K, Linderoth C, Löf M, Åsberg K, Henriksson P, Bendtsen M
Development of an Intervention Targeting Multiple Health Behaviors Among High School Students: Participatory Design Study Using Heuristic Evaluation and Usability Testing
JMIR Mhealth Uhealth 2020;8(10):e17999
URL: http://mhealth.jmir.org/2020/10/e17999/
doi:10.2196/17999
PMID:33118942

©Ulrika Müssener, Kristin Thomas, Catharina Linderoth, Marie Löf, Katarina Åsberg, Pontus Henriksson, Marcus Bendtsen. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 29.10.2020. This is an open-access article distributed

XSL•FO
RenderX

under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Archetypes of Gamification: Analysis of mHealth Apps

Manuel Schmidt-Kraepelin[1*], MSc; Philipp A Toussaint[1*], MSc; Scott Thiebes[1], MSc; Juho Hamari[2], PhD; Ali Sunyaev[1], PhD

[1]Department of Economics and Management, Karlsruhe Institute of Technology, Karlsruhe, Germany

[2]Gamification Group, Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland

[*]these authors contributed equally

Corresponding Author:
Ali Sunyaev, PhD
Department of Economics and Management
Karlsruhe Institute of Technology
Kaiserstraße 89
Karlsruhe, D-76133
Germany
Phone: 49 72160846037
Fax: 49 72160846581
Email: sunyaev@kit.edu

Abstract

Background: Nowadays, numerous health-related mobile apps implement gamification in an attempt to draw on the motivational potential of video games and thereby increase user engagement or foster certain health behaviors. However, research on effective gamification is still in its infancy and researchers increasingly recognize methodological shortcomings of existing studies. What we actually know about the phenomenon today stems from fragmented pieces of knowledge, and a variety of different perspectives. Existing research primarily draws on conceptual knowledge that is gained from research prototypes, and isolated from industry best practices. We still lack knowledge on how gamification has been successfully designed and implemented within the industry and whether certain gamification approaches have shown to be particularly suitable for certain health behaviors.

Objective: We address this lack of knowledge concerning best practices in the design and implementation of gamification for health-related mobile apps by identifying archetypes of gamification approaches that have emerged in pertinent health-related mobile apps and analyzing to what extent those gamification approaches are influenced by the underlying desired health-related outcomes.

Methods: A 3-step research approach is employed. As a first step, a database of 143 pertinent gamified health-related mobile apps from the Apple App Store and Google Play Store is set up. Second, the gamification approach of each app within the database is classified based on an established taxonomy for gamification in health-related apps. Finally, a 2-step cluster analysis is conducted in order to identify archetypes of the most dominant gamification approaches in pertinent gamified health-related mobile apps.

Results: Eight archetypes of gamification emerged from the analysis of health-related mobile apps: (1) competition and collaboration, (2) pursuing self-set goals without rewards, (3) episodical compliance tracking, (4) inherent gamification for external goals, (5) internal rewards for self-set goals, (6) continuous assistance through positive reinforcement, (7) positive and negative reinforcement without rewards, and (8) progressive gamification for health professionals. The results indicate a close relationship between the identified archetypes and the actual health behavior that is being targeted.

Conclusions: By unveiling salient best practices and discussing their relationship to targeted health behaviors, this study contributes to a more profound understanding of gamification in mobile health. The results can serve as a foundation for future research that advances the knowledge on how gamification may positively influence health behavior change and guide practitioners in the design and development of highly motivating and effective health-related mobile health apps.

(JMIR Mhealth Uhealth 2020;8(10):e19280) doi:10.2196/19280

KEYWORDS

mHealth; smartphones; mobile phones; gamification; quantified-self; exergames, persuasive technology

XSL•FO
RenderX

Introduction

Following the proliferation of smartphones and other smart devices into people's everyday lives during the last decade, mobile app stores (eg, the Apple App Store and Google Play Store) now provide users with a plethora of different health-related mobile apps (mHealth apps) [1]. Typical mHealth apps available on these stores include apps that support users in pursuing healthy diets [2], apps that motivate their users to increase their physical activity [3], or apps that help managing chronic diseases properly [4], to name but a few. Surveys show that 58.23% (934/1604) of US smartphone users have downloaded at least one such mHealth app to their smartphones in the past [5,6], making mHealth apps a multibillion Dollar business with annual growth rates of 30% or more [7]. Yet, despite mHealth apps' potential to positively influence users' health behavior and their high download numbers, research also suggests that in the past a majority of users have failed to frequently use such apps or even stopped using them after a short period, for example, due to high data entry burden, or loss of interest [5,8].

Gamification presents itself as a promising approach to overcome a loss of interest, increase user engagement [9], raise the quality of health behaviors [10], and motivate users to use mHealth apps for a sustained period [11]. It refers to the overall proliferation of games in culture, society, and technology and describes that technologies are being transformed and designed to afford positive experiences, motivational enforcement, and skill accruement [12]. Drawing on the motivational potential of (video) games to foster certain health behavior outcomes has a long tradition in health care. The development of so-called serious games (ie, "games that do not have entertainment, enjoyment or fun as their primary purpose" [13]), for example, can be traced back as far as to the 1970s [14]. Standing in the long tradition of gaming in health care, gamification has rapidly gained interest by health care researchers and professionals over the last decade. While some health behaviors, such as exercise and exercise programs in themselves are pervasively gameful in nature, mHealth apps supporting such behaviors are increasingly and more explicitly gamified further [15].

Research on gamification and its design is still in its infancy and rapidly evolving. Researchers increasingly recognize methodological shortcomings of existing studies on the effectiveness of gamification and that research on gamification has mainly advanced without an agenda, theoretical guidance, or a clear picture of the field [16,17]. What we actually know about the phenomenon today stems from fragmented pieces of knowledge, and a variety of different perspectives [16]. One of these perspectives has led to the development of a variety of frameworks and processes for designing and implementing effective gamification (see Mora et al [18] and Morschheuser et al [19] for an overview). Another perspective is concerned with providing means (eg, taxonomies) to conceptually classify existing gamification approaches (eg, [20-22]). However, these research streams usually produce research results based on research prototypes. While experiments on simple prototypes of gamification may improve the internal validity of the corpus of gamification research on how any specific gamification features affect behavior, the state of the apps of gamification in practice may differ from research prototypes. Therefore, research is needed to map the types of gamification available and improve the possible external validity as well [16,17,23]. In order to better grasp the phenomenon of gamification and its influence on peoples' health behavior through mHealth apps in practice, we currently lack knowledge about how gamification is actually implemented in real-world mHealth apps and whether certain industry best practices have emerged. From our point of view, research and practice will benefit from such knowledge for various reasons. First, in conjunction with conceptual design knowledge, such deeper knowledge on gamification approaches in pertinent, real-world mHealth apps can guide developers in designing and implementing suitable gamification approaches with regard to their targeted health behavior and aid them in unleashing the full motivational potential of gamified mHealth apps. Second, for researchers in the field of gamified mHealth apps, such knowledge can serve as an indicator that shows whether, and if so, how extant research insights into the design of effective gamification in mHealth apps have been transferred to real-world systems. Third, such research contributes to a deeper understanding of the interplay between the effectiveness of gamification and its application context because it helps to get a better picture of whether certain gamification approaches have shown to be particularly suitable for certain health behaviors in practice. Providing practitioners and researchers with such knowledge requires scrutinizing the status quo of and understanding which, if any, dominant gamification approaches are being designed and implemented in mHealth apps. We therefore ask:

(Research Question) *What are dominant gamification approaches in mHealth apps?*

In literature, several studies exist that apply gamification to mHealth apps. However, most studies investigate the psychological or behavioral effects that occur when introducing specific game design elements (eg, points, leaderboards) or a combination of different elements to a certain mHealth app [24]. The stream of research that is closest to our work has been concerned with investigating the extent to which game design elements have been implemented in real-world mHealth apps and analyzing their relationships to various other constructs such as app popularity [15], user ratings [25,26], or the use of health behavior theory constructs [15]. To the best of our knowledge, no research exists that has explicitly examined the emergence of specific dominant gamification approaches in the domain of mHealth apps and the relationship to their targeted health behaviors.

In order to answer our research question, we draw on the taxonomy of gamification approaches in mHealth apps proposed by Schmidt-Kraepelin et al [20]. The taxonomy enables us to classify gamification approaches employed in pertinent gamified mHealth apps from the Apple App Store and Google Play Store. In addition, we use the 2-step clustering approach of Punj and Stewart [27], which has been employed in similar research [28,29], for identifying archetypes of gamification approaches in pertinent mHealth apps. In doing so, we are able to unveil established best practices in the design and implementation of gamification for mHealth apps and to analyze to what extent

XSL•FO
RenderX

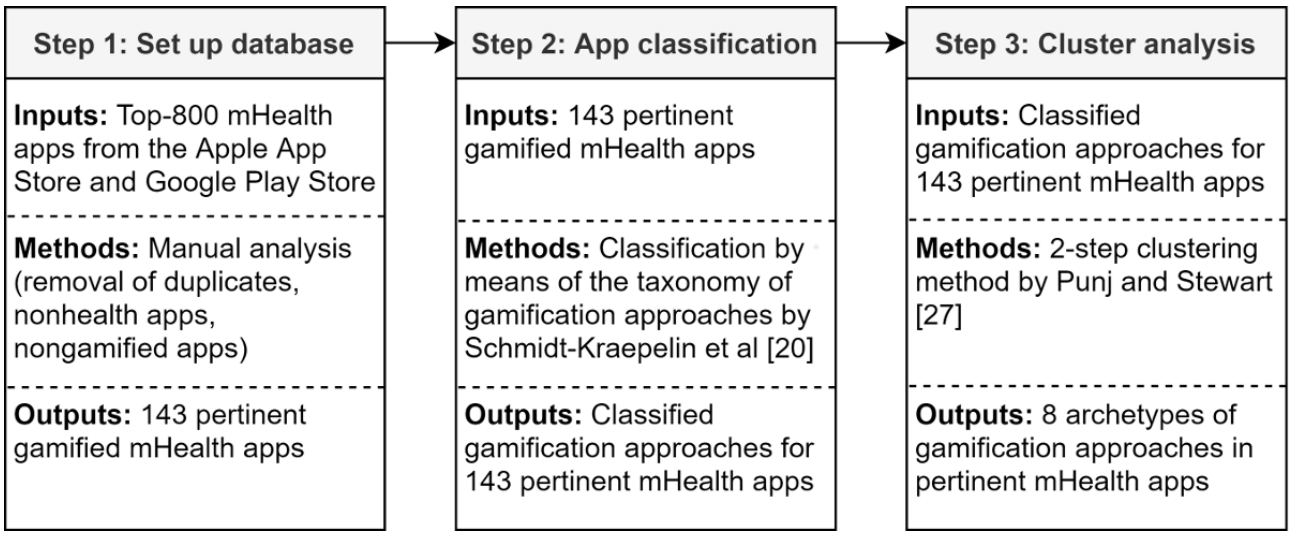
those gamification approaches are influenced by the underlying desired health-related outcomes.

Methods

Overview

To answer our research question, we employed a 3-step research approach that was informed by the studies of Remane et al [28] and Thiebes et al [29] and is shown in Figure 1. As a first step, we set up a database of pertinent gamified mHealth apps from the Apple App Store and Google Play Store. Second, we classified the gamification approaches of the gamified mHealth apps within our sample based on the taxonomy provided by Schmidt-Kraepelin et al [20]. Finally, we conducted a 2-step cluster analysis in order to identify archetypes of the most dominant gamification approaches in pertinent gamified mHealth apps.

Figure 1. Overview of the 3-step research approach.

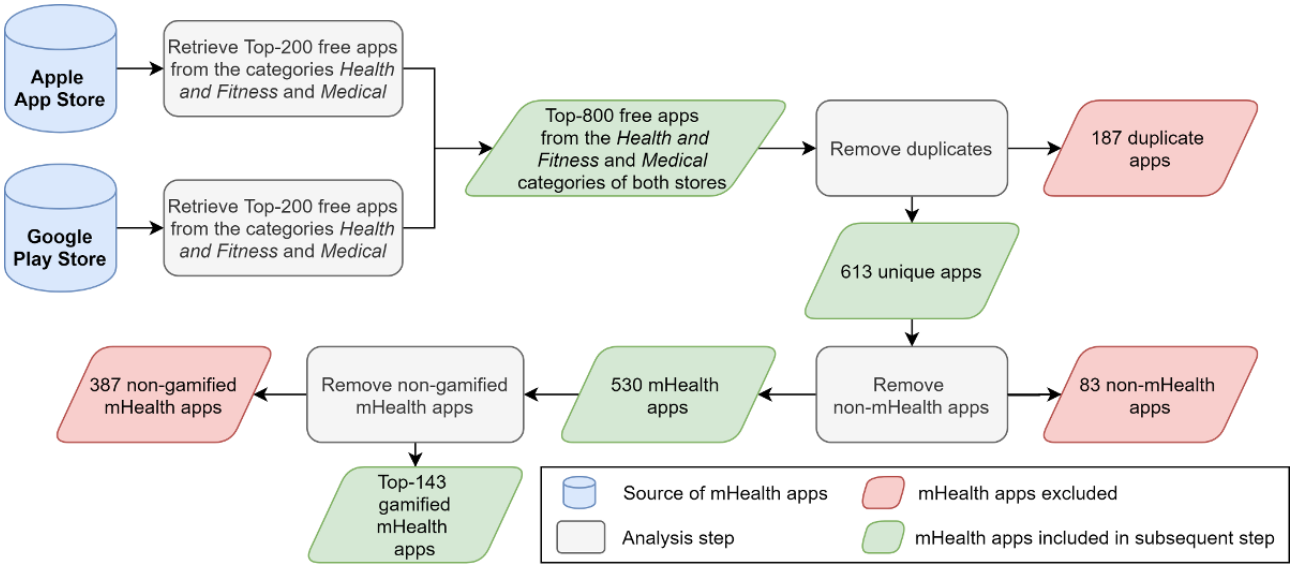


Step 1: Setting up a Database of Pertinent Gamified mHealth Apps

In the first step of this study, we set up a database of pertinent gamified mHealth apps for the 2 prevailing mobile operating systems: iOS and Android. For this purpose, we decided to sample mHealth apps from the Apple App Store and the Google Play Store because they are by far the 2 largest mobile app stores in terms of app downloads (Apple App Store: 29.6 billion in 2018; Google Play Store: 75.5 billion in 2018) and revenue (Apple App Store: US $46.6 billion in 2018; Google Play Store: US $24.8 billion in 2018) [30,31]. mHealth apps are usually found in the categories *Health and Fitness* and *Medical*, which exist on both the Apple App Store and the Google Play Store [32]. Both categories offer separate rankings, listing the top apps in terms of downloads for paid and free apps. As about 95% of apps downloaded to mobile devices are free and more than 95% of revenue in mobile app stores is generated through freemium apps (ie, apps that offer basic functionalities for free, but charge money for additional features) [33], we decided to consider only free apps within this study. Because the top rankings tend to change slightly from day to day, all rankings were recorded on January 18, 2018. To ensure a representative sample of mHealth apps, we exported the top 200 apps for both categories in both app stores. We decided to concentrate our analysis on the most popular apps since they hold the best chance of having successfully employed gamification because users tend to download and use them the most. This approach led to a set of 800 potentially relevant apps. If an app was recorded multiple times in the initial database (ie, once for the Apple App Store and once for the Google Play Store) and both versions offered identical functionality, one version of the app was excluded from further analysis. Overall, 187 duplicates were excluded. Furthermore, we examined all apps closely in order to determine (1) whether they were mHealth apps and (2) whether they were gamified (ie, made use of gamification). We excluded all apps from analysis that did not meet both requirements. To determine whether an app was an mHealth app, we followed the suggestions proposed by Stepanovic and Mettler [11]. Accordingly, an app had to either support patients in the treatment of a given medical condition (eg, diabetes), support users in pursuing healthy lifestyles (eg, weight control), or support medical professionals in the patient treatment or education. Overall, 530 apps were classified as mHealth apps. In a subsequent step, an app was considered as gamified, if it used gamification mechanics such as game-like rewards or incentives to increase motivation and sustain habits of users over time. As a reference point for gamification mechanics, the list proposed by Thiebes et al [22] was used. These game-like rewards and incentives had to support some underlying health-related functionality. For example, quizzes were only considered if their purpose was to motivate users to conduct the main health-related activity of the app. In this step, 143 of the remaining 530 apps were coded as gamified apps which constituted our final set of pertinent gamified mHealth apps. A complete list of apps included in and excluded from our database can be found in Multimedia Appendix 1, whereas an overview of the database setup process is given in Figure 2.

XSL•FO
RenderX

Figure 2. Setting up a database of pertinent gamified mHealth apps.



Step 2: Classifying Gamification Approaches in mHealth Apps

For deriving meaningful clusters of gamification approaches, we needed a classification scheme to determine their characteristics. One type of such classification schemes are taxonomies, which serve as important tools for structuring knowledge in many scientific disciplines. Extant literature has already proposed taxonomies related to gamification. However, they often do not distinguish the concepts of gaming or serious games and gamification (eg, [34]), or only categorize single game design elements but not holistic gamification approaches (eg, [22]). In addition, many classification schemes consist of a set of dimensions (eg, game design elements) without providing concrete characteristics (ie, manifestations of these dimensions). Such taxonomies often leave too much room for interpretation, which impedes classifying real-world objects. In this work, we draw on the taxonomy proposed by Schmidt-Kraepelin et al [20] as our classification scheme, because it (1) provides the opportunity of classifying holistic gamification approaches as it is not limited to single game elements, (2) has been explicitly built for health care apps, and (3) provides a concrete set of dimensions with mutually exclusive and collectively exhaustive characteristics. Further, the taxonomy has been derived based on the guidelines proposed by Nickerson et al [35], which combine inductive and deductive reasoning and have been extensively used to develop taxonomies for phenomena in health care [29]. Past research has shown that taxonomies developed based on the guidelines by Nickerson et al [35] are suitable classification schemes with regard to the development of archetypes through cluster analysis [28,29]. The taxonomy proposed by Schmidt-Kraepelin et al [20] consists of 12 dimensions, each consisting of 2 to 3 mutually exclusive characteristics, with a total of 30 characteristics. The dimensions included in the taxonomy are (1) gamification concept-to-user communication, (2) user identity, (3) rewards, (4) competition, (5) target group, (6) collaboration, (7) goal setting, (8) narrative, (9) reinforcement, (10) level of integration, (11) persuasive intent, and (12) user advancement. The complete taxonomy is shown in Table 1. A detailed description of the taxonomy can be found in Multimedia Appendix 2.

To classify the apps in our database along the taxonomy, each app was downloaded, tested, and used to experience all features. Depending on the complexity of the mHealth app, the required analysis time was usually 15 to 30 minutes. Each mHealth app was coded by one of two researchers (MS-K and PT). Prior to data analysis, the researchers were trained in the understanding of the taxonomy by the authors who originally developed the taxonomy. In order to ensure a high level of coding reliability, both researchers coded an initial set of 20 gamified mHealth apps independently and subsequently discussed their results. This initial coding and the subsequent discussion were supervised by the original authors of the taxonomy. Afterward, the two researchers coded their assigned set of mHealth apps on their own. In cases where a researcher was uncertain about a coding, the respective gamified mHealth app was discussed with the other researcher and the authors of the original taxonomy. Final classifications for all 143 mHealth apps can be found in Multimedia Appendix 3.

XSL•FO
RenderX

Table 1. Taxonomy of gamification approaches for health apps proposed by Schmidt-Kraepelin et al [20].

Dimension	Rationale	Characteristics
Gamification concept-to-user communication	How does the gamification approach communicate with the user?	• Direct • Mediated
User identity	How is the user's identity represented in the gamification approach?	• Virtual character • Self-selected
Rewards	Which rewards can users earn by playing and progressing within the gamification approach?	• Internal • Internal and external • No
Competition	How do users compete with each other within the gamification approach?	• Direct • Indirect • No
Target Group	Who is the targeted audience of the gamification approach?	• Patients • Healthy individuals • Health professionals
Collaboration	Which form of collaboration does the gamification approach offer?	• Cooperative • Supportive only • No
Goal setting	Who sets goals within the gamification approach?	• Self-set • Externally set
Narrative	How does the gamification approach behave over time?	• Continuous • Episodical
Reinforcement	How does the gamification approach attempt to reinforce its users?	• Positive • Positive–negative
Persuasive intent	Which type of health-related change does the gamification approach aim to evoke?	• Compliance change • Behavior change • Attitude change
Level of integration	To which extent is the gamification approach cohesively related to the underlying health-related activities?	• Independent • Inherent
User advancement	How does the gamification approach consider the overall user advancement?	• Presentation only • Progressive • No

Step 3: Cluster Analysis

In the third step of our methodology, we derived archetypes of gamification approaches for mHealth apps, utilizing cluster analysis. Cluster analysis is a process of finding distinct groups of objects (ie, clusters) in data for which the objects within 1 group are as similar as possible, and as dissimilar as possible from objects in the other groups based on a predetermined set of attributes [36]. Many different clustering methods exist and choosing the approach best suited for the present problem can be cumbersome and error prone. For example, the researcher has to consider what similarity or dissimilarity measure to choose and how many clusters to generate [27]. While in general, iterative partitioning algorithms, such as k-means, yield better performance than hierarchical clustering methods, they usually require defining a priori how many clusters the researcher wants to produce. To overcome the weaknesses of both approaches and increase clustering performance, we followed the 2-step approach proposed by Punj and Stewart [27]. In the first step, a hierarchical method is used to determine a preliminary solution, from which a candidate number of clusters can be deduced. In the second step, this candidate number is then used as a starting point in an iterative partitioning algorithm in order to arrange the included objects into their final cluster solution. Because the objects within this study (ie, gamification approaches) are classified through the application of a taxonomy that is similarly structured like the taxonomies of Remane et al [28] and Thiebes et al [29], we followed the clustering approaches of those studies and utilized Ward's method for step 1 and the k-means algorithm for step 2. Both steps were conducted with IBM's statistical analysis software SPSS Statistics version 25.0.

The dendrogram produced by Ward's method indicated that the 1-7-, 11-,12-, or 13-cluster solutions would all be suitable candidate numbers of clusters. Reviewing the scree plot, with

XSL•FO
RenderX

use of the elbow rule [37], the cluster solutions of size 6, 13, and 8 stood out to have the most explanatory power in this particular order (cf. Multimedia Appendix 4). Hence, we narrowed the search space to 6-, 8-, and 13-cluster solutions. Having determined the preliminary cluster solutions, we used k-means to derive our final cluster solution in the second stage. The 6- and 13-cluster solutions produced by k-means both comprised clusters with size 1 or 2, which impeded a meaningful interpretation of these solutions. In addition to the small cluster sizes, all 3 solutions showed low to no significance (ie, how relevant a certain characteristic is for the cluster solution) of the characteristics Mediated (dimension: *gamification concept-to-user communication*) and virtual character (dimension: *user identity*) in the clustering process (cf. Multimedia Appendix 5). This finding was not surprising as only 1 out of 143 and 3 out of 143 objects, respectively, showed these characteristics and therefore would only influence clustering slightly if at all. Further testing confirmed that the low relevance of these characteristics and hence their respective dimensions was present for all partitioning sizes ranging from 2 to 19 clusters.

Consequently, in order to achieve relevance of all dimensions in the clustering process and to increase the explanatory power of our results, we omitted the dimensions *gamification concept-to-user communication* and *user identity* with their respective characteristics from the clustering, thus leaving 10 dimensions and 26 characteristics. With these new limitations the dendrogram produced by Ward's method pointed toward 6, 8, 10, and 16 as preferred cluster solutions. Again, examining the scree plot with the elbow rule indicated 6-, 11-, 13-, and 8-cluster solutions to have the most explanatory power in this particular order (cf. Multimedia Appendix 6). Therefore, we selected 6, 8, 10, and 11 as our preliminary cluster solutions. In contrast to the k-means clustering of all 12 dimensions, the results for the 10 dimensions painted a much clearer picture. Moreover, relevance of characteristics was greatly improved throughout all possible cluster solutions (cf. Multimedia Appendix 7), which strengthened us in our decision to omit the 2 irrelevant dimensions from analysis. Because the 10- and 11-cluster solutions had small clusters of size 4 and smaller, we disregarded them for further analysis. Manual inspection for explanatory power showed that while we were able to find meaningful interpretations for both, 6- and 8-cluster solutions, the 8-cluster solution represented a more detailed and fine-grained picture for the landscape of gamification approaches in mHealth apps. Therefore, we selected the 8-cluster solution as the most suitable one for this study and report it below.

Results

Archetypes of Gamification Approaches for mHealth Apps

Each cluster of the 8-cluster solution contains 6-35 mHealth apps of the total 143 apps that incorporated gamification approaches from our database. Thereby, each cluster can be differentiated with regard to its most salient characteristics. Because the taxonomy development method by Nickerson et al [35] results in characteristics that are mutually exclusive and collectively exhaustive, the data can be interpreted as percentages [29]. A detailed overview of the results of the cluster analysis can be found in Multimedia Appendix 8, whereas Figure 3 illustrates the main characteristics of all clusters by the means of bar charts. For example, 24% (5/21) of apps in cluster 6 target patients, while 76% (16/21) have healthy individuals as their target group, and no apps (0%) were designed for health professionals. Based on the 8 identified clusters we derived 8 distinct archetypes of gamification approaches that are being implemented in mHealth apps. In the following, we elaborate on each archetype by highlighting its most representative characteristics, describing the most prevalent health behaviors that the archetype is used for and providing typical examples of mHealth apps for each archetype.

XSL•FO
RenderX

Figure 3. Cluster characteristics.

Archetype 1: Competition and Collaboration

The first archetype is characterized by the fact that it is the only archetype that contains both competitive and collaborative gamification elements. It reinforces health behaviors with positive and negative motivational experiences. While the underlying health activity is mostly independent from the gamification approach, the *competition* and *collaboration* elements as well as internal *rewards* aim to help users in reaching externally and self-set goals, thus evoking compliance and behavioral change. Archetype 1 exclusively targets healthy individuals. It may draw on a continuous *narrative* but often relies on an episodical style. This archetype is primarily implemented in apps such as Sweatcoin Pays You To Get Fit, Pedometer, Step Counter & Weight Loss, or Fitbit. The targeted health behavior of these apps is physical activity and fitness. Consequently, they are primarily found in the category *Health and Fitness*. A typical example that implements archetype 1 is

XSL•FO
RenderX

Nike+. This app offers daily and monthly running challenges in which users may participate and compete against each other while also providing elements of cooperative collaboration.

Archetype 2: Pursuing Self-Set Goals Without Rewards

A unique characteristic of archetype 2 is that it exclusively draws on goals that are set by the users themselves. Furthermore, archetype 2 does not offer any type of explicit *rewards* for specific health-related activities (eg, badges, vouchers) to their users. Users usually do not have the opportunity for cooperative *collaboration*, but may be motivated through supportive *collaboration* (eg, connection to social media) and various forms of *competition*, which results in possibilities for positive and negative *reinforcement*. While the *narrative* is usually episodical, *user advancement* within the archetype is only presentational (eg, through consecutive user levels), but not progressive (eg, through unlocking new content or progressive levels of difficulty). Similar to archetype 1, archetype 2 is mainly implemented in apps that target physical activity, fitness, and nutrition. Apps such as Freeletics Bodyweight, Planet Tracker, and Fooducate Healthy Weight Loss & Calorie Counter are typical examples that implement archetype 2. Another example is the app Runtastic Running & Fitness Tracker. It allows users to track runs or other fitness activities and share them with others. Users may also send a "digital cheer" through the app to motivate peers while they are physically active.

Archetype 3: Episodical Compliance Tracking

Archetype 3 typically does not draw on negative *reinforcement* or any type of social component such as *competition* or *collaboration*. An important characteristic of this archetype is that the narrative is always episodical, which means that it is clearly divided into different stages or that the user progress is partially or fully reset after a certain amount of time. Within these episodes, the *user advancement* may be presentational or progressive. Moreover, archetype 3 in some cases may also offer internal *rewards*. The main purpose of most apps that implement archetype 3 is to support healthy individuals in staying compliant to specific rules or guidelines for a certain period. Targeted health behaviors of archetype 3 are broad and include physical activity, fitness, an overall healthy lifestyle, nutrition, female health, pregnancy, meditation, mental health, and therapy adherence. Archetype 3 is implemented in apps such as BetterMe: Weight Loss Plan, Lifesum: Calorie Counter, Food & Nutrition Tracker, and Water Drink Reminder. The app Pregnancy Week by Week also implements archetype 3. It provides parents with insights into their baby's development during pregnancy and gives advice on nutrition, exercise, and medical needs of the current week.

Archetype 4: Inherent Gamification for External Goals

Archetype 4 uses gamification to help healthy individuals to meet externally set health goals. It mainly aims to foster a compliance or behavioral change within the users by drawing on internal or external *rewards*. Archetype 4 does not contain competitive or collaborative elements, but rather relies on positive *reinforcement*. It aims to improve the users' health through an episodic *narrative* of externally set goals. Archetype 4 only has a presentational *user advancement*. A unique characteristic of this archetype is that its gamification elements are inherently connected to the mHealth apps themselves (ie, the health-related activity is partially or fully embedded in the gamification approach and cannot be performed without interacting with the gamification approach). Similar to archetype 3, the targeted health behaviors of apps that implement archetype 4 are rather broad. They include physical activity, fitness, meditation, mental health, health navigation (eg, directories of physicians), and therapy adherence. Typical example apps are Ada – Your Health Companion, 30 Day Fit Challenge Workout, and Map My Fitness Workout Trainer. Another example is Moodpath. This app provides users with a mood diary, which helps them to better understand their feelings and thoughts with episodical interactive sessions on topics such as depressions, sadness, and anxiety.

Archetype 5: Internal Rewards for Self-Set Goals

Archetype 5 utilizes gamification in an attempt to evoke a change in the users' behavior or attitude. This is achieved by omitting elements of *competition*, but providing users with various forms of *collaboration*. A unique characteristic of this archetype is that it solely draws on internal *rewards* (eg, points, badges). Users set goals themselves and may experience positive and negative *reinforcement* as a result of their decisions. Many apps that implement archetype 5 focus on supporting meditation and mental health. However, it is also used to support female health, pregnancy, and therapy adherence. Typical examples are apps such as Smoke Free - Quit Smoking Now, and Aura: Calm Anxiety & Stress Chat. The app Headspace: Guided Meditation offers a range of meditation programs revolving around different life aspects such as Foundation, Sport, Health, Relationships, or Performance. Within each session the user may learn new meditation techniques and strategies to improve their mental well-being.

Archetype 6: Continuous Assistance Through Positive Reinforcement

Archetype 6 targets healthy individuals and patients that require continuous assistance with a certain health issue. It offers no *rewards* and does not provide progressive *user advancement*. In addition, it does not incorporate any competitive elements and *reinforcement* is only positive. While this archetype is not limited to a specific *persuasive intent* or *goal setting*, the *level of integration* is primarily independent. Overall, archetype 6 is in many aspects similar to archetype 3. However, the main difference is that archetype 6 draws on a continuous *narrative* in order to meet users' needs for continuous assistance. The targeted health behaviors of apps that implement archetype 6 include physical activity, fitness, nutrition, female health, pregnancy, and therapy adherence. Typical examples of such apps are mySugr: the blood sugar tracker made just for you, DreamMapper, and Kindara: Fertility Tracker. The app Pill Reminder supports users in tracking their medication with the help of notifications and a cooperative function that grants caretakers insight into the medication adherence of the patient.

XSL•FO
RenderX

Archetype 7: Positive and Negative Reinforcement Without Rewards

In contrast to the archetypes described so far, archetype 7 solely targets health professionals. It does not provide any *competition*, *collaboration*, or *rewards* and goals are set externally. While archetype 7 implies positive and negative *reinforcement* strategies, *user advancement* is only presentational and not progressive. Archetype 7 is only implemented in mHealth apps that support medical education of health professionals such as Teach Me Anatomy, NCLEX-RN Pocket Prep, or EMT-B Pocket Prep. Another typical example is the app ATI TEAS Pocket Prep. This app aims to provide users with necessary knowledge and skills through a simple knowledge database and a classical quiz structure.

Archetype 8: Progressive Gamification for Health Professionals

Archetype 8 is similar to archetype 7 in the sense that it solely targets health professionals. However, these 2 archetypes also differ to some extent. The most prominent difference is that archetype 8 primarily draws on a progressive *user advancement*, which means that either new content may be progressively unlocked or levels of difficulty may increase while users deepen their knowledge and skills. In contrast to archetype 7, archetype 8 is also utilized as a means to change the attitude of users. Similar to archetype 7, archetype 8 is solely implemented in mHealth apps that support medical education of health professionals. Apps that implement this archetype include Airway Ex, Prognosis: Your Diagnosis, or ATI TEAS 6 Practice Test. The app Touch Surgery is another typical example. It lets students compete against each other indirectly by ranking how well they performed a surgery case within the mHealth app and gives them the possibility to communicate and rate surgery cases of each other.

Discussion

Principal Results

Overall, the findings of our study help to better grasp the phenomenon of gamification and its influence on peoples' health behavior through mHealth apps in real-world systems. Analysis of our derived archetypes reveals some interesting insights into the current landscape of gamification approaches that are being utilized in pertinent mHealth apps. Our results paint a heterogeneous landscape of different gamification approaches for mHealth apps and help to explain the relationship between industry best practices for gamification and targeted health behaviors. The results also indicate that gamification approaches in real-world mHealth apps differ in some aspects from gamification approaches that are deployed in research prototypes, which underlines the value of consistently contrasting conceptual knowledge against real-world observations. The main findings of this study are discussed in the following.

First, our results give insight into the variety of different gamification approaches and paint a heterogenous landscape of different gamification approaches for mHealth apps. While some archetypes seem closer to classical notions of gamification in health (eg, archetype 1), other approaches have received substantially less attention in extant research. Overall, our results highlight the richness of opportunities and different perspectives that gamification provides when it comes to augmenting mHealth apps with motivational affordances. Gamification cannot only be applied with regard to a lot of different health-related outcomes, but it can also come in various forms and shapes.

Second, our results also indicate a close relationship between the identified archetypes of gamification approaches in mHealth apps and the actual health behavior change that is being targeted by such mHealth apps. We further analyzed this finding by investigating the relationship between the most significant archetype characteristics and the targeted health behavior. The results of this additional analysis, in general, support our findings (cf. Multimedia Appendices 9 and 10). It is noticeable that certain archetypes of gamification approaches are in practice primarily used for certain health behavior changes. For developers of gamified mHealth apps these insights provide valuable points of reference for implementing gamification that targets to support a specific health behavior. For example, when it comes to getting people to live a more active life and increase their physical activity, our results indicate that archetype 1 and archetype 2 are implemented frequently. These gamification approaches are characterized by a high degree of competition, which could be inappropriate in other, more serious, health contexts and could even lead to negative effects (eg, demoralization of users due to overemphasizing of peer pressure) [38]. The two archetypes differ in particular with regard to the dimension of *goal setting* (ie, whether goals are set by the user or determined externally) and *collaboration* (ie, whether collaboration is cooperative, supportive only, or no possibility for collaboration is implemented). A completely different form of gamification approaches, however, is used in apps that are intended to support future health professionals in their training. These typically implement archetype 7 and archetype 8. They are characterized by the fact that the goals to be achieved are set externally and the *narrative* is episodic. However, both archetypes are also fundamentally different in some characteristics. For example, while archetype 7 does not offer any type of *rewards* and draws only on presentational *user advancement*, archetype 8 aims to motivate their users with internal *reward*s and adjust difficulty levels in a progressive *user advancement*. In contrast to the previously discussed gamification approaches for fitness apps, archetype 7 and archetype 8 do not provide any form of *competition*. This could be an expression of the fact that the use of *competition* in learning environments is often seen as problematic and controversial in research [39]. On the one hand, researchers report that *competition* is used in classrooms to draw the attention of learners and motivate learning [39,40]. On the other hand, researchers have often raised concerns that incautious implementations of *competition* in educational settings may create anxiety and impede performance [41]. Also, other clusters indicate a close relationship between the targeted health behavior and the implemented gamification approach. For example, when it comes to support mental health activities (archetype 5), developers of gamified mHealth apps primarily avoid using gamification that potentially puts additional pressure on users

XSL•FO
RenderX

(eg, competition, externally set goals) and instead aim to foster positive experiences by allowing self-set goals and *collaboratio*n. For health behaviors that require continuous assistance (archetype 6: eg, chronic disease management, medication adherence), especially continuous narratives seem to be a popular approach to meet users' motivational needs. By contrast, developers of mHealth apps that implement this archetype primarily do not use elements of *competition* or any type of *rewards*. Contrasting the previously discussed archetypes, the range of targeted health behaviors for archetype 3 and archetype 4 is substantially broader. Instead of being tightly related to a specific health behavior, these archetypes seem to be applicable for various different health behaviors. For example, archetype 3 uses episodic narratives in order to motivate users to stay compliant with a temporary health behavior (eg, following pregnancy guidelines). By contrast, apps that implement archetype 4 draw on gamification approaches that are inherently related to the targeted health behavior in order to motivate users to achieve externally set goals (eg, completing a specific 30 days fitness challenge).

Third, the manual analysis and interpretation of our 8-cluster solution indicates that archetypes of gamification approaches for mHealth apps seem to be less clearly definable in comparison to other use cases of this method (in this regard, we particularly refer to the studies by Thiebes et al [29] and Remane et al [28]). This observation is also supported by statistical indicators. For example, in our study of gamification approaches for mHealth apps the analysis of the dendrogram resulted in considerably more cluster sizes as suitable candidates than the dendrogram in the study by Thiebes et al [29]. A similar picture was drawn when applying the elbow rule (ie, comparing the slopes of cluster size candidates within the scree plots), where the 6-cluster solution was clearly identified as the dominating solution by Thiebes et al [29], while we were left with multiple possible solutions that performed relatively equal with regard to statistical measures. From our point of view, there are different explanations for this observation. For example, it might be the case that our taxonomy of gamification approaches allows more meaningful combinations of characteristics than the taxonomies for business models proposed by Thiebes et al [29] and Remane et al [28], which leads to a more heterogeneous landscape and less clearly separable archetypes. Furthermore, it should also be noted that the development and provision of mHealth apps in most cases require significantly less capital and economic risk taking than direct-to-consumer genetic testing services [29] or carsharing services [28]. As a consequence, providers of gamified mHealth apps may simply be more willing to experiment around with their gamification approaches and do not necessarily have to draw on best practices. Lastly, the concept of gamification is still rather new and has just recently become increasingly popular for mHealth apps [24]. Thus, best practices for gamification approaches in mHealth apps might still be in an emerging phase and become clearer in the future.

Fourth, when comparing our classification results with the classification results that emerged during the taxonomy development process [20], large differences between theoretical considerations on gamification in mHealth apps and their implementation in practice became apparent. These differences become particularly clear in the dimensions *gamification concept-to-user communication* and *user identity*. During the taxonomy development process, in which exclusively 27 mHealth apps presented in research papers (primarily research prototypes) were examined, 5 of 27 apps were classified with a mediated *gamification concept-to-user communication* and 6 of 27 apps were classified with a virtual character as the characteristic for *user identity*. By contrast, our classification of a total of 143 pertinent mHealth apps offered in the Apple App Store and Google Play Store showed only 3 apps with a mediated *gamification concept-to-user-communication* and only 1 app with a virtual character as form of *user identity*. In the case of virtual characters, this may be explained to some extent by the consideration of user preferences by developers of mHealth apps. A recent study showed that avatars (the most common form of virtual characters in gamification approaches) are among those game design elements that are least preferred by users of mHealth apps [42]. However, existing research has shown that the successful use of avatars in gamification approaches is complex and requires a profound consideration of individual user needs [43]. Only by incorporating these individual needs and avoiding one-size-fits-all solutions, a desired level of emotional attachment and extraneous cognitive load may be achieved [43]. Our results could be interpreted as an indicator that most developers of mHealth apps are not willing to put this effort into the development of virtual characters, especially because users seem to attach less importance to their inclusion in comparison to other game design elements.

Implications

Our study yields several implications for research and practice. For research, our work contributes to a more comprehensive understanding of mHealth apps. We utilized a systematic classification of gamification approaches in order to develop archetypes that have emerged in pertinent mHealth apps. In doing so, we strengthen the previously formulated notion [28,29] that combining classifications based on the guidelines by Nickerson et al [35] with cluster analysis methods such as Ward's method and k-means is a promising approach to uncover archetypes in emerging contexts. Furthermore, our research demonstrates that the taxonomy proposed by Schmidt-Kraepelin et al [20] is applicable not only for research prototypes but also for real-world systems. However, our results also indicate that some dimensions and characteristics (eg, direct *gamification concept-to-user communication* and virtual characters as *user identity*) seem to be much less relevant in practice than previously assumed. Compared with other archetypes of mHealth apps that have been proposed in the literature [44], the archetypes presented here focus on a rather narrow aspect of mHealth apps (ie, their use of gamification for motivational purposes). This allowed us to clearly state how gamification has predominantly been instantiated in mHealth apps. In particular, our results indicate that specific best practices for gamification approaches have emerged for certain targeted health behavior changes. Hereby, our research contributes to answering the call for more research that helps to select suitable persuasive architectures for different mHealth apps [9].

XSL•FO
RenderX

For practice, our research yields important implications for providers and developers of mHealth apps, who want to utilize gamification in order to motivate desired health behaviors. The presented archetypes provide these stakeholders with blueprints of potential gamification approaches that have been utilized in pertinent mHealth apps. In particular, when regarding the close connection between gamification approach archetypes and underlying health behaviors (eg, physical activity, education of health professionals, mental health guidance), such blueprints can be used as guidance for designing suitable gamification approaches. Although our work does not account for the efficiency of these archetypes with regard to motivational or health behavior outcomes, it does provide practitioners with best practices that have been established among the most popular mHealth apps. Our results thereby are in line with the most prominent gamification design frameworks that emphasize the importance of specifying the targeted (health) behavior before designing a gamification approach [18]. We would also like to stress that our provided blueprints cannot substitute following established design frameworks. Rather, they should be used as an additional means to triangulate the results of following such design frameworks. Overall, our work contributes to a better understanding of how gamification is being applied in real-world mHealth apps. Such knowledge becomes increasingly important given that more and more apps draw on motivational techniques in order to achieve a desired level of user motivation and stand out from the mass of available mHealth apps.

Limitations and Future Research

The findings of this study should be interpreted in consideration of some key limitations. First, gamification is a relatively young and constantly evolving phenomenon that just recently started to draw increasing attention by researchers and practitioners alike. Disregarding this aspect, our archetypes represent only a snapshot of the current landscape of gamification approaches in pertinent mHealth apps. It is likely that in the near future, new insights from research and practice may guide the development of innovative and more effective gamification approaches (eg, through stronger consideration of individual and contextual factors that determine the effectiveness of gamification). Future research may answer this limitation by closely examining the evolution of knowledge on effective gamification approaches in mHealth apps and how they are designed. Second, we limited our analysis to free mHealth apps on the Apple App Store and Google Play Store which might have led to the exclusion of relevant paid apps or mHealth apps from other app stores. However, it would have been complicated and costly to deal with the different revenue and pricing models of freemium, premium, and others. With an industry trend toward free mobile apps [15], 95% of downloaded apps being free, and more than 95% of revenue in mobile app stores being generated through freemium business models [33], we are confident that this sample inherits the majority of gamification approaches used in mHealth apps today. Furthermore, we think that the categorical decision of only including free apps from the Apple App Store and Google Play Store ensures the reproducibility of the results. Nonetheless, investigating paid apps or alternative app stores could potentially reveal further insights. Third, the sample of 143 gamification approaches is notably smaller than that of similar studies in other contexts (eg, the 277 direct-to-consumer genetic testing services investigated by Thiebes et al [29]). This may have led to some clusters being underrepresented in our analysis. In addition, we focus our analysis on the most popular mHealth apps only, because our goal was to identify the best practices of gamification approaches. However, this certainly limits the explanatory power of our results as mHealth apps that have a very small and specific target group (eg, people with very specific and rare diseases) and still implement valuable gamification approaches may be underrepresented in our sample. It may be an interesting avenue for future research to assess whether less popular apps utilize gamification in a different way. In this regard, additional measures for mHealth app popularity (eg, user rating or number of users) could be taken into account in order to include more niche mHealth apps in the analysis. Such research could additionally contribute to the upcoming research stream that aims to assess the potential effect that gamification may have on mHealth app success [25]. Fourth, the coding of each mHealth app by only 1 researcher might have led to false classification in terms of the gamification approaches, which then could have impacted the cluster analysis. However, both researchers were trained prior to the coding to have an identical understanding of the gamification approaches confined within the taxonomy in order to minimize human error in this process. Furthermore, with larger groups of objects the cluster analysis tends to be less sensitive toward outliers, which decreases the effects of false coding. Additionally, our approach included 2 different clustering methods (ie, Ward's Method and k-means) to further minimize the risk of creating inaccurate cluster solutions due to outliers. Finally, it must be noted that we omitted the 2 dimensions *gamification concept-to-user communication* and *user identity* when conducting the cluster analysis. However, we are confident that this was the best possible approach to achieve meaningful results because omitting both dimensions during the analysis resulted in substantially higher explanatory power of the clusters and only a negligible number of objects (of the 143 apps analyzed, only 1 showed the characteristic virtual character in the dimension *user identity* and only 3 showed the characteristic mediated in the dimension gamification *concept-to-user communication* showed the inferior characteristics.

Conclusions

Gamification is a relatively young and constantly evolving phenomenon that is increasingly being utilized in the health care sector and thus becomes more and more important for researchers, health professionals, and providers of digital health services. In this study, we propose 8 rigorously developed archetypes of gamification approaches that illustrate how gamification is being implemented in mHealth apps and how their design is determined by the targeted health behavior. In doing so, we unveil salient best practices, and thereby contribute to a more profound understanding of gamification in mHealth apps. Our results can serve as a foundation for future research that advances our knowledge on how gamification may positively influence health behavior change and guide practitioners in the design and development of highly motivating and effective mHealth apps.

XSL•FO
RenderX

Acknowledgments

The authors acknowledge the support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of included and excluded mHealth apps.
[XLSX File (Microsoft Excel File), 60 KB - mhealth_v8i10e19280_app1.xlsx]

Multimedia Appendix 2
Detailed description of taxonomy proposed by Schmidt-Kraepelin et al [20].
[DOCX File , 28 KB - mhealth_v8i10e19280_app2.docx]

Multimedia Appendix 3
Taxonomy coding overview.
[XLSX File (Microsoft Excel File), 29 KB - mhealth_v8i10e19280_app3.xlsx]

Multimedia Appendix 4
Cluster analysis details Ward's Method 12 dimensions.
[DOCX File , 55 KB - mhealth_v8i10e19280_app4.docx]

Multimedia Appendix 5
Cluster analysis details k-means method 12 dimensions.
[DOCX File , 44 KB - mhealth_v8i10e19280_app5.docx]

Multimedia Appendix 6
Cluster analysis details Ward's Method 10 dimensions.
[DOCX File , 56 KB - mhealth_v8i10e19280_app6.docx]

Multimedia Appendix 7
Cluster analysis details k-means method 10 dimensions.
[DOCX File , 44 KB - mhealth_v8i10e19280_app7.docx]

Multimedia Appendix 8
Overview of clusters.
[DOCX File , 114 KB - mhealth_v8i10e19280_app8.docx]

Multimedia Appendix 9
Relationship between archetype characteristics and targeted health behavior.
[DOCX File , 81 KB - mhealth_v8i10e19280_app9.docx]

Multimedia Appendix 10
Targeted health behaviors per archetypes.
[XLSX File (Microsoft Excel File), 110 KB - mhealth_v8i10e19280_app10.xlsx]

References

1. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. Journal of the American Medical Informatics Association 2015 Apr 21;22(e1):e28-e33. [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]
2. West JH, Belvedere LM, Andreasen R, Frandsen C, Hall PC, Crookston BT. Controlling Your "App"etite: How Diet and Nutrition-Related Mobile Apps Lead to Behavior Change. JMIR mHealth uHealth 2017 Jul 10;5(7):e95 [FREE Full text] [doi: 10.2196/mhealth.7410] [Medline: 28694241]

XSL•FO
RenderX

3. Knight E, Stuckey MI, Prapavessis H, Petrella RJ. Public Health Guidelines for Physical Activity: Is There an App for That? A Review of Android and Apple App Stores. JMIR mHealth uHealth 2015;3(2):e43 [FREE Full text] [doi: 10.2196/mhealth.4003] [Medline: 25998158]
4. 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]
5. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR mHealth uHealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
6. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health App Use Among US Mobile Phone Users: Analysis of Trends by Chronic Disease Status. JMIR mHealth uHealth 2017 Dec 19;5(12):e197 [FREE Full text] [doi: 10.2196/mhealth.7832] [Medline: 29258981]
7. State of Mobile - 2020. App Annie. URL: https://www.appannie.com/download/de/go/state-of-mobile-2020/2001_State_of_Mobile_2020_Main_EN.pdf [accessed 2020-09-23]
8. Yuan S, Ma W, Kanthawala S, Peng W. Keep Using My Health Apps: Discover Users' Perception of Health and Fitness Apps with the UTAUT2 Model. Telemed J E Health 2015 Sep;21(9):735-741. [doi: 10.1089/tmj.2014.0148] [Medline: 25919238]
9. Cugelman B. Gamification: What It Is and Why It Matters to Digital Health Behavior Change Developers. JMIR Serious Games 2013;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
10. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: A systematic analysis of gamified health behavior change support systems literature. Int J Med Inform 2016 Dec;96:62-70. [doi: 10.1016/j.ijmedinf.2016.02.006] [Medline: 26944611]
11. Stepanovic S, Mettler T. Gamification applied for health promotion: does it really foster long-term engagement? A scoping review. Atlanta: AIS; 2018 Presented at: European Conference on Information Systems; Jun 23-28; Portsmouth, UK p. 23-28.
12. Hamari J. Gamification. In: Ritzer GR, editor. The Blackwell Encyclopedia of Sociology. New York, NY, USA: John Wiley & Sons; 2019.
13. Michael D, Chen S. Serious games: Games that educate, train,inform. Independence, KY: Course Technology PTR; 2005.
14. Ricciardi F, De Paolis LT. A Comprehensive Review of Serious Games in Health Professions. International Journal of Computer Games Technology 2014;2014:9. [doi: 10.1155/2014/787968]
15. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a Fad? Gamification in Health and Fitness Apps. JMIR Serious Games 2014;2(2):e9 [FREE Full text] [doi: 10.2196/games.3413] [Medline: 25654660]
16. Koivisto J, Hamari J. The rise of motivational information systems: A review of gamification research. International Journal of Information Management 2019 Apr;45:191-210. [doi: 10.1016/j.ijinfomgt.2018.10.013]
17. Nacke LE, Deterding S. The maturing of gamification research. Computers in Human Behavior 2017 Jun;71:450-454. [doi: 10.1016/j.chb.2016.11.062]
18. Mora A, Riera D, González C, Arnedo-Moreno J. Gamification: a systematic review of design frameworks. J Comput High Educ 2017 May 23;29(3):516-548. [doi: 10.1007/s12528-017-9150-4]
19. Morschheuser B, Hassan L, Werder K, Hamari J. How to design gamification? A method for engineering gamified software. Information and Software Technology 2018 Mar;95:219-237. [doi: 10.1016/j.infsof.2017.10.015]
20. Schmidt-Kraepelin M, Thiebes S, Tran M, Sunyaev A. What?s in the Game? Developing a Taxonomy of Gamification Concepts for Health Apps. : University of Hawai'i at Manoa; 2018 Presented at: Hawaii International Conference on System Sciences; Jan 3-6; Waikoloa, HI, USA p. 1217-1226. [doi: 10.24251/hicss.2018.150]
21. Liu D, Santhanam R, Webster J. Toward Meaningful Engagement: A Framework for Design and Research of Gamified Information Systems. MIS Quarterly 2017 Apr 4;41(4):1011-1034. [doi: 10.25300/misq/2017/41.4.01]
22. Thiebes S, Lins S, Basten D. Gamifying Information Systems-A Synthesis of Gamification Mechanics and Dynamics. 2014 Presented at: European Conference on Information Systems; Jun 9-11; Tel Aviv, Israel p. 1-17.
23. Rapp A, Hopfgartner F, Hamari J, Linehan C, Cena F. Strengthening gamification studies: Current trends and future opportunities of gamification research. International Journal of Human-Computer Studies 2019 Jul;127:1-6. [doi: 10.1016/j.ijhcs.2018.11.007]
24. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: A systematic review of the literature. Internet Interventions 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
25. Schmidt-Kraepelin M, Thiebes S, Sunyaev A. Investigating the Relationship between User Ratings and Gamification – A Review of Mhealth Apps in the Apple App Store and Google Play Store. : University of Hawai'i at Manoa; 2019 Presented at: Hawaii International Conference on System Sciences; Jan 8-11; Wailea, HI, USA. [doi: 10.24251/hicss.2019.181]
26. Mendiola MF, Kalnicki M, Lindenauer S. Valuable Features in Mobile Health Apps for Patients and Consumers: Content Analysis of Apps and User Ratings. JMIR mHealth uHealth 2015 May 13;3(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]

XSL•FO
RenderX

27. Punj G, Stewart DW. Cluster Analysis in Marketing Research: Review and Suggestions for Application. Journal of Marketing Research 1983 May;20(2):134-148. [doi: 10.2307/3151680]
28. Remane G, Nickerson R, Hanelt A, Tesch J, Kolbe L. A taxonomy of carsharing business models. New York City: AIS; 2016 Presented at: International Conference on Information Systems; Dec 11-14; Dublin, Ireland.
29. Thiebes S, Toussaint PA, Ju J, Ahn J, Lyytinen K, Sunyaev A. Valuable Genomes: Taxonomy and Archetypes of Business Models in Direct-to-Consumer Genetic Testing. J Med Internet Res 2020 Jan 21;22(1):e14890 [FREE Full text] [doi: 10.2196/14890] [Medline: 31961329]
30. Iqbal M. App Download and Usage Statistics (2019). Business of Apps. 2019. URL: https://www.businessofapps.com/data/app-statistics/ [accessed 2020-09-23]
31. Iqbal M. App Revenue Statistics (2019). Business of Apps. 2019. URL: https://www.businessofapps.com/data/app-revenues/ [accessed 2020-09-23]
32. Xu W, Liu Y. mHealthApps: A Repository and Database of Mobile Health Apps. JMIR mHealth uHealth 2015;3(1):e28 [FREE Full text] [doi: 10.2196/mhealth.4026] [Medline: 25786060]
33. Kearl M. 30 Essential Stats On In-App Purchases And Monetization: $37 billion by 2017. Braze. 2016. URL: https://www.braze.com/blog/in-app-purchase-stats/ [accessed 2020-09-23]
34. McCallum S. Gamification and serious games for personalized health. Stud Health Technol Inform 2012;177:85-96. [Medline: 22942036]
35. Nickerson RC, Varshney U, Muntermann J. A method for taxonomy development and its application in information systems. Eur J Inf Syst 2012 Jun 19;22(3):336-359. [doi: 10.1057/ejis.2012.26]
36. Kaufman L, Rousseeuw P. Finding groups in data: an introduction to cluster analysis. 3rd edition ed. Hoboken, NJ: John Wiley & Sons; 2009.
37. Yim O, Ramdeen KT. Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. The Quantitative Methods for Psychology 2015 Feb 01;11(1):8-21. [doi: 10.20982/tqmp.11.1.p008]
38. Schmidt-Kraepelin M, Thiebes S, Stepanovic S, Mettler T, Sunyaev A. Gamification in Health Behavior Change Support Systems – A Synthesis of Unintended Side Effects. Siegen: University of Siegen; 2019 Presented at: Wirtschaftsinformatik; Feb 23-27; Siegen, Germany.
39. Santhanam R, Liu D, Shen WM. Research Note—Gamification of Technology-Mediated Training: Not All Competitions Are the Same. Information Systems Research 2016 Jun;27(2):453-465. [doi: 10.1287/isre.2016.0630]
40. Hanus MD, Fox J. Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. 2015 Jan;80:152-161. [doi: 10.1016/j.compedu.2014.08.019]
41. Kohn A. No contest: The case against competition. New York, NY: Mariner Books; 1992.
42. Schmidt-Kraepelin M, Thiebes S, Schöbel S, Sunyaev A. Users' Game Design Element Preferences in Health Behavior Change Support Systems for Physical Activity: A Best-Worst-Scaling Approach. New York, NY: AIS; 2019 Presented at: International Conference on Information Systems; Dec 15-19; Munich, Germany.
43. Schöbel S, Janson A, Mishra A. A Configurational View on Avatar Design – The Role of Emotional Attachment, Satisfaction, and Cognitive Load in Digital Learning. New York, NY: AIS; 2019 Presented at: International Conference on Information Systems; Dec 15-19; Munich, Germany. [doi: 10.2139/ssrn.3524079]
44. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014;5(3):229 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]

Edited by G Eysenbach; submitted 11.04.20; peer-reviewed by F Gorman, R Robbins, R Hervás; comments to author 12.06.20; revised version received 03.07.20; accepted 24.07.20; published 19.10.20.

Please cite as:
Schmidt-Kraepelin M, Toussaint PA, Thiebes S, Hamari J, Sunyaev A
Archetypes of Gamification: Analysis of mHealth Apps
JMIR Mhealth Uhealth 2020;8(10):e19280
URL: https://mhealth.jmir.org/2020/10/e19280
doi:10.2196/19280
PMID:33074155

©Manuel Schmidt-Kraepelin, Philipp A. Toussaint, Scott Thiebes, Juho Hamari, Ali Sunyaev. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 19.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited.

XSL•FO
RenderX

The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Weight Management Apps in Saudi Arabia: Evaluation of Features and Quality

Dalal M Alshathri[1], MSc; Abeer S Alhumaimeedy[2], PhD; Ghada Al-Hudhud[3], PhD; Aseel Alsaleh[4], PhD; Sara Al-Musharaf[5], PhD; Ghadeer S Aljuraiban[5], PhD

[1]Department of Nutrition, Security Forces Hospital Program, Riyadh, Saudi Arabia

[2]Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

[3]Department of Information Technology, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

[4]Family and Community Medicine Department, College of Medicine, Arabian Gulf University, Manama, Bahrain

[5]Department of Community Health Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia

Corresponding Author:
Ghadeer S Aljuraiban, PhD
Department of Community Health Sciences
College of Applied Medical Sciences
King Saud University
Turki Alawwal
Riyadh
Saudi Arabia
Phone: 966 50 301 2345
Email: galjuraiban@ksu.edu.sa

Abstract

Background: Weight management apps may provide support and management options for individuals with overweight and obesity. Research on the quality of weight management mHealth apps among the Saudi population is insufficient despite frequent use.

Objective: The aims of this study were to explore user perceptions of weight management apps, explore reasons for starting and stopping app use, appraise the quality of weight management apps available in the App Store, and compare the features currently available within the app market and those that are most desirable to weight management app users.

Methods: A web-based survey consisted of 31 open and closed questions about sociodemographic information, general health questions, app use, app user perceptions, and discontinuation of app use. The quality of the weight management apps available on the App Store was assessed using the Mobile App Rating Scale and evidence-based strategies. We also used six sigma evaluations to ensure that the quality measured by the tools consistently meets customer expectations.

Results: Data from the survey were analyzed. Of the respondents, 30.17% (324/1074) had used a weight management app, 18.16% (195/1074) used the apps and stopped, and 51.68% (555/1074) had never used a weight management app. Of apps mentioned, 23 met the inclusion criteria. The overall average Mobile App Rating Scale quality of apps was acceptable; 30% (7/23) received a quality mean score of 4 or higher (out of 5), and 30% (7/23) did not meet the acceptability score of 3 or higher. Evidence-based strategy results showed that feedback was not observed in any of the apps, and motivation strategy was observed in only 1 app. The sigma results of evidence-based strategies reflect that most of the apps fail to pass the mean.

Conclusions: App users desired a feature that allows them to communicate with a specialist, which is a missing in the available free apps. Despite the large number and accessibility of weight management apps, the quality and features of most are variable. It can be concluded from six sigma results that passing the mean does not ensure that the quality is consistently distributed through all app quality properties and Mobile App Rating Scale and evidence-based strategies do not give developers an indication of the acceptance of their apps by mobile users. This finding stresses the importance of reevaluating the passing criterion, which is ≥50% for designing an effective app.

(JMIR Mhealth Uhealth 2020;8(10):e19844) doi:10.2196/19844

KEYWORDS

mHealth; eHealth; smartphone; obesity; weight management; mobile apps; MARS; six sigma

XSL•FO
RenderX

Introduction

Obesity, a multifactorial health problem related to behavioral factors such as physical activity and diet [1], is a major public health concern worldwide, and its incidence nearly doubled from 1975 to 2016 according to a long-term analysis of trends using BMI [2]. In 2013, the Saudi Health Interview Survey reported that the prevalence of obesity was 29%, higher in women than men (33.5% vs 24.1%), and expected to continue to increase [3].

One of the factors causing individuals to have a more sedentary lifestyle is the use of smartphones [4]. In 2016, the percentage of Saudi people using smartphones reached 88% [5], and in 2018, the penetration rate of mobile cellular subscriptions was approximately 129% of the population according to the Communication and Information Technology Commission [6]. Data from large-scale surveys have showed that weight management apps were some of the most popular among medical and public health apps (mHealth apps) [7-9]. The use of these weight management apps showed effectiveness according to a systematic review and meta-analysis of 12 clinical trials that reported significant weight losses of 1 kg relative to traditional weight reduction interventions or intensive consulting [10].

A recent survey in Saudi Arabia aimed to explore weight management app use, barriers to use, and reasons for discontinued use among smartphone users [11]. The study demonstrated that more than 40% of participants used weight management apps and more than half of the app users were overweight or obese. However, a limited number of studies have assessed the quality of weight management apps using clear, identifiable, and justifiable quality assessment measures [7].

The quality evaluation of apps plays a vital role in assisting with the development and improvement of mHealth apps. Recently, several software product measures and metrics have been used to evaluate app quality, but many of these metrics are technical and highly dependent on the software type [12]. The Mobile App Rating Scale (MARS) is one of many tools for evaluating mHealth apps in smartphones [13], and it has been used in different studies to evaluate a variety of mHealth apps [14-16]. Apps can also be evaluated for their level of adherence to evidence-based strategies (EBS) by characterizing them depending on the presence or absence of app features [17]. Additionally, six sigma is a unique data-driven process used to test and analyze the policies, procedures, and measures of a quality plan to help software engineers easily detect quality discrepancies in apps [18]. In software engineering, six sigma is used to evaluate and control software quality such as that for mobile weight management apps [18].

Given the limited available studies that evaluate the quality of popular weight management apps, the aim of this study was to identify app users' perceptions and reasons for starting and stopping app use and evaluate the quality of weight management apps using two measurement tools: MARS and EBS. We also used six sigma methodology to ensure that the quality measured by the tools consistently meets customer expectations and determine whether the quality of the app is related to app use among users in Saudi Arabia.

Methods

Design and Sample

This cross-sectional web-based survey was conducted with Saudi smartphone users. Over 4 months, smartphone users were invited to participate in an anonymous, web-based survey hosted on a Microsoft platform. The survey link was advertised through social media and university portals. To increase the number of participants and their diversity, Twitter ads were used to promote the survey link in 8 Saudi cities, including Baha, Eastern, Tabuk, Asir, Makkah, Almadeinah, Hail, and Jazan. The ad ran during the second week of September 2019 for 7 days. The survey was open to anyone who wanted to participate, and both app users and nonusers were invited. Survey responses were collected over a 4-month period. The King Saud University institutional review board approved the study (reference E-19-4001) in May 2019. No personal identifying information was collected, and participants had the right to refuse to participate in the study. Informed consent was a requirement for participation. The online survey was in accordance with the Checklist for Reporting Results of Internet e-Surveys (CHERRIES, Multimedia Appendix 1) [19].

Participants aged younger than 18 years, those with inconsistent or illogical responses (ie, participant reported a weight of –10 kg), and those who were app users and failed to answer more than 50% of the questions on app use were excluded.

The sample size was calculated based on Saudi Arabia's current population of 10.7 million adults aged 18 years and older [20] and that 88% of adult Saudis were smartphone users at the time of the research [5]. A confidence level of 95% and a precision of 5% were used. Given these parameters, the minimum sample required for the analyses to have a power of 95% was 385 individuals.

Survey Items

The survey consisted of 31 questions encompassing the following domains: (1) sociodemographic characteristics (age, sex, residency, nationality, income, education, and employment), (2) general health (tobacco use, weight, height, medical diagnoses, physical activity, and diet), (3) app use (app name, reason for download, use pattern, app features, and whether app was recommended by health care provider), (4) app user perceptions (effectiveness, security, and accuracy), and (5) reason for discontinuing use of the app. These questions were obtained from a questionnaire used in the United States (Multimedia Appendix 2) [9].

The questionnaire was tested on a small sample of app users and nonusers (n=10). Questions and question order were revised based on feedback and results from the test to reduce response bias and enhance response time.

The final web-based questionnaire was presented in 3 steps. The first screen asked participants for their informed consent. Consent was a required response before the respondent could advance to the next screen. The second screen contained

XSL•FO
RenderX

sociodemographic and general health questions. The last screen contained questions on app use, stop use, and user perceptions. The back icon on each screen allowed participants to edit previous answers. The survey comprised a mix of open and closed questions and took between 5 and 9 minutes to complete.

App Search Strategy

In Saudi Arabia, the use of iOS devices increased by 3.6% and Android decreased by 3% from June to October 2019 [21]. Therefore, our search was limited to iOS users (Apple App Store). The App Store was searched twice for apps using English and Arabic search terms (weight loss, diet, and weight management). Through the search strategy, we aimed to find a large proportion of apps within our designated time frame. Thus, we used 6 search terms (3 in English and 3 in Arabic) with an arbitrary decision to include the first 10 displayed apps from each search. The search was done in November 2019 using a newly created Apple ID without any search history. The search was not filtered by any attribute as this feature does not exist in the App Store (iOS 13). The first 10 displayed apps from each search were reviewed based on predefined inclusion criteria: free, language is English or Arabic, made for the average consumer, and related to weight loss.

Quality Assessment Tools for Weight Management Apps

Mobile App Rating Scale

The quality of the weight management apps was assessed independently by two investigators (DA and GA) using MARS; any discrepancy was reviewed by a third investigator (ASA) and resolved by consensus. MARS contains 23 items rated on a 5-point scale (1.0=inadequate, 2.0=poor, 3.0=acceptable, 4.0=good, and 5.0=excellent) [13]. A total of 19 questions formed the objective quality section, which was divided into 4 scales: engagement, functionality, aesthetics, and information quality. Four questions formed the subjective quality section that evaluated user satisfaction. Apps were evaluated on an iPhone (iOS 13), and their star ratings from the App Store were obtained for further analysis. For the subjective scale, an average rating was taken.

Evidence-Based Strategy Assessment

App evidence-based strategies were characterized depending on the presence or absence of app features [17]. Strategies and indicators of adherence were as follows:

- Presence of self-monitoring capabilities for weight, meals, nutrition (including protein, fats, carbohydrates, fiber, and water), and physical activity
- Presence of goal setting with or without customization
- Healthy eating support including information, education, and skills development
- Physical activity support including information, education, and skills development
- Social support such as online communication with other users
- Weight and/or health assessment with or without personalization
- Motivational strategies including prompts, rewards, or a gamified design
- Personalized feedback

Apps were independently reviewed for their level of adherence to EBS by two investigators (DA and GA) with discrepancies reviewed by a third investigator (ASA) and resolved by consensus.

Six Sigma Evaluations

In software engineering, six sigma is used to evaluate and control software quality [22]. Software engineers use six sigma statistical methods like run, control charts, and process capability index (Cpk) to examine the software quality based on the quality measurements and quality standards. Six sigma is a data-driven, problem-solving process that consists of 5 stages, the Define-Measure-Analyze-Improve-Control process, to achieve six sigma goals. The results are then analyzed within the context of the software, and improvements are suggested based on the analysis outcomes (ie, if the sigma level is less than 6) [22].

Two researchers (ASA and GH) independently applied six sigma to evaluate apps quality based on quality data of MARS. Results were compared with MARS means, and the relation of the quality of the app to app use was identified.

Six sigma evaluation measures the quality of apps based on their behavior in all defined quality attributes, and based on this behavior it predicts app behavior on an undefined one. More specifically, we used the Cpk of six sigma to evaluate how close the app quality is to customer expectations considering its natural variability [22]. Cpk is a statistical measure of a software quality (ie, the ability of the software to meet software quality standard measures). The Cpk measures the natural variation of software quality relative to the quality standards limits. In addition, it allows the comparison of different software with respect to how well software meets quality standard limits. In a relative manner and within the context of this investigation, we used the limits of MARS (5 is considered excellent; 3 is considered acceptable) as the quality standard limits [13]. Thus, if a certain app scores high in six sigma, then that app will continually meet quality attribute limits of MARS and is expected to meet the limits of other quality attributes when considered [22]. The larger the Cpk, the higher the app quality. To calculate the Cpk, we used the following equation:

Where USL is the upper limit of customer expectations, LSL is the lower limit of the customer expectations, μ is the data mean, and σ is the standard deviation of the sample data.

Based on the Cpk, we can determine the sigma level of the app according to the following specification [22]:

- Cpk 0.33: sigma level 1
- Cpk 0.67: sigma level 2
- Cpk 1: sigma level 3
- Cpk 1.33: sigma level 4
- Cpk 1.67: sigma level 5
- Cpk 2: sigma level 6
- Cpk <0: limit is irrelevant

XSL•FO
RenderX

Levels below 3 are considered poor quality, and levels 3 and above are considered good quality. Level 6 represents the best quality [22].

Statistical Analysis

Data analysis was performed using SPSS Statistics version 24.0 (IBM Corporation). Numerical variables are represented as means and standard deviations, and categorical variables are represented as percentages. BMI was calculated as the weight in kilograms divided by the square of the height in meters (kg/m^2). The BMI categories were classified as follows: underweight (BMI<18.5), normal weight (18.5≤BMI≤24.9), overweight (25.0≤BMI≤29.9), moderate obesity (30.0≤BMI≤34.9), and severe obesity (BMI≥35).

Participants were stratified based on the following app use categories: (1) users (participants reported that they use a weight management app), (2) ex-users (participants reported that they used weight management apps and then stopped), and (3) nonusers (participants reported that they never used a weight management app). Analysis was done by user category with the data on app use and app user perceptions. Simple logistic regression analysis was applied to identify the association between BMI and use of a weight management app. Linear regression analysis was used to determine the relationship between the ranking of the apps in the App Store with the MARS score and also between the MARS score and the EBS criteria.

Results

Sociodemographic Characteristics and Health Status

The results of the sociodemographic characteristics for all participants stratified by use pattern (user, ex-user, or nonuser) are presented in Table 1. A total of 1209 people responded to the survey. Of the participants who read the welcome page and proceeded to consent, 98.68% (1193/1209) agreed to participate in the survey. The data were excluded in the analysis if the respondents were non-Saudi, aged younger than 18 years, provided inconsistent or illogical answers (ie, participant reported his weight as –10 kg), or if the participants are app users and did not answer more than 50% of questions on app use, which left 1074 responses for further analysis.

Of the respondents, 30.17% (324/1074) used a weight management app, 18.16% (195/1074) used an app and stopped, and 51.68% (555/1074) had never used a weight management app. The majority of the respondents were aged 18 to 31 years (785/1074, 73.09%); 69.93% (751/1074) were female and 30.07% (323/1074) were male. The majority of the participants were residents of the central region (706/1074, 65.74%).

Regarding health, only 53.35% (573/1074) thought that their general health was very good or excellent, and 35.20% (378/1074) reported that they never engage in physical activity for at least 15 minutes. A total of 44.51% (478/1074) had a BMI in the normal range, and 48.23% (518/1074) had overweight or moderate or severe obesity. Of respondents, 11.17% (120/1074) were smokers, and only 9.22% (99/1074) reported that a health care provider had recommended a weight management app to them. The most prevalent medical diagnoses that the respondents reported having were depression (69/1074, 6.42%), diabetes (41/1074, 3.82%), and hypertension (29/1074, 2.70%).

A simple logistic regression model was applied to test whether BMI predicted the use of a weight management app. Results revealed that high BMI was significantly associated with the use of a weight management app ($^2_1$=5.88, P<.02). The odds ratio for an increase in BMI was 1.18 (95% CI 1.03-1.35).

XSL•FO
RenderX

Table 1. Sociodemographic and health status characteristics of participants stratified by use patterns (n=1074).

Characteristic	User (n=324)	Ex-user (n=195)	Nonuser (n=555)	Total (n=1074)
Age in years, n (%)				
18-31	200 (61.73)	150 (76.92)	435 (78.38)	785 (73.09)
32-45	108 (33.33)	41 (21.03)	110 (19.82)	259 (24.12)
≥46	16 (4.94)	4 (2.05)	10 (1.80)	30 (2.79)
Gender, n (%)				
Female	233 (71.91)	144 (73.85)	374 (67.39)	751 (69.93)
Male	91 (28.09)	51 (26.15)	181 (32.61)	323 (30.07)
Region of country, n (%)				
Central	204 (62.96)	135 (69.23)	367 (66.13)	706 (65.74)
Southern	14 (4.32)	5 (2.56)	19 (3.42)	38 (3.54)
Eastern	37 (11.42)	10 (5.13)	43 (7.75)	90 (8.38)
Northern	33 (10.19)	24 (12.31)	56 (10.09)	113 (10.52)
Western	33 (10.19)	15 (7.69)	62 (11.17)	110 (10.24)
Living abroad	3 (0.93)	6 (3.08)	8 (1.44)	17 (1.58)
Education, n (%)				
High school	55 (16.98)	34 (17.44)	106 (19.10)	195 (18.16)
Bachelor’s degree	207 (63.89)	125 (64.10)	330 (59.46)	662 (61.64)
Postgraduate	62 (19.14)	36 (18.46)	119 (21.44)	217 (20.20)
Employment, n (%)				
Student	192 (59.26)	126 (64.62)	277 (49.91)	595 (55.40)
Not employed	21 (6.48)	18 (9.23)	52 (9.37)	91 (8.47)
Retired	0 (0)	1 (0.51)	11 (1.98)	12 (1.12)
Employee	111 (34.26)	50 (25.64)	215 (38.74)	376 (35.01)
Household income per month (SR), n (%)				
<5000	175 (54.01)	108 (55.38)	265 (47.75)	548 (51.02)
5001-10,000	16 (4.94)	20 (10.26)	58 (10.45)	94 (8.75)
10,001-20,000	59 (18.21)	31 (15.90)	107 (19.28)	197 (18.34)
>20,001	66 (20.37)	29 (14.87)	104 (18.74)	199 (18.53)
General health status, n (%)				
Excellent	87 (26.85)	41 (21.03)	124 (22.34)	252 (23.46)
Very good	99 (30.56)	60 (30.77)	162 (29.19)	321 (29.89)
Good	81 (25.00)	51 (26.15)	151 (27.21)	283 (26.35)
Average	52 (16.05)	39 (20.00)	106 (19.10)	197 (18.34)
Poor	5 (1.54)	4 (2.05)	12 (2.16)	21 (1.96)
Exercise frequency in the past week[a], n (%)				
None	78 (24.07)	75 (38.46)	225 (40.54)	378 (35.20)
1 day	34 (10.49)	33 (16.92)	91 (16.40)	158 (14.71)
2 days	60 (18.52)	37 (18.97)	83 (14.96)	180 (16.76)
3-4 days	88 (27.16)	30 (15.38)	107 (19.28)	225 (20.95)
5-6 days	64 (19.75)	20 (10.26)	49 (8.83)	133 (12.38)
Nutrition status of diet, n (%)				
Excellent	26 (8.02)	10 (5.13)	23 (4.14)	59 (5.49)

XSL•FO
RenderX

Characteristic	User (n=324)	Ex-user (n=195)	Nonuser (n=555)	Total (n=1074)
Very good	73 (22.53)	19 (9.74)	84 (15.14)	176 (16.39)
Good	103 (31.79)	70 (35.90)	175 (31.53)	348 (32.40)
Fair	84 (25.93)	66 (33.85)	182 (32.79)	332 (30.91)
Poor	38 (11.73)	30 (15.38)	91 (16.40)	159 (14.80)
BMI, n (%)				
Underweight	16 (4.94)	14 (7.18)	48 (8.65)	78 (7.26)
Normal	139 (42.90)	91 (46.67)	248 (44.68)	478 (44.51)
Overweight	92 (28.40)	54 (27.69)	160 (28.83)	306 (28.49)
Moderate obesity	50 (15.43)	20 (10.26)	65 (11.71)	135 (12.57)
Severe obesity	27 (8.34)	16 (8.21)	34 (6.13)	77 (7.17)
Smoking, n (%)				
Yes	32 (9.88)	21 (10.77)	67 (12.07)	120 (11.17)
No	292 (90.12)	174 (89.23)	488 (87.93)	954 (88.83)
App recommended by provider, n (%)				
Yes	52 (16.05)	26 (13.33)	21 (3.87)	99 (9.22)
No	265 (81.79)	165 (84.62)	396 (71.35)	826 (76.91)

[a]At least 15 minutes of exercise or physical activity.

User Perceptions and Use Patterns

Analyses of user perceptions and patterns of use of weight management app are presented in Multimedia Appendix 3. All data were obtained and analyzed from the user group only, where participants reported using a weight management app in the previous 6 months. The frequency of use was 2 or more times per day for 27.8% (90/324) of respondents, once a day for 20.1% (65/324), and a few times each week for 20.1% (65/324). The most common reasons for wanting to download a weight management app were to monitor food intake (319/860, 37.1%) and lose weight (258/860, 30.0%). The most common reasons for downloading a particular weight management app were recommendations from friends and family (153/294, 52.0%), and its rank in the App Store (65/294, 22.1%).

The most reported desirable features were (1) the possibility to be monitored by a specialist (323/976, 33.1%), (2) barcode identification of calorie content (191/976, 19.6%), (3) availability of nutrition information on numerous food items (153/976, 15.7%), (4) weekly or monthly progress report (152/976, 15.6%), and (5) constant reminders to follow a chosen diet or exercise plan (157/976, 16.1%). Most users agreed or strongly agreed that apps that suggested exercise and diet plans helped them lose weight (246/324, 75.9%).

A large proportion of weight management app users agreed or strongly agreed that apps were effective for losing weight (185/324, 57.1%). Regarding accuracy, 48.8% (158/324) of app users believed that apps are accurate, whereas 1.9% (6/324) did not use an app that recorded their data. Of the current weight management app users, only 7.1% (23/324) believed that weight management apps were not secure. A large proportion of app users (180/324, 55.6%) noted that they would never pay anything for a weight management app.

Reasons for Discontinuing Use

Analyses of the reasons for discontinuing use are presented in Multimedia Appendix 3. Of the participants, 18.16% (195/1074) had downloaded weight management apps that they no longer use. The most reported reasons for discontinuing use were (1) loss of interest (64/195, 32.8%), (2) hidden costs (53/195, 27.2%), (3) monitoring by a specialist was not offered (27/195, 13.8%), (4) difficulty using the app (21/195, 10.8%), and (5) language barrier (18/195, 9.2%).

Apps that Participants Used for Weight Management

Of the users, 53.2% (179/324) listed the name of the apps they used (Multimedia Appendix 4). However, because the question on app name was open-ended, some participants wrote ambiguous names or cited more than one app. The total number of weight management apps mentioned by users (more than once) was 267. The most reported apps were MyFitnessPal (145/267, 54.3%), health apps that come with a smartphone (16/267, 6.0%), StepsApp Pedometer (13/267, 4.9%), Soarrate (10/267, 3.8%), and Fitbit (10/267, 3.8%).

App Search

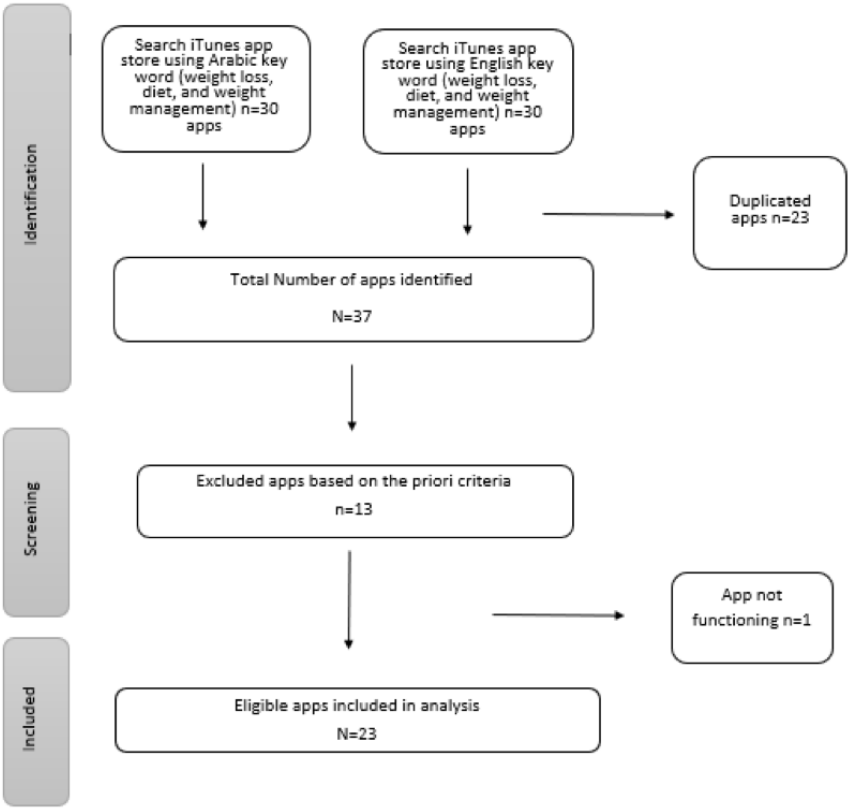
A total of 60 apps were identified from the search in the Saudi App Store. Of these apps, we excluded 23 that were duplicates, 12 that were not free, 1 app that did not function, and 1 app that appeared but was described in the app store as a book. Figure 1 provides a description of the search process.

A total of 23 apps met the inclusion criteria. The app ratings in the App Store varied from 2.1 to 5 (out of 5). Of these, 65% (15/23) received 4 or more stars, and the number of users that rated each app were 1000 or more (5/15, 33%), 100 to 1000 (5/15, 33%), or less than 100 (5/15, 33%). Of the apps, 48% (11/23) were updated by the app developers 1 day to 2 months

XSL•FO
RenderX

after data collection. Of the apps, 52% (12/23) were only available in Arabic and, among these, 25% (3/12) were apps for purchase. Approximately 30% (7/23) of the apps were available in English, and 86% (6/7) of these were apps for purchase. Only 17% (4/23) of the apps were available in both English and Arabic (Multimedia Appendix 5).

Figure 1. Flow diagram of app search.



App Quality Scores

Mobile App Rating Scale

Table 2 presents the final scores, means, and standard deviations for the following: (1) 4 subscales (engagement mean, functionality mean, aesthetics mean, and information mean), (2) overall quality (mean of 4 subscales), and (3) subjective quality of the 23 apps. The average MARS quality of the 23 apps was acceptable, and scores varied from 1.7 to 4.4 (out of 5.0). The reliability of the objective scales calculated as Cronbach alpha = .95. The average MARS quality score for the apps was 3.3 (SD 0.8) out of 5.0. The mean of the subjective scores was 2.5 (SD 1.1). Of the 4 subscales, functionality had the highest median score of 3.6 (SD 0.8), and information had the lowest median score of 2.9 (SD 0.9). Of the apps, 30% (7/23) received a quality mean score of 4.0 or higher; only one of these apps was originally an Arabic app. Of the apps, 30% (7/23) did not meet the acceptability score of 3.0 or higher, and all of these were in the Arabic language only (Table 2).

XSL•FO
RenderX

Table 2. The Mobile App Rating Scale mean scores for weight management apps[a].

App	Engagement	Functionality	Aesthetics	Information	MARS[b]	Subjective
Lose Weight for Men	4.4	5.0	4.3	4.0	4.4	4.0
Rashaqa adad alsoarat	4.6	4.8	4.7	3.5	4.4	3.8
MyFitnessPal	4.4	4.0	4.3	4.4	4.3	4.3
Fitbit	4.4	4.3	4.7	3.7	4.3	3.5
StepsApp Pedometer	3.8	4.8	4.3	3.8	4.2	2.8
Lose It! – Calorie Counter	4.4	4.0	4.0	3.4	4.0	3.3
Calorie Counter by FatSecret	4.2	4.3	3.7	3.6	4.0	4.0
Pacer Pedometer	4.8	3.5	3.7	3.5	3.9	3.8
Lifesum – Diet & Food Diary	4.2	3.8	4.0	3.5	3.9	3.0
7 Minute Workout: Fitness App	3.8	4.0	4.3	3.6	3.9	3.0
FUDC – Follow-Up Diet and Calories	4.2	4.0	3.3	2.8	3.6	2.5
mDiet	3.0	4.0	3.0	3.2	3.3	2.5
Adaad alsoaraat	4.2	3.0	3.0	3.1	3.3	2.3
Soarrate	3.4	3.3	3.0	3.6	3.3	2.5
Weight Tracker	3.2	3.8	2.6	2.5	3.0	2.5
My Diet Coach – Weight Loss	3.4	3.0	2.6	3.1	3.0	1.8
Tmarin manzliah	3.0	3.8	2.3	2.3	2.9	1.3
Alwazan almethali	2.6	3.0	3.0	2.4	2.8	1.3
Hesab alwazan almethali	2.6	3.0	3.0	1.6	2.6	1.3
Monabeh alsoaraat	2.6	2.5	2.3	1.8	2.3	1.3
Diet	1.8	2.8	2.0	2.0	2.2	1.3
Rajeem 7kilo fi esboaa	1.4	2.5	1.6	1.2	1.7	1.0
Rajem sareea	1.4	2.5	1.6	1.2	1.7	1.0
Subscales mean (SD)	3.5 (1.0)	3.6 (0.8)	3.3 (1.0)	2.9 (0.9)	3.3 (0.8)	2.5 (1.1)

[a]All items were rated on a 5-point scale from 1=inadequate to 5=excellent.

[b]MARS: Mobile App Rating Scale.

The MARS scores for the apps available only in the Arabic language ranged between 1.7 and 4.4 and for apps available in only English, the scores ranged between 3.0 and 4. Apps available in English and Arabic had scores that ranged between 3.0 and 4.4 (Multimedia Appendix 6).

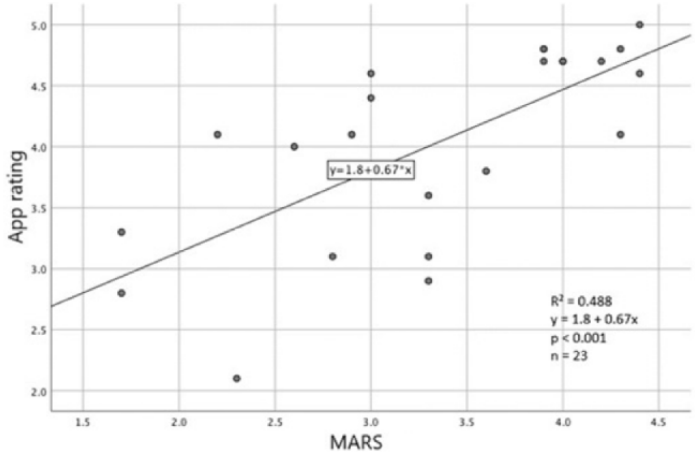
A simple linear regression was applied to predict the relation between the star rating in the App Store and the MARS score and showed a significant association ($F_{1,21}$=20.018, P<.001) with an R^2 of .488 (Figure 2).

For engagement and aesthetics, Lose Weight for Men, Rashaqa adad alsoarat, MyFitnessPal, Fitbit, Lose It! – Calorie Counter, and Lifesum – Diet & Food Diary were the highest rated apps. Despite the high functionality of the Arabic weight management apps, Multimedia Appendix 6 shows that the lowest engagement mean was for Arabic apps.

Of the 23 apps, 14 were mentioned more than once by app users in the survey; 11 apps were found in the survey responses but did not appear in our app search. Multimedia Appendix 7 shows the MARS mean scores of the apps and the number of users who reported using the apps.

XSL•FO
RenderX

Figure 2. Regression analyses of the association between the star rating in the App Store and the Mobile App Rating Scale score.



Assessment of Evidence-Based Strategies

Table 3 provides the frequency of evidence-based strategies within the included apps. The most common strategies were as follows (no free apps or versions offered personalized feedback to the user):

- Self-monitoring (17/23, 74%) allowed the user to track weight and/or physical activity over time. A few apps included more comprehensive tracking options such as for nutrition, sleep, and cardiometabolic indicators
- Weight and/or health assessment features (17/23, 74%) enable the app to assess the BMI and/or calorie requirement
- Goal setting (13/23, 57%) mainly consisted of goals for weight loss, calorie balance, water intake, or physical activity
- Healthy eating support (12/23, 52%) was mostly in the form of healthy eating guidelines, meal plans, and nutritional information on specific foods
- Physical activity support (9/23, 39%) including physical activity tips and plans
- Motivational strategies (7/23, 30%) included prompts, gamification, or use of rewards such as points for meeting weight goals or physical activity levels
- Social support components (8/23, 35%) included the ability to communicate online with other users

Of the apps that were available in Arabic, 50% (6/12) had one or two EBS features, 17% (2/12) had 3 features, 25% (3/12) had 5 features, and 8% (1/12) had 6 features. Of these, 75% (9/12) of the apps provided weight/health assessment and 66% (8/12) provided healthy eating support. Feedback was not observed in any of the apps, and motivation strategy was observed in 1 app.

Self-monitoring, goal setting, and weight/health assessment were observed in the majority of apps (3/4, 75%) available in English and Arabic. However, the percentages for healthy eating, social support, and feedback were 25% (1/4), 25% (1/4), and 0%, respectively.

Of the apps not available in Arabic, 43% (3/7) had 3 features, 29% (2/7) had 5 features, and 29% (2/7) had 7 features. Of these, 100% (7/7) of the apps provided self-monitoring and 71% (5/7) provided goal setting, weight/health assessment, and motivation strategy.

XSL•FO
RenderX

Table 3. Assessment of evidence-based strategies for weight management apps[a].

App name	Self-monitoring	Goal setting	Healthy eating support	Physical activity support	Social support	Weight /health assessment	Motivational strategies	Personalized feedback	Total EBS[b] within app (n=8) n (%)
MyFitnessPal	1	1	1	1	1	1	1	0	7 (88)
Fitbit	1	1	1	1	1	1	1	0	7 (88)
Rashaqa adad alsoarat	1	1	0	1	1	1	1	0	6 (75)
FUDC - Follow-Up Diet and Calories	1	1	1	0	1	1	0	0	5 (63)
Pacer Pedometer	1	1	0	1	1	0	1	0	5 (63)
Calorie Counter by FatSecret	1	1	1	0	1	1	0	0	5 (63)
Adaad alsoaraat	1	1	1	1	0	1	0	0	5 (63)
Soarrate	1	1	1	0	1	1	0	0	5 (63)
Weight Tracker	1	1	1	0	0	1	0	0	4 (50)
Lose It! - Calorie Counter	1	1	0	0	1	1	0	0	4 (50)
My Diet Coach - Weight Loss	1	1	0	0	0	0	1	0	3 (38)
mDiet	0	1	1	0	0	1	0	0	3 (38)
Lifesum - Diet & Food Diary	1	0	0	0	0	1	1	0	3 (38)
Tmarin manzliah	1	0	0	1	0	1	0	0	3 (38)
7 Minute Workout: Fitness App	1	0	0	1	0	1	0	0	3 (38)
Lose Weight for Men	1	0	0	1	0	1	0	0	3 (38)
StepsApp Pedometer	0	1	0	1	0	0	1	0	3 (38)
Alwazan almethali	1	0	0	0	0	1	0	0	2 (25)
Diet	0	0	1	0	0	1	0	0	2 (25)
Hesab alwazan almethali	1	0	0	0	0	1	0	0	2 (25)
Monabeh alsoaraat	0	0	1	0	0	0	0	0	1 (13)
Rajeem 7kilo fi esboaa	0	0	1	0	0	0	0	0	1 (13)
Rajem sareea	0	0	1	0	0	0	0	0	1 (13)
Total apps	17	13	12	9	8	17	7	0	N/A[c]

[a]1=presence, 0=absence.

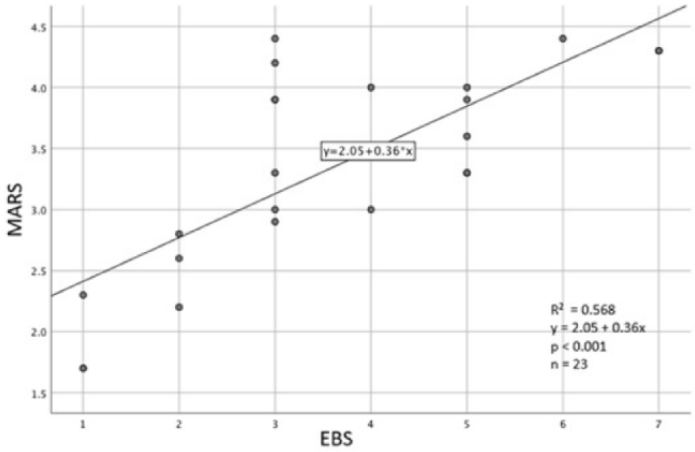
[b]EBS: evidence-based strategies.

[c]N/A: Not applicable.

A simple linear regression was calculated to predict the MARS scale based on EBS. A significant regression equation was found ($F_{1,21}$=27.66, P<.001), with an R^2 of .568 (Figure 3). Multimedia Appendix 8 shows the EBS scores of the apps and the number of users who reported using the apps.

XSL•FO
RenderX

Figure 3. Regression analyses of the association between the Mobile App Rating Scale score and the evidence-based strategy criteria.



Six Sigma and the Process Capability Index

Table 4 shows the Cpk calculations for each of the 23 apps found in the App Store and their sigma levels. Note that we used the limits mean in Cpk calculations. This is because a calculation based on data mean mostly provides a negative Cpk, which leads to the conclusion that the score levels of MARS are irrelevant to the actual app quality. Therefore, we adjusted the analysis assumption to provide results that are more relevant.

The quality thresholds of the two popular measurements tools, MARS and EBS, used in this analysis are presented in Table 5. Six sigma was used to evaluate the thresholds of these tools based on data from Table 2, Table 3, and Multimedia Appendix 4.

The calculation for the EBS Cpk was as follows: μ=3.608695652, σ=1.777105165, and Cpk=–0.073397335. The calculation for the MARS Cpk was as follows: μ=3.347826087, σ=0.845775365, and Cpk=0.137083715.

The sigma results indicate that EBS has no level, as its Cpk is negative. This means that most of the apps failed to pass the mean indicating that newly developed apps will probably fail as well. The sigma level of MARS at level 1 means that their evaluation criteria are more realistic and closer to the natural structure of weight management apps.

XSL•FO
RenderX

Table 4. Six sigma behavioral evaluation for the quality model: engagement, functionality, aesthetics, and information.

App name	Engagement	Functionality	Aesthetics	Information	μ[a]	σ[b]	Cpk[c]	Sigma level
Pacer Pedometer	4.8	3.5	3.7	3.5	3.88	0.62	0.53	2
Rashaqa adad alsoarat	4.6	4.8	4.7	3.5	4.40	0.61	0.55	2
Lose It! - Calorie Counter	4.4	4.0	4.0	3.4	3.95	0.41	0.81	3
Fitbit	4.4	4.3	4.7	3.7	4.28	0.42	0.79	3
Lose Weight for Men	4.4	5.0	4.3	4.0	4.43	0.42	0.79	3
MyFitnessPal	4.4	4.0	4.3	4.4	4.28	0.19	1.76	6
Adaad alsoaraat	4.2	3.0	3.0	3.1	3.33	0.59	0.57	2
FUDC - Follow-Up Diet and Calories	4.2	4.0	3.3	2.8	3.58	0.64	0.52	2
Calorie Counter by FatSecret	4.2	4.3	3.7	3.6	3.95	0.35	0.95	3
Lifesum - Diet & Food Diary	4.2	3.8	4.0	3.5	3.88	0.30	1.12	4
StepsApp Pedometer	3.8	4.8	4.3	3.8	4.18	0.48	0.70	3
7 Minute Workout: Fitness App	3.8	4.0	4.3	3.6	3.93	0.30	1.12	4
Soarrate	3.4	3.3	3.0	3.6	3.33	0.25	1.33	4
My Diet Coach - Weight Loss	3.4	3.0	2.6	3.1	3.03	0.33	1.01	4
Weight Tracker	3.2	3.8	2.6	2.5	3.03	0.60	0.55	2
Tmarin manzliah	3.0	3.8	2.3	2.3	2.85	0.71	0.47	2
mDiet	3.0	4.0	3.0	3.2	3.30	0.48	0.70	3
Hesab alwazan almethali	2.6	3.0	3.0	1.6	2.55	0.66	0.50	2
Monabeh alsoaraat	2.6	2.5	2.3	1.8	2.30	0.36	0.94	3
Alwazan almethali	2.6	3.0	3.0	2.4	2.75	0.30	1.11	4
Diet	1.8	2.8	2.0	2.0	2.15	0.44	0.75	3
Rajeem 7kilo fi esboaa	1.4	2.5	1.6	1.2	1.68	0.57	0.58	2
Rajem sareea	1.4	2.5	1.6	1.2	1.68	0.57	0.58	2

[a]μ: data mean.

[b]σ: standard deviation of the sample data.

[c]Cpk: process capability index.

XSL•FO
RenderX

Table 5. Comparison of app quality according to mobile users (survey outcomes), evidence-based strategy, and the Mobile App Rating Scale.

App name	Limits 4.0-8.0 EBS[a]	Limits 3.0-5.0 MARS[b]	Survey	Within limit of N≥50 EBS	Within limit of acceptable=3 MARS
MyFitnessPal	7.0	4.3	54.3	TRUE	TRUE
Fitbit	7.0	4.3	3.8	TRUE	TRUE
Rashaqa adad alsoarat	6.0	4.4	1.5	TRUE	TRUE
FUDC - Follow-Up Diet and Calories	5.0	3.6	0	TRUE	TRUE
Pacer Pedometer	5.0	3.9	0.7	TRUE	TRUE
Calorie Counter by FatSecret	5.0	4.0	1.9	TRUE	TRUE
Adaad alsoaraat	5.0	3.3	3.0	TRUE	TRUE
Soarrate	5.0	3.3	3.8	TRUE	TRUE
Weight Tracker	4.0	3.0	0.7	TRUE	TRUE
Lose It! - Calorie Counter	4.0	4.0	3.4	TRUE	TRUE
My Diet Coach - Weight Loss	3.0	3.0	0.7	FALSE	TRUE
mDiet	3.0	3.3	2.2	FALSE	TRUE
Lifesum - Diet & Food Diary	3.0	3.9	3.0	FALSE	TRUE
Tmarin manzliah	3.0	2.9	0	FALSE	FALSE
7 Minute Workout: Fitness App	3.0	3.9	0	FALSE	TRUE
Lose Weight for Men	3.0	4.4	0	FALSE	TRUE
StepsApp Pedometer	3.0	4.2	4.9	FALSE	TRUE
Alwazan almethali	2.0	2.8	0.7	FALSE	FALSE
Diet	2.0	2.2	0	FALSE	FALSE
Hesab alwazan almethali	2.0	2.6	0	FALSE	FALSE
Monabeh alsoaraat	1.0	2.3	0	FALSE	FALSE
Rajeem 7kilo fi esboaa	1.0	1.7	0	FALSE	FALSE
Rajem sareea	1.0	1.7	0	FALSE	FALSE

[a]EBS: evidence-based strategy.

[b]MARS: Mobile App Rating Scale.

Discussion

Principal Findings

In this survey, we identified user perceptions of weight management apps and reasons for using such apps. We also assessed the quality of weight management apps available at the Saudi App Store using the MARS quality score, an EBS assessment, and six sigma evaluations. The user surveys and evaluation of features indicated that personalized feedback is the most common feature lacking among commercial apps, and it is the feature that users most desired. According to the majority of ex-users, the reasons for stopping their use of weight management apps included a loss of interest and hidden costs.

Regarding app appraisal, the EBS and MARS quality scores showed that the quality of weight management app features was variable. The behavior of a weight management app in all 4 MARS quality attributes predicts its behavior toward other quality attributes, and as a result determines its overall quality based on stronger judgment than the mean only. For instance, Rashaqa adad alsoarat had the second best score in MARS but only achieved level 2 as its behavior to the different quality attributes was variable. MyFitnessPal was not the best in MARS; however, it achieved level 6 because of consistent behavior in all 4 quality attributes, resulting in 54% of weight management app users using MyFitnessPal.

Six sigma evaluations indicated that several apps have high scores for engagement and functionality, but these are not matched with MARS. The six sigma results lead us to the following question: If we apply a MARS or EBS evaluation to an app, does this indicate that the app will be used by mobile users?

Measurement tools like MARS and EBS evaluate software based on the mean score only. More specifically, if the app passes the mean, it is assumed to be of good quality, and good quality apps should find their way to the market, but this is not the case. In fact, passing the mean does not ensure the quality is consistently distributed through all of the app's quality properties, and in turn that the app has good quality. Using MARS only allows an app to score high if it meets one out of

XSL•FO
RenderX

the 4 quality attributes of MARS. However, the sigma results of EBS reflect that most of the apps failed to pass the mean, indicating that newly developed apps will likely fail. This stresses the importance of reevaluating the passing criterion, which is ≥50%. We can infer from Multimedia Appendix 7 and 8 that neither the MARS nor the EBS tool gives developers an indication of the acceptance of their app by mobile users. Thus, based on Cpk results, six sigma is a better tool to identify the quality of a weight management app and if it actually meets MARS quality attributes.

Comparison With Prior Work

We targeted the general population to approach different types of users. This approach was previously adopted in studies in the United States [9], Saudi Arabia [11], Germany [23], and China [24]. In our study, 30% of the survey participants were app users, a percentage that is lower than in other studies [9,11,24], possibly because we excluded users who had not used the app in the previous 6 months, unlike in other studies. We found that the highest use of weight management apps was among women; similar studies have also found that women are more involved in weight control and healthy eating than men [25,26].

The majority of users believe that weight management apps are effective. The effectiveness of weight management apps was established in a systematic review and meta-analysis of 8 randomized controlled trials comparing the use of weight management apps for weight loss to traditional care or intensive consulting [10]. The systematic review and meta-analysis results indicated a significant effect of weight management apps through a 1 kg reduction in body weight.

To the best of our knowledge, evaluating weight management apps in Arabic using EBS has not been done previously. However, in 2016, Arabic apps were evaluated based on 13 evidence-informed practices [27]; the difference between these two tools is that EBS represents broader criteria than evidence-based practices. The advantage of using EBS is in describing the overarching evidence-based quality of the current market for weight management apps. Evaluating Arabic apps using different strategies from before extends the current literature. In our sample, the most desirable features reported by the app users are the possibility to be monitored by a specialist and barcode identification of calorie content. These features reflect two strategies of evidence-based features, which are personalized feedback and healthy eating support [17].

The average number of evidence-based features present in an app was between 3 and 4, which was more than in a previous study [17] and could be explained by the fact that app content is improving over time. In our findings, the most popular feature is self-monitoring, which is consistent with the findings of previous studies [17]. In contrast, the majority of the apps in the Saudi App Store lack the personalized feedback feature.

Weight management apps in Arabic have limited strategies. However, a comparison with a study conducted on Arabic weight management apps found that our results show improvements in the Arabic apps [27]. Overall, weight management features of apps found in the Saudi App Store have weak adherence to EBS, which may be the result of the lack of health care expert involvement during app development.

In our study, the average MARS quality scores for the 23 weight management apps available in the App Store varied significantly, with 7 apps not meeting the minimum acceptability score of 3.0. None of the apps received the maximum score of 5.0. These findings are similar to that of a previous study that examined 23 weight management apps available in the App Store and Google Play [28]. However, the maximum quality score was higher than our findings, which may be because of the inclusion of paid apps [28].

The weakest MARS subscale was for the quality of information. In contrast, the functionality subscale had the highest median score, a result that is in line with previous studies that used MARS to assess the quality of mindful-eating mobile apps [14] and weight management apps [28]. Another study that evaluated apps for managing tinnitus [28] also found that functionality had the highest MARS subscale.

MARS indicates that some apps, like the StepsApp Pedometer, have high scores for engagement and functionality that are not reflected in the EBS results. Several apps have powerful features and efficiency but require careful future evaluation for long-term use.

In our survey, the MyFitnessPal app was the most cited by the participants. This app was mentioned more than 140 times, and the second most cited app was mentioned only 16 times. Furthermore, the assessment of the quality and features of weight management apps showed that MyFitnessPal has good quality traits and 88% of the EBS. Other weight management apps had equal quality and evidence-based features but were less popular among the participants. This lack of popularity could be the result of other factors that impact app popularity. For example, we asked app users about their reasons for downloading a specific app from the App Store. The most reported reason was recommendations from friends or relatives. Therefore, app users recommending a particular app to their friends and relatives could cause a snowball effect and lead to an increase in the use of a particular app.

Strengths and Limitations

Our study, which included a large number of participants across Saudi Arabia, shows the overall quality of weight management apps, reports on the areas of the apps that are weak or strong, and notes the app strategies that are absent or present. Such information can assist app developers in enhancing their current apps or developing new, better apps.

Although this study represents the first appraisal of weight management apps downloaded from the App Store in Saudi Arabia, we only included free apps or free versions. Because Android apps were not reviewed, and the app search was limited to the Saudi App Store, the findings cannot be generalized to all smartphone apps. As a result, the possibility exists that we missed additional apps or features. Also, MARS includes items that may be a source of subjective bias; however, having two independent reviewers applying MARS, with a third reviewer resolving discrepancy, helped reduce bias, and the same method was used when applying EBS. Another limitation of our study

XSL•FO
RenderX

is that the survey data are self-reported, which could be a source of error. Furthermore, classifying participants based on BMI categories does not reflect body composition, therefore, we were unable to report results on fat and fat-free mass. In addition, our study was limited by the recruitment strategy, so we may have missed individuals with limited internet access or those not using social media.

Conclusions

Despite the large quantity and easy accessibility of weight management apps, the quality and features of the majority of apps from the App Store included in the study remains low. Improvements made to Arabic apps have been limited, and the information content needs to be enhanced. In general, we found that the weakest areas of apps from the App Store are information quality and graphic design. App users wanted a feature that allows them to communicate with a specialist, so this feature should be considered by app developers in the future. Additionally, we can infer that MARS and EBS do not give developers an indication of the acceptance of their apps by mobile users. This stresses the importance of reevaluating the passing criterion and approaching users when developing an app. Our findings lead to the recommendations that significant attention should be paid to supporting the maintainability of weight management apps in the future.

Acknowledgments

The authors extend their appreciation to the Deputyship for Research & Innovation, “Ministry of Education” in Saudi Arabia for funding this research work through the project number IFKSUHI-1441-346.

Authors' Contributions

GA, DA, ASA, and GH performed the analysis, interpreted the data, and prepared the manuscript. AA and SA revised the work critically for important intellectual content; GA and DA were responsible for final content; and all authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Checklist for reporting results of internet e-surveys (CHERRIES).
[DOCX File , 21 KB - mhealth_v8i10e19844_app1.docx]

Multimedia Appendix 2
Questionnaire: weight-management apps in Saudi Arabia: feature assessment, and quality evaluation.
[DOCX File , 23 KB - mhealth_v8i10e19844_app2.docx]

Multimedia Appendix 3
Users perception, pattern of use, and reasons for discounting use.
[DOCX File , 34 KB - mhealth_v8i10e19844_app3.docx]

Multimedia Appendix 4
Apps that participants used for weight management (n=267).
[DOCX File , 31 KB - mhealth_v8i10e19844_app4.docx]

Multimedia Appendix 5
Weight management apps and their ratings in the App Store.
[DOCX File , 32 KB - mhealth_v8i10e19844_app5.docx]

Multimedia Appendix 6
The Mobile App Rating Scale mean scores for weight management apps stratified by app language.
[DOCX File , 30 KB - mhealth_v8i10e19844_app6.docx]

Multimedia Appendix 7
The Mobile App Rating Scale mean scores and the number of users who reported using the apps.
[DOCX File , 21 KB - mhealth_v8i10e19844_app7.docx]

Multimedia Appendix 8

XSL•FO
RenderX

Evidence-based strategy assessment scores and the number of users who reported using the apps.
[DOCX File , 21 KB - mhealth_v8i10e19844_app8.docx]

References

1. Katzmarzyk PT, Lear SA. Physical activity for obese individuals: a systematic review of effects on chronic disease risk factors. Obes Rev 2012 Feb;13(2):95-105. [doi: 10.1111/j.1467-789X.2011.00933.x] [Medline: 21951422]
2. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. Lancet 2017 Dec 16;390(10113):2627-2642 [FREE Full text] [doi: 10.1016/S0140-6736(17)32129-3] [Medline: 29029897]
3. Memish ZA, El Bcheraoui C, Tuffaha M, Robinson M, Daoud F, Jaber S, et al. Obesity and associated factors—Kingdom of Saudi Arabia, 2013. Prev Chronic Dis 2014 Oct 09;11:E174 [FREE Full text] [doi: 10.5888/pcd11.140236] [Medline: 25299980]
4. Fennell C, Barkley J, Lepp A. The relationship between cell phone use, physical activity, and sedentary behavior in adults aged 18-80. Comput Human Behav 2019;90:53-59 [FREE Full text] [doi: 10.1016/j.chb.2018.08.044]
5. Annual Report 2016. Communications and Information Technology Commission. 2016. URL: https://www.citc.gov.sa/en/mediacenter/annualreport/Documents/PR_REP_012Eng.pdf [accessed 2019-11-12]
6. Global and regional information and communication technologies estimates. Communications and Information Technology Commission. 2018 Jan 1. URL: https://www.citc.gov.sa/ar/mediacenter/newsletter/Documents/PR_ENL_033_Ar.pdf [accessed 2019-12-28]
7. Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. JMIR Mhealth Uhealth 2014;2(2):e19 [FREE Full text] [doi: 10.2196/mhealth.3088] [Medline: 25099179]
8. Nikolaou CK, Lean MEJ. Mobile applications for obesity and weight management: current market characteristics. Int J Obes (Lond) 2016 Nov 22;41(1):200-202. [doi: 10.1038/ijo.2016.186] [Medline: 27780974]
9. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
10. Flores MG, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
11. Aljuraiban GS. Use of weight-management mobile phone apps in Saudi Arabia: a web-based survey. JMIR Mhealth Uhealth 2019 Feb 22;7(2):e12692 [FREE Full text] [doi: 10.2196/12692] [Medline: 30794205]
12. Ruan C, Yang J. Software quality evaluation model based on weighted mutation rate correction incompletion G1 combination weights. Math Probl Eng 2014;2014:9 [FREE Full text] [doi: 10.1155/2014/541292]
13. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
14. Lyzwinski LN, Edirippulige S, Caffery L, Bambling M. Mindful eating mobile health apps: review and appraisal. JMIR Ment Health 2019 Aug 22;6(8):e12820 [FREE Full text] [doi: 10.2196/12820] [Medline: 31441431]
15. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the Mobile App Rating Scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718 [FREE Full text] [doi: 10.2196/10718] [Medline: 30361196]
16. Stec MA, Arbour MW, Hines HF. Client-centered mobile health care applications: using the Mobile Application Rating Scale instrument for evidence-based evaluation. J Midwifery Womens Health 2019 May;64(3):324-329. [doi: 10.1111/jmwh.12941] [Medline: 30887711]
17. Rivera J, McPherson A, Hamilton J, Birken C, Coons M, Iyer S, et al. Mobile apps for weight management: a scoping review. JMIR Mhealth Uhealth 2016 Jul 26;4(3):e87 [FREE Full text] [doi: 10.2196/mhealth.5115] [Medline: 27460502]
18. Alhuhud G, Altamimi W. Quality evaluation of mobile game: Miftah Alfasaha. Mobile Information Systems 2016;2016 [FREE Full text] [doi: 10.1155/2016/4581279]
19. Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet e-surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
20. Demography Survey. General Authority for Statistics. 2016. URL: https://www.stats.gov.sa/sites/default/files/en-demographic-research-2016_2.pdf [accessed 2019-11-11]
21. Mobile operating system market share Saudi Arabia. StatCounter. 2019. URL: https://gs.statcounter.com/os-market-share/mobile/saudi-arabia [accessed 2019-12-15]
22. Six Sigma Resources for Six Sigma Quality. iSixSigma. 2015. URL: http://www.isixsigma.com/ [accessed 2019-12-18]
23. Ernsting C, Dombrowski SU, Oedekoven M, Kanzler M, Kuhlmey A, Gellert P. Using smartphones and health apps to change and manage health behaviors: a population-based survey. J Med Internet Res 2017 Apr 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]

XSL•FO
RenderX

24. Xie Z, Nacioglu A, Or C. Prevalence, demographic correlates, and perceived impacts of mobile health app use amongst Chinese adults: cross-sectional survey study. JMIR Mhealth Uhealth 2018 Dec 26;6(4):e103 [FREE Full text] [doi: 10.2196/mhealth.9002] [Medline: 29699971]
25. Wardle J, Haase AM, Steptoe A, Nillapun M, Jonwutiwes K, Bellisle F. Gender differences in food choice: the contribution of health beliefs and dieting. Ann Behav Med 2004 Apr;27(2):107-116. [Medline: 15053018]
26. Fagerli RA, Wandel M. Gender differences in opinions and practices with regard to a "healthy diet". Appetite 1999 Apr;32(2):171-190. [doi: 10.1006/appe.1998.0188] [Medline: 10097024]
27. Alnasser AA, Amalraj RE, Sathiaseelan A, Al-Khalifa AS, Marais D. Do Arabic weight-loss apps adhere to evidence-informed practices? Transl Behav Med 2016 Dec;6(3):396-402 [FREE Full text] [doi: 10.1007/s13142-015-0347-7] [Medline: 27528528]
28. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016;13(1):35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet e-Surveys
Cpk: process capability index
EBS: evidence-based strategy
LSL: lower limit of customer expectations
MARS: Mobile App Rating Scale
mHealth: mobile health
USL: upper limit of customer expectations
μ: data mean
σ: standard deviation of the sample data

Edited by L Buis; submitted 04.05.20; peer-reviewed by J Alvarez Pitti, L Tully, C Fernández; comments to author 17.06.20; revised version received 16.07.20; accepted 04.09.20; published 26.10.20.

Please cite as:
Alshathri DM, Alhumaimeedy AS, Al-Hudhud G, Alsaleh A, Al-Musharaf S, Aljuraiban GS
Weight Management Apps in Saudi Arabia: Evaluation of Features and Quality
JMIR Mhealth Uhealth 2020;8(10):e19844
URL: http://mhealth.jmir.org/2020/10/e19844/
doi:10.2196/19844
PMID:33104013

©Dalal M Alshathri, Abeer S Alhumaimeedy, Ghada Al-Hudhud, Aseel Alsaleh, Sara Al-Musharaf, Ghadeer S Aljuraiban. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 26.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating Asthma Mobile Apps to Improve Asthma Self-Management: User Ratings and Sentiment Analysis of Publicly Available Apps

Marlene Camacho-Rivera[1], BSc, MPH, MSc, SCD; Huy Vo[2], PhD; Xueqi Huang[2], MSc; Julia Lau[2], BSc; Adeola Lawal[3], BSc; Akira Kawaguchi[2], PhD

[1]Department of Community Health Sciences, SUNY Downstate Health Sciences University, Brooklyn, NY, United States
[2]Department of Computer Science, Grove School of Engineering, City College of New York, New York, NY, United States
[3]Department of Community Health and Social Medicine, CUNY School of Medicine, New York, NY, United States

Corresponding Author:
Marlene Camacho-Rivera, BSc, MPH, MSc, SCD
Department of Community Health Sciences
SUNY Downstate Health Sciences University
450 Clarkson Avenue, MSC 43
Brooklyn, NY, 11203
United States
Phone: 1 7182704386
Email: marlene.camacho-rivera@downstate.edu

Abstract

Background: The development and use of mobile health (mHealth) apps for asthma management have risen dramatically over the past two decades. Asthma apps vary widely in their content and features; however, prior research has rarely examined preferences of users of publicly available apps.

Objective: The goals of this study were to provide a descriptive overview of asthma mobile apps that are publicly available and to assess the usability of asthma apps currently available on the market to identify content and features of apps associated with positive and negative user ratings.

Methods: Reviews were collected on June 23, 2020, and included publicly posted reviews until June 21, 2020. To characterize features associated with high or low app ratings, we first dichotomized the average user rating of the asthma app into 2 categories: a high average rating and a low average rating. Asthma apps with average ratings of 4 and above were categorized as having a high average rating. Asthma apps with average ratings of less than 4 were categorized as having a low average rating. For the sentiment analysis, we modeled both 2-word (bi-gram) and 3-word (tri-gram) phrases which commonly appeared across highly rated and lowly rated apps.

Results: Of the 10 apps that met the inclusion criteria, a total of 373 reviews were examined across all apps. Among apps reviewed, 53.4% (199/373) received high ratings (average ratings of 4 or 5) and 47.2% (176/373) received low ratings (average ratings of 3 or less). The number of ratings across all apps ranged from 188 (AsthmaMD) to 10 (My Asthma App); 30% (3/10) of apps were available on both Android and iOS. From the sentiment analysis, key features of asthma management that were common among highly rated apps included the tracking of peak flow readings (n=48), asthma symptom monitoring (n=11), and action plans (n=10). Key features related to functionality that were common among highly rated apps included ease of use (n=5). Users most commonly reported loss of data (n=14) and crashing of app (n=12) as functionality issues among poorly rated asthma apps.

Conclusions: Our study results demonstrate that asthma app quality, maintenance, and updates vary widely across apps and platforms. These findings may call into question the long-term engagement with asthma apps, a crucial factor for determining their potential to improve asthma self-management and asthma clinical outcomes.

(JMIR Mhealth Uhealth 2020;8(10):e15076) doi:10.2196/15076

KEYWORDS

mHealth; asthma apps; sentiment analysis; user ratings; smartphone; mobile phone

XSL•FO
RenderX

Introduction

Background

Asthma is the leading chronic disease in children and adolescents and one of the most common among adults, with more than 25 million Americans impacted [1]. Nationally, asthma accounts for 11 million doctor's office visits, 439,435 discharges from hospital inpatient care, and 1.7 million emergency department visits each year [2]. Further, economic costs associated with asthma are significant, both for the United States and for asthmatics and their families. From 2008 to 2013, the annual economic cost of asthma was more than US $81.9 billion, including medical costs and loss of work and school days: US $3 billion in losses due to missed work and school days, US $29 billion due to asthma-related mortality, and US $50.3 billion in medical costs [2].

Many factors combine to contribute to poor asthma rates and worsen outcomes. Provider practice behaviors, suboptimal access to health care, lack of patient knowledge regarding proper medication use, and patients' difficulty adhering to medical regimens contribute to poor asthma outcomes [3-8]. With appropriate medical care including education, patients can achieve asthma control; however, there are competing challenges to providing effective patient education, such as time constraints and prioritizing other issues above asthma education [9].

Mobile health apps and devices to monitor and track health-related questions are a rapidly growing field within the public health, data science, and technology sectors [10-16], and offer potential opportunities to address some of the barriers to patient asthma education [17-20].

However, promotion of new technologies is predicated on the hypothesis that enabling people to quantify their own behaviors will drive health behavior change through contextualization and goal setting [21]. Mobile apps (configured to work independently or with wearable devices) can be used to record and track health-related behaviors, provide tailored education, and send reminders and prompts. Nationally, approximately 80% of US adults own a smartphone and close to 60% of smartphone owners also report installing one or more health apps onto their smartphones [22,23].

In recent years, there has been a proliferation of new mobile apps for the self-assessment and self-management of asthma [24,25]. However, uptake of asthma apps has been sparse, despite evidence of their efficacy in impacting asthma outcomes [26]. Results from randomized controlled studies show that use of mobile health apps improves asthma symptoms and medication compliance, which should, in turn, reduce emergency department visits and hospital admissions [27-29]. To our knowledge, there has not been a comprehensive review of asthma apps which considers user ratings, features, and reviews as a tool to identify potential barriers and facilitators of app use among individuals self-managing asthma.

Objectives

The aims of this study were (1) to provide a descriptive overview of asthma mobile apps that are publicly available and (2) to assess the usability of asthma apps currently available on the market to identify content and features of apps associated with positive and negative user ratings.

Methods

Asthma App Selection

To examine the average user rating and reviews for asthma apps, the following inclusion criteria were applied for our search: apps must be available for download on Android or iOS platforms, written in English, available within the United States, able to be downloaded onto a smartphone or tablet, and had at least one update within the last 5 years. Apps must also have a primary focus on asthma, with respect to either asthma education or asthma self-management. Because of the potential for bias in ratings due to small sample size, only apps that had more than 10 written reviews were included. Apps were not excluded based on cost for use or intended audience. We started with a comprehensive list of apps that strictly matched our inclusion criteria. The list was expanded by using the following keywords commonly associated with asthma education or asthma management to identify additional asthma apps: *asthma management*, *asthma game*, *asthma quiz*, *peak flow*, and *asthma education*. All apps were evaluated by the research team based on inclusion criteria; of the 51 apps initially reviewed, 29 apps had been updated in the last 5 years. Of the 29 apps recently updated, 19 were excluded based on having less than 10 written reviews available. The final list of 10 apps included in the analysis is presented in Table 1.

XSL•FO
RenderX

Table 1. List of asthma mobile apps included in analysis.

Name	Developer	Android	iOS	Last update	Update times	Category
My Asthma App	Asthma and Respiratory Foundation NZ	N/A[a]	X	May 5, 2019	6	Education
Propeller	Reciprocal Labs	X	N/A	March 28, 2019	25	Management
SaniQ Asthma	Qurasoft GmbH	X	X	March 19, 2019	25	Management
Asthma Tracker	Kantonsspital Baselland	N/A	X	March 13, 2019	12	Management
AirCasting	HabitatMap	N/A	X	March 7, 2019	N/A	Environmental Data
Peak Flow	Ben Hills	X	N/A	April 26, 2018	N/A	Management
My Asthma Pal	Children's Medical Center of Dallas	X	X	March 7, 2019	6	Management
asthmaTrack	dangerDown LLC	N/A	X	January 18, 2018	28	Management
Breathcount asthma control	Segfoltas	X	N/A	January 9, 2017	N/A	Management
AsthmaMD	AsthmaMD, Inc.	X	X	March 10, 2017	21	Management

[a]N/A: Not available

User Ratings

To identify specific sentiments within language characteristics of user reviews that are associated with high or low app ratings, we first dichotomized the average user rating of the asthma app into 2 categories: a high average rating and a low average rating. Asthma apps with average ratings of 4 and above were categorized as having a high average rating. Asthma apps with average ratings of less than 4 were categorized as having a low average rating.

Sentiment Analysis of User Reviews: N-Gram models

An *N-gram* is a contiguous sequence of n items from a piece of article, sentence, or speech. An *N-gram model* is a probabilistic language model for predicting the next item given the sequence of *(n–1)–gram*, which simulates characteristics of the language used in a certain corpus. Its simplicity and scalability enable its application in sentiment analysis of the reviews [30]. For this analysis, we modeled both 2-word (bi-gram) and 3-word (tri-gram) phrases that commonly appeared across highly rated and lowly rated reviews. Because of the limited number of reviews, the tri-gram model did not yield expressive results, as it identified the same tri-gram *peak flow meter* in both the good and bad reviews. The bi-grams extracted from the model showed distinct patterns in good reviews and bad reviews. As such, results from bi-gram models for highly rated reviews and lowly rated reviews are discussed in the next section. We removed the stop words in the preprocessing step, and separated the reviews into good reviews (ratings >3) and bad reviews (ratings ≤3). Reviews were collected on June 23, 2020, and contain reviews from January 08, 2010, to June 21, 2020. Of the 10 apps that met the inclusion criteria, a total of 373 reviews were examined across all apps.

Results

Asthma App Descriptive Results

Of the 10 apps included in the analysis, 6 were available at no cost to users and 4 apps were available for purchase, with a maximum cost of US $2.99; 50% of apps (n=5) were updated in the past 2 years.

As shown in Figures 1-3, among apps reviewed, 53.4% (199/373) received high ratings (average ratings of 4 or 5) and 47.2% (176/373) received low ratings (average ratings of ≤3). The number of ratings across all apps ranged from 188 (AsthmaMD) to 10 (My Asthma App). Three of the apps were available on both Android and iOS platforms.

Stream graph of the number of review counts from 2010 to 2020 is shown in Figure 1. The number of reviews for each app is displayed in a distinct color to visually track the change in number of reviews over time. The height of the colored area suggests the number of reviews an app gets in the year. The width of the colored area suggests the total number of reviews an app gets over the years. As shown in Figure 1, the dominant app is AsthmaMD, which reached a peak number of reviews during 2012 to 2014.

XSL•FO
RenderX

Figure 1. Stream graph of the review counts for asthma apps from 2010 to 2020.

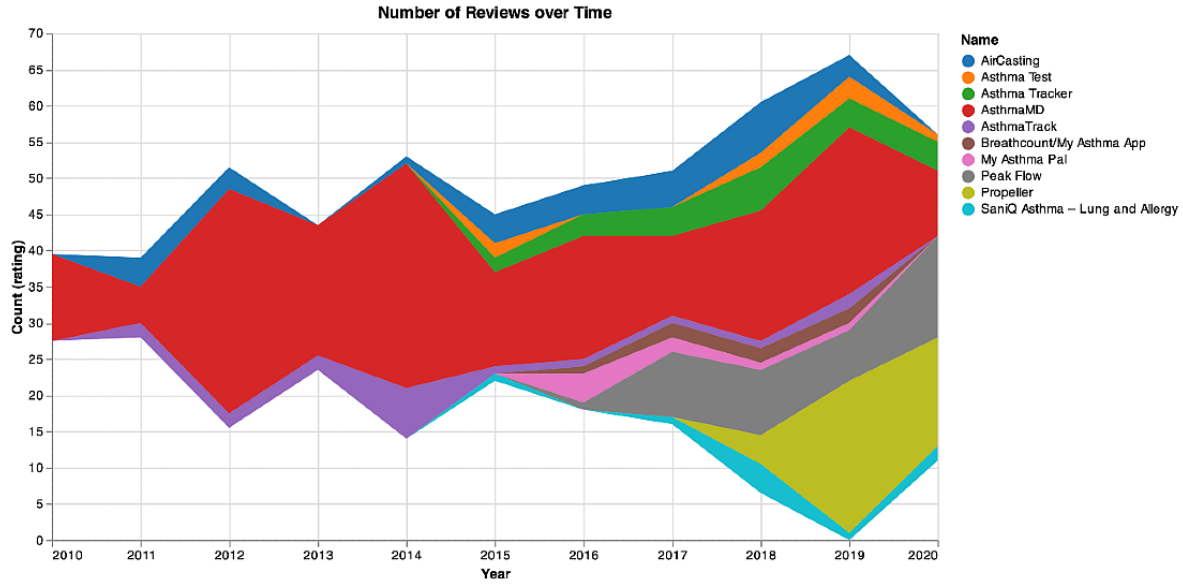
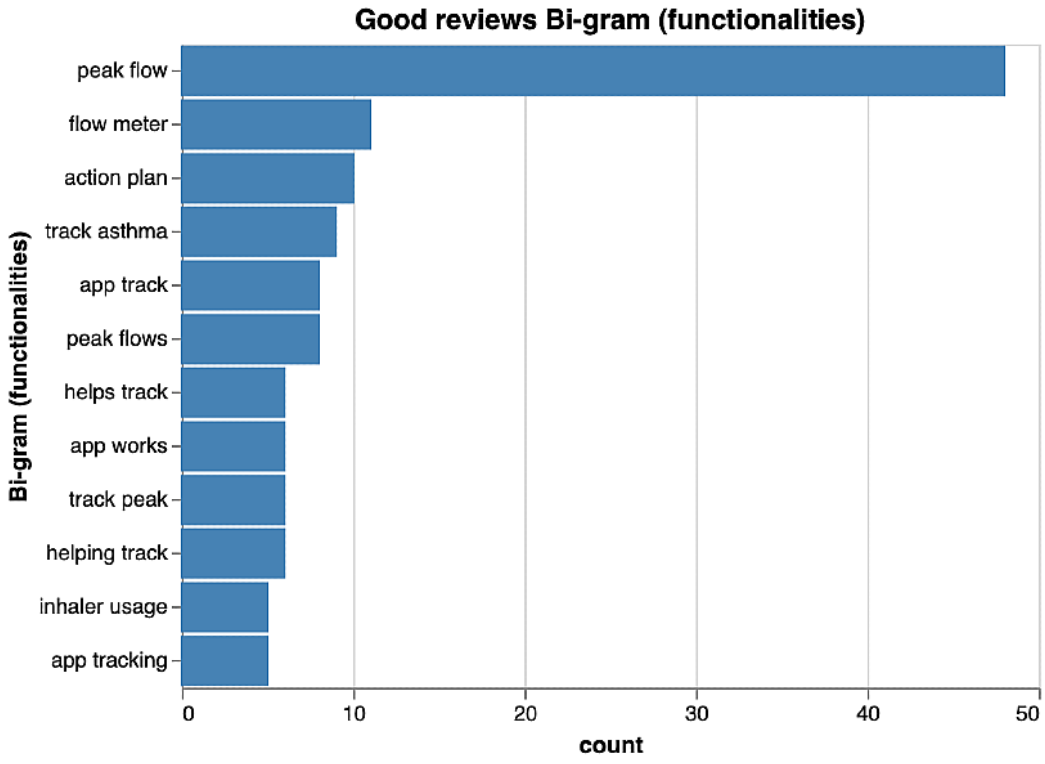
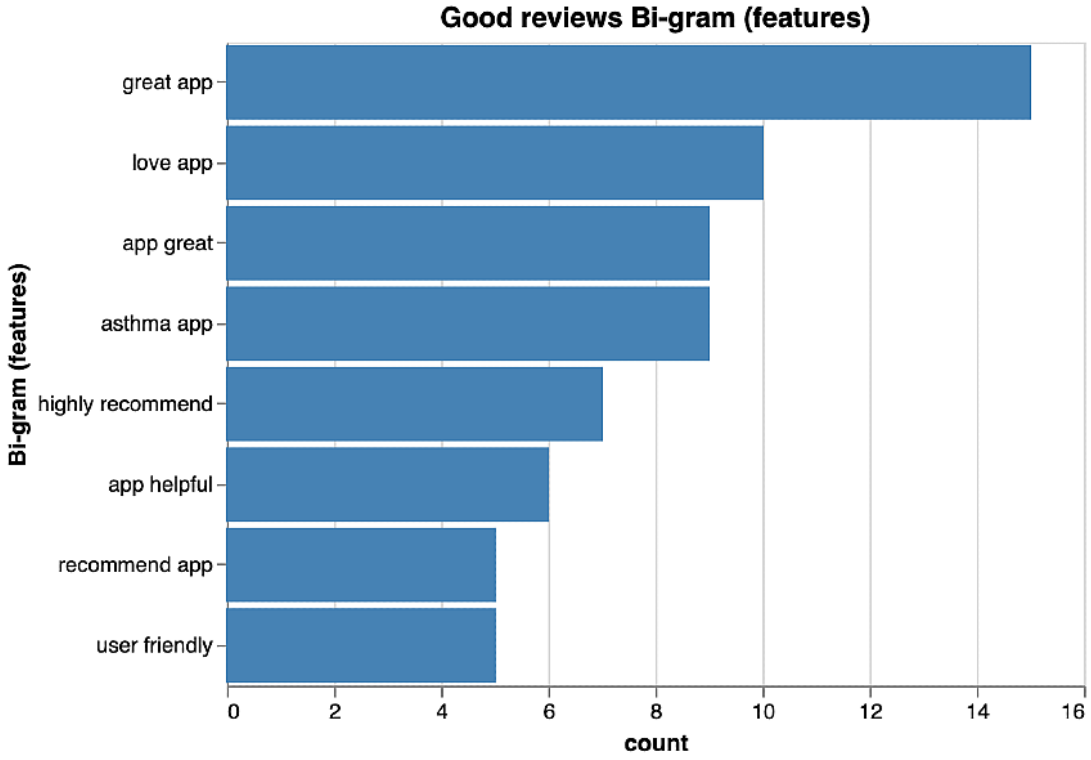


Figure 2. Bi-gram results of functionalities of highly rated asthma mobile apps.



XSL•FO
RenderX

Figure 3. Bi-gram results of features of highly rated asthma mobile apps.



Sentiment Analysis of User Reviews

Distinct patterns of language were observed when comparing highly rated and poorly rated apps. Among the highly rated reviews, bi-grams emerged with respect to features as well as functionality of the app. Figures 2 and 3 depict the frequencies of the most commonly occurring bi-grams among highly rated apps. Several key features of asthma management that were common among highly rated apps included the tracking of peak flow readings (n=48), asthma symptom monitoring (n=11), and action plans (n=10). Key features related to functionality that were common among highly rated apps included ease of use (n=5), as illustrated in Figure 2.

Among poorly rated asthma apps, bi-grams (Figures 4 and 5) predominantly focused on functionality issues encountered by users (Figure 4). The functionality keywords that had the most common occurrences in bad reviews were create account (n=8), lost data (n=6), open app (n=5). By identifying the most important words (or keywords) in negative reviews through the term frequency–inverse document frequency measure in Figure 6, we concluded that users most commonly reported loss of data (n=14) and crashing of app (n=12) as functionality issues among poorly rated asthma apps.

XSL•FO
RenderX

Figure 4. Bi-gram results of functionalities of poorly rated asthma mobile apps.

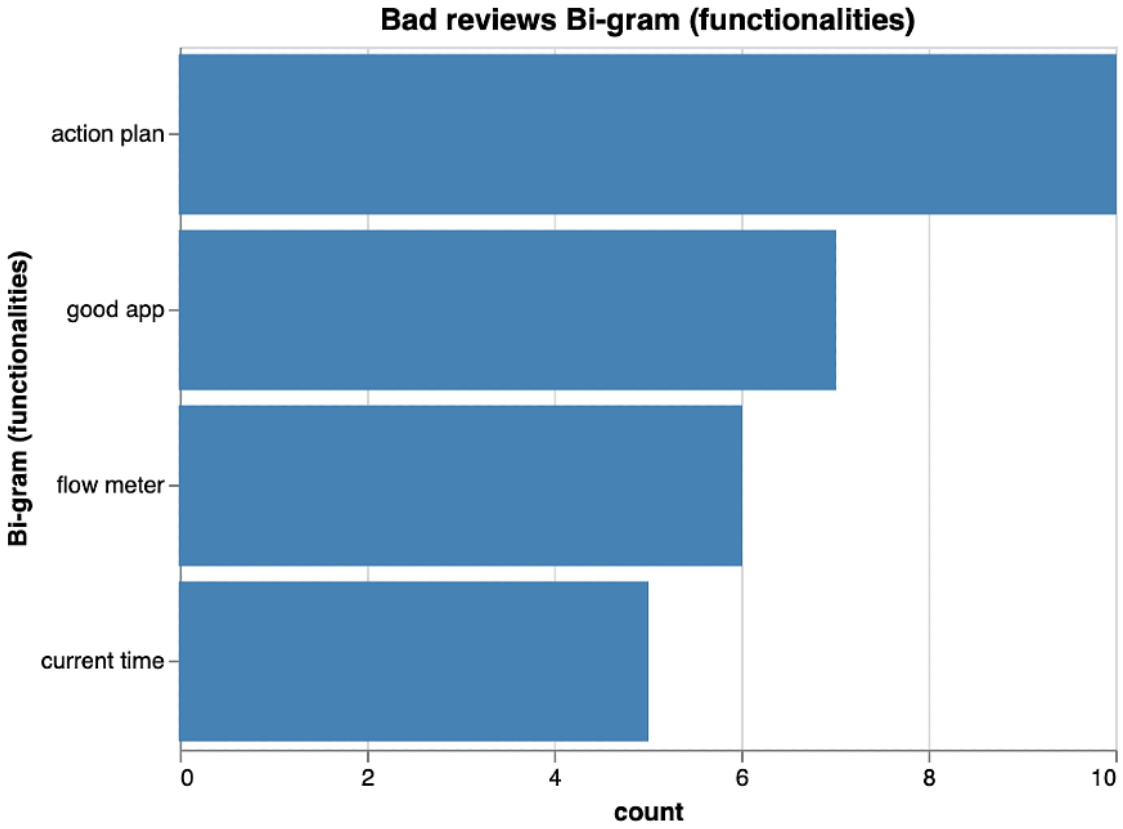
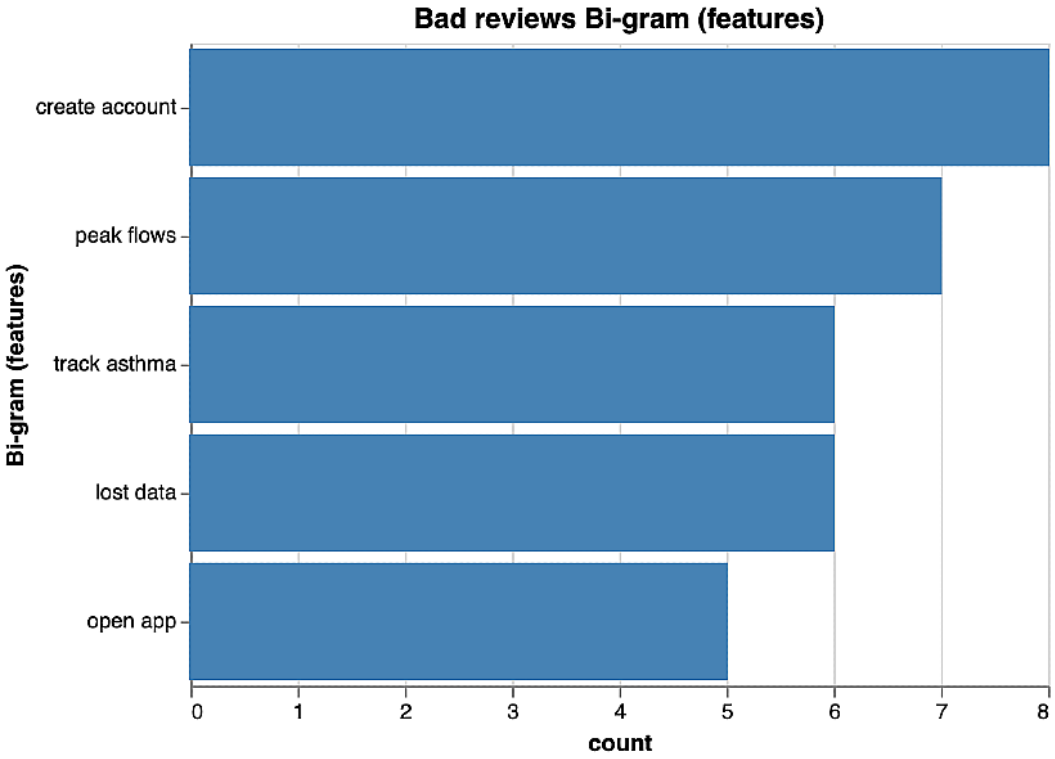
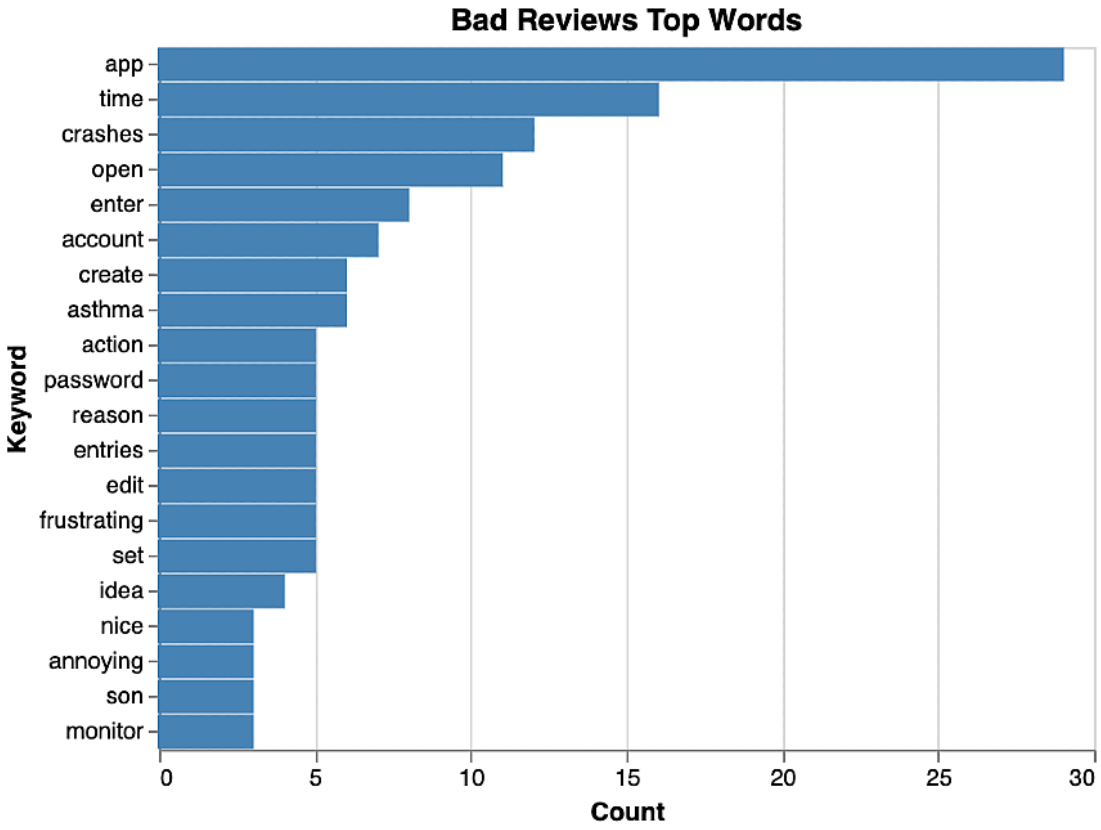


Figure 5. Bi-gram results of features of poorly rated asthma apps.



XSL•FO
RenderX

Figure 6. Top words in reviews of poorly rated asthma apps.



Discussion

Principal Results

This study analyzed ratings and features of publicly available asthma apps to identify user preferences. Our descriptive results confirmed those of prior studies which observed the limited availability of publicly available, up-to-date asthma apps on the market [26]. In terms of available functions, we observed that most asthma apps offered functions associated with several recommendations of effective self-management including asthma education, peak flow monitoring, and an action plan [31]. Consistent with prior reviews of asthma apps, our results confirmed that asthma apps offered either education or management, with few offering a combination of both [25].

Furthermore, our study results demonstrated that asthma app quality varied widely, ranging anywhere from an average user rating of 1.5 out of 5 to 5 out of 5. It also appeared that individual asthma apps varied in their user rating between Android and iOS platforms. These findings may call into question the long-term engagement with asthma apps, a crucial factor for determining their value [32,33]. In other domains such as diabetes self-management, researchers have observed that long-term engagement of app users is generally limited [34]. However, chronic diseases such as asthma require long-term self-management.

Recommendations for App Features and Functions

With the *N-grams* model, we were able to capture several characteristics related to features and functionality that users like or dislike about the apps; for example, *peak flow* and *action plan* were frequently mentioned among features, whereas *easy to use* and *doesn't work* were frequently mentioned around functionality. Previous studies of asthma apps confirmed peak flow tracking, symptom tracking, and medication tracking as some of the most common functions in highly rated apps [25]. Additional function recommendations which were not captured in our sentiment analysis but have been identified in previous studies of asthma apps include medical appointment tracking, health snapshots, and clinical asthma questionnaires [20,26,35]. Additional categories which have been identified through formative research of specific asthma apps include the incorporation of notification features such as those related to medication reminders, medical appointment reminders, and peak flow reading reminders [36-40].

One potential way to improve long-term engagement, which has been successfully applied to physical activity, is interactions with virtual coaches [41,42]. Thus, developers of upcoming asthma apps might consider the implementation of virtual coaches to enhance long-term engagement. Another potential method of improving long-term engagement, in particular among children and adolescents with asthma, could be through gamification and use of contingent rewards [43-45]. Lastly, application of behavior change techniques including specific goal setting, provision of performance feedback, and barrier identification may also improve acceptability and long-term engagement with asthma apps [46-48].

Limitations

Our study did have several limitations, which should be considered when interpreting its findings. The popularity of an asthma app may not yield information regarding the quality of

XSL•FO
RenderX

the asthma content presented in the app with respect to the clinical or scientific literature. Prior studies have observed variability in the quality of asthma information with respect to its consistency with National Asthma Education and Prevention Program guidelines [3]. Future studies should examine the concordance between asthma education and asthma management clinical tools within popular asthma apps.

Limitations of sentiment analysis include its sensitivity to review content; as such, user ratings may not always correspond to descriptive reviewer feedback. Further, displays of reviews were limited to downloads with a minimum threshold of 10 reviews. However, sentiment analysis has been increasingly applied across a variety of health outcomes [49-52]. Finally, this study did not include assessment of apps using a validated instrument (eg, Mobile Application Rating Scale). However, this type of analysis was conducted in a previous review of asthma apps [25].

Conclusions

Our study extends previous research in this field by focusing on the experiences and reviews of asthmatics' interactions with publicly available asthma apps. Our study results have several implications with respect to informing the development of asthma mobile apps and their recommendation for clinical use. As use of asthma apps have been found to have an impact on several clinical outcomes, including but not limited to control, quality of life, medication adherence, and patient-reported outcomes, improvements to asthma apps should include a focus on user-centered design and experiences. Combining big-data analytic approaches with qualitative data from users may yield additional insights to improve usability and long-term engagement with asthma apps. Further, collaborations between asthma app developers, clinicians, and researchers should include considerations regarding data security, privacy features, and sharing of personal health information which would also increase patient and provider confidence regarding recommending the use of asthma mobile apps to improve asthma self-management.

Acknowledgments

This study was supported by a City University of New York Interdisciplinary Research Grant to MC-R and AK. MC-R is supported by TRANSPORT – The Translational Program of Health Disparities Research Training (5S21MD012474-02).

Authors' Contributions

MC-R and HV led the drafting of the manuscript and supervised the process. HV, XH, and JL contributed to the data analysis. MC-R, AL, and AK contributed to the discussion of data.

Conflicts of Interest

None declared.

References

1. Centers for Disease Control and Prevention. Asthma Surveillance Data. Atlanta, GA: Centers for Disease Control and Prevention URL: https://www.cdc.gov/asthma/asthmadata.htm [accessed 2020-07-10]
2. Nurmagambetov T, Kuwahara R, Garbe P. The Economic Burden of Asthma in the United States, 2008-2013. Ann Am Thorac Soc 2018 Mar;15(3):348-356. [doi: 10.1513/AnnalsATS.201703-259OC] [Medline: 29323930]
3. Fletcher MJ, Tsiligianni I, Kocks JWH, Cave A, Chunhua C, Sousa JCD, et al. Improving primary care management of asthma: do we know what really works? NPJ Prim Care Respir Med 2020 Jun 17;30(1):29 [FREE Full text] [doi: 10.1038/s41533-020-0184-0] [Medline: 32555169]
4. Madhok N, Kipperman S, Tom C, Sin S, Rastogi D. Parental Perception and Knowledge of Asthma Severity and Management of Athmatic Children in the Bronx. Journal of Allergy and Clinical Immunology 2009 Feb;123(2):S161. [doi: 10.1016/j.jaci.2008.12.606]
5. McQuaid EL. Barriers to medication adherence in asthma: The importance of culture and context. Ann Allergy Asthma Immunol 2018 Jul;121(1):37-42. [doi: 10.1016/j.anai.2018.03.024] [Medline: 29580846]
6. McQuaid EL, Everhart RS, Seifer R, Kopel SJ, Mitchell DK, Klein RB, et al. Medication adherence among Latino and non-Latino white children with asthma. Pediatrics 2012 Jun;129(6):e1404-e1410 [FREE Full text] [doi: 10.1542/peds.2011-1391] [Medline: 22566417]
7. Tackett AP, Farrow M, Kopel SJ, Coutinho MT, Koinis-Mitchell D, McQuaid EL. Racial/ethnic differences in pediatric asthma management: the importance of asthma knowledge, symptom assessment, and family-provider collaboration. J Asthma 2020 Jul 01:1-12. [doi: 10.1080/02770903.2020.1784191] [Medline: 32546119]
8. Everhart RS, Koinis Mitchell D, Friedman D, Kopel S, Canino G, Fritz G, et al. Pediatric asthma management within Latino and non-Latino White families. Fam Syst Health 2014 Jun;32(2):167-175. [doi: 10.1037/fsh0000005] [Medline: 24611719]
9. Singh S, Surani S, McGuinness S, Eudicone J, Gilbert I, Subramanian S. Current practice patterns, challenges, and educational needs of asthma care providers in the United States. J Asthma 2020 May 19:1-10. [doi: 10.1080/02770903.2020.1761980] [Medline: 32336241]

XSL•FO
RenderX

10. Indraratna P, Tardo D, Yu J, Delbaere K, Brodie M, Lovell N, et al. Mobile Phone Technologies in the Management of Ischemic Heart Disease, Heart Failure, and Hypertension: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e16695 [FREE Full text] [doi: 10.2196/16695] [Medline: 32628615]
11. Tully L, Burls A, Sorensen J, El-Moslemany R, O'Malley G. Mobile Health for Pediatric Weight Management: Systematic Scoping Review. JMIR Mhealth Uhealth 2020 Jun 03;8(6):e16214 [FREE Full text] [doi: 10.2196/16214] [Medline: 32490849]
12. Koepp J, Baron MV, Hernandes Martins PR, Brandenburg C, Kira ATF, Trindade VD, et al. The Quality of Mobile Apps Used for the Identification of Pressure Ulcers in Adults: Systematic Survey and Review of Apps in App Stores. JMIR Mhealth Uhealth 2020 Jun 16;8(6):e14266 [FREE Full text] [doi: 10.2196/14266] [Medline: 32470916]
13. de Cock C, van Velthoven M, Milne-Ives M, Mooney M, Meinert E. Use of Apps to Promote Childhood Vaccination: Systematic Review. JMIR Mhealth Uhealth 2020 May 18;8(5):e17371 [FREE Full text] [doi: 10.2196/17371] [Medline: 32421684]
14. Lau N, O'Daffer A, Colt S, Yi-Frazier JP, Palermo TM, McCauley E, et al. Android and iPhone Mobile Apps for Psychosocial Wellness and Stress Management: Systematic Search in App Stores and Literature Review. JMIR Mhealth Uhealth 2020 May 22;8(5):e17798 [FREE Full text] [doi: 10.2196/17798] [Medline: 32357125]
15. Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile Apps for Mental Health Issues: Meta-Review of Meta-Analyses. JMIR Mhealth Uhealth 2020 May 29;8(5):e17458 [FREE Full text] [doi: 10.2196/17458] [Medline: 32348289]
16. Nuamah J, Mehta R, Sasangohar F. Technologies for Opioid Use Disorder Management: Mobile App Search and Scoping Review. JMIR Mhealth Uhealth 2020 Jun 05;8(6):e15752 [FREE Full text] [doi: 10.2196/15752] [Medline: 32501273]
17. Marcano BJS, Huckvale K, Greenfield G, Car J, Gunn LH. Smartphone and tablet self management apps for asthma. Cochrane Database Syst Rev 2013;11:CD010013. [doi: 10.1002/14651858.CD010013.pub2] [Medline: 24282112]
18. Mulvaney SA, Ho Y, Cala CM, Chen Q, Nian H, Patterson BL, et al. Assessing adolescent asthma symptoms and adherence using mobile phones. J Med Internet Res 2013;15(7):e141 [FREE Full text] [doi: 10.2196/jmir.2413] [Medline: 23864345]
19. Kagen S, Garland A. Asthma and Allergy Mobile Apps in 2018. Curr Allergy Asthma Rep 2019 Feb 02;19(1):6 [FREE Full text] [doi: 10.1007/s11882-019-0840-z] [Medline: 30712150]
20. Cook KA, Modena BD, Simon RA. Improvement in Asthma Control Using a Minimally Burdensome and Proactive Smartphone Application. J Allergy Clin Immunol Pract 2016;4(4):730-737.e1 [FREE Full text] [doi: 10.1016/j.jaip.2016.03.005] [Medline: 27107690]
21. Harris SK, Aalsma MC, Weitzman ER, Garcia-Huidobro D, Wong C, Hadland SE, et al. Research on Clinical Preventive Services for Adolescents and Young Adults: Where Are We and Where Do We Need to Go? J Adolesc Health 2017 Mar;60(3):249-260 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.10.005] [Medline: 28011064]
22. Rising CJ, Jensen RE, Moser RP, Oh A. Characterizing the US Population by Patterns of Mobile Health Use for Health and Behavioral Tracking: Analysis of the National Cancer Institute's Health Information National Trends Survey Data. J Med Internet Res 2020 May 14;22(5):e16299 [FREE Full text] [doi: 10.2196/16299] [Medline: 32406865]
23. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
24. Ramsey RR, Plevinsky JM, Kollin SR, Gibler RC, Guilbert TW, Hommel KA. Systematic Review of Digital Interventions for Pediatric Asthma Management. J Allergy Clin Immunol Pract 2020 Apr;8(4):1284-1293. [doi: 10.1016/j.jaip.2019.12.013] [Medline: 31870809]
25. Tinschert P, Jakob R, Barata F, Kramer J, Kowatsch T. The Potential of Mobile Apps for Improving Asthma Self-Management: A Review of Publicly Available and Well-Adopted Asthma Apps. JMIR Mhealth Uhealth 2017 Aug 02;5(8):e113 [FREE Full text] [doi: 10.2196/mhealth.7177] [Medline: 28768606]
26. Jácome C, Almeida R, Pereira AM, Araújo L, Correia MA, Pereira M, INSPIRERS group. Asthma App Use and Interest Among Patients With Asthma: A Multicenter Study. J Investig Allergol Clin Immunol 2020 Apr;30(2):137-140 [FREE Full text] [doi: 10.18176/jiaci.0456] [Medline: 32327403]
27. Johnson KB, Patterson BL, Ho Y, Chen Q, Nian H, Davison CL, et al. The feasibility of text reminders to improve medication adherence in adolescents with asthma. J Am Med Inform Assoc 2016 May;23(3):449-455. [doi: 10.1093/jamia/ocv158] [Medline: 26661717]
28. Mosnaim G, Li H, Martin M, Richardson D, Belice PJ, Avery E, et al. A tailored mobile health intervention to improve adherence and asthma control in minority adolescents. J Allergy Clin Immunol Pract 2015;3(2):288-290.e1 [FREE Full text] [doi: 10.1016/j.jaip.2014.10.011] [Medline: 25609351]
29. Real FJ, Beck AF, DeBlasio D, Zackoff M, Henize A, Xu Y, et al. Dose Matters: A Smartphone Application to Improve Asthma Control Among Patients at an Urban Pediatric Primary Care Clinic. Games Health J 2019 Oct;8(5):357-365. [doi: 10.1089/g4h.2019.0011] [Medline: 31157983]
30. Yazdani A, Safdari R, Golkar A, R Niakan Kalhori S. Words prediction based on N-gram model for free-text entry in electronic health records. Health Inf Sci Syst 2019 Dec;7(1):6. [doi: 10.1007/s13755-019-0065-5] [Medline: 30886701]
31. Wu AC. The Promise of Improving Asthma Control Using Mobile Health. J Allergy Clin Immunol Pract 2016;4(4):738-739. [doi: 10.1016/j.jaip.2016.04.003] [Medline: 27393782]

XSL•FO
RenderX

32. Tatara N, Arsand E, Skrøvseth SO, Hartvigsen G. Long-term engagement with a mobile self-management system for people with type 2 diabetes. JMIR Mhealth Uhealth 2013;1(1):e1 [FREE Full text] [doi: 10.2196/mhealth.2432] [Medline: 25100649]
33. Leigh S, Ashall-Payne L, Andrews T. Barriers and Facilitators to the Adoption of Mobile Health Among Health Care Professionals From the United Kingdom: Discrete Choice Experiment. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e17704 [FREE Full text] [doi: 10.2196/17704] [Medline: 32628118]
34. Adu MD, Malabu UH, Malau-Aduli AE, Drovandi A, Malau-Aduli BS. User Retention and Engagement With a Mobile App Intervention to Support Self-Management in Australians With Type 1 or Type 2 Diabetes (My Care Hub): Mixed Methods Study. JMIR Mhealth Uhealth 2020 Jun 11;8(6):e17802 [FREE Full text] [doi: 10.2196/17802] [Medline: 32525491]
35. Rhee H, Allen J, Mammen J, Swift M. Mobile phone-based asthma self-management aid for adolescents (mASMAA): a feasibility study. Patient Prefer Adherence 2014;8:63-72 [FREE Full text] [doi: 10.2147/PPA.S53504] [Medline: 24470755]
36. Farzandipour M, Nabovati E, Heidarzadeh Arani M, Akbari H, Sharif R, Anvari S. Enhancing Asthma Patients' Self-Management through Smartphone-Based Application: Design, Usability Evaluation, and Educational Intervention. Appl Clin Inform 2019 Oct;10(5):870-878. [doi: 10.1055/s-0039-1700866] [Medline: 31724144]
37. Roberts C, Sage A, Geryk L, Sleath B, Carpenter D. Adolescent Preferences and Design Recommendations for an Asthma Self-Management App: Mixed-Methods Study. JMIR Form Res 2018 Sep 13;2(2):e10055 [FREE Full text] [doi: 10.2196/10055] [Medline: 30684424]
38. Ramsey RR, Carmody JK, Holbein CE, Guilbert TW, Hommel KA. Examination of the uses, needs, and preferences for health technology use in adolescents with asthma. J Asthma 2019 Sep;56(9):964-972. [doi: 10.1080/02770903.2018.1514048] [Medline: 30207802]
39. Geryk LL, Roberts CA, Sage AJ, Coyne-Beasley T, Sleath BL, Carpenter DM. Parent and Clinician Preferences for an Asthma App to Promote Adolescent Self-Management: A Formative Study. JMIR Res Protoc 2016 Dec 06;5(4):e229 [FREE Full text] [doi: 10.2196/resprot.5932] [Medline: 27923777]
40. Hsia BC, Wu S, Mowrey WB, Jariwala SP. Evaluating the ASTHMAXcel Mobile Application Regarding Asthma Knowledge and Clinical Outcomes. Respir Care 2020 Jun 02. [doi: 10.4187/respcare.07550] [Medline: 32487751]
41. Ammari WG, Obeidat NM, Anani AR, AlKalbani RJ, Sanders M. ATTACHED, DETACHED and WITHOUT inhaler technique coaching tools to optimize pMDI use competence, asthma control and quality-of-life in asthmatic adults. J Thorac Dis 2020 May;12(5):2415-2425 [FREE Full text] [doi: 10.21037/jtd.2020.03.50] [Medline: 32642147]
42. Horwitz D, Kestenbom I, Goldbart A, Chechik T, Dizitzer Y, Golan-Tripto I. The effect of a coaching program on asthma control and health care utilization in children with asthma. J Asthma 2019 Oct 08:1-8. [doi: 10.1080/02770903.2019.1672721] [Medline: 31591919]
43. Hamari J, Koivisto J, Sarsa H. Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. New York: IEEE; 2014 Presented at: 2014 47th Hawaii International Conference on System Sciences; January 6-9, 2014; Waikoloa, HI p. 3025-3034 URL: https://people.uta.fi/~kljuham/2014-hamari_et_al-does_gamification_work.pdf [doi: 10.1109/hicss.2014.377]
44. Edwards EA, Caton H, Lumsden J, Rivas C, Steed L, Pirunsarn Y, et al. Creating a Theoretically Grounded, Gamified Health App: Lessons From Developing the Cigbreak Smoking Cessation Mobile Phone Game. JMIR Serious Games 2018 Nov 29;6(4):e10252 [FREE Full text] [doi: 10.2196/10252] [Medline: 30497994]
45. Rudin RS, Fanta CH, Predmore Z, Kron K, Edelen MO, Landman AB, et al. Core Components for a Clinically Integrated mHealth App for Asthma Symptom Monitoring. Appl Clin Inform 2017 Oct;8(4):1031-1043 [FREE Full text] [doi: 10.4338/ACI-2017-06-RA-0096] [Medline: 29241243]
46. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
47. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. JMIR Mhealth Uhealth 2016;4(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4461] [Medline: 27122452]
48. Al-Durra M, Torio M, Cafazzo JA. The use of behavior change theory in Internet-based asthma self-management interventions: a systematic review. J Med Internet Res 2015;17(4):e89 [FREE Full text] [doi: 10.2196/jmir.4110] [Medline: 25835564]
49. Gibbons J, Malouf R, Spitzberg B, Martinez L, Appleyard B, Thompson C, et al. Twitter-based measures of neighborhood sentiment as predictors of residential population health. PLoS One 2019;14(7):e0219550 [FREE Full text] [doi: 10.1371/journal.pone.0219550] [Medline: 31295294]
50. Mammen JR, Java JJ, Rhee H, Butz AM, Halterman JS, Arcoleo K. Mixed-methods content and sentiment analysis of adolescents' voice diaries describing daily experiences with asthma and self-management decision-making. Clin Exp Allergy 2019 Mar;49(3):299-307. [doi: 10.1111/cea.13250] [Medline: 30113733]
51. Liu J, Davidson E, Bhopal R, White M, Johnson M, Netto G, et al. Adapting health promotion interventions to meet the needs of ethnic minority groups: mixed-methods evidence synthesis. Health Technol Assess 2012;16(44):1-469 [FREE Full text] [doi: 10.3310/hta16440] [Medline: 23158845]

XSL•FO
RenderX

52. Liu S, Chen B, Kuo A. Monitoring Physical Activity Levels Using Twitter Data: Infodemiology Study. J Med Internet Res 2019 Jun 03;21(6):e12394 [FREE Full text] [doi: 10.2196/12394] [Medline: 31162126]

Abbreviations

mHealth: mobile health
NAEPP: National Asthma Education and Prevention Program

Edited by G Eysenbach, L Buis; submitted 17.06.19; peer-reviewed by Q Chen, R Gore; comments to author 15.05.20; revised version received 10.07.20; accepted 03.08.20; published 29.10.20.

Please cite as:
Camacho-Rivera M, Vo H, Huang X, Lau J, Lawal A, Kawaguchi A
Evaluating Asthma Mobile Apps to Improve Asthma Self-Management: User Ratings and Sentiment Analysis of Publicly Available Apps
JMIR Mhealth Uhealth 2020;8(10):e15076
URL: http://mhealth.jmir.org/2020/10/e15076/
doi:10.2196/15076
PMID:33118944

©Marlene Camacho-Rivera, Huy Vo, Xueqi Huang, Julia Lau, Adeola Lawal, Akira Kawaguchi. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 29.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Mobile App Specifically Designed to Facilitate Exercise in Parkinson Disease: Single-Cohort Pilot Study on Feasibility, Safety, and Signal of Efficacy

Merrill R Landers[1], PT, DPT, PhD; Terry D Ellis[2], PT, PhD

[1]Department of Physical Therapy, School of Integrative Health Sciences, University of Nevada, Las Vegas, Las Vegas, NV, United States

[2]Department of Physical Therapy and Athletic Training, College of Health & Rehabilitation Sciences: Sargent College, Boston University, Boston, MA, United States

Corresponding Author:
Merrill R Landers, PT, DPT, PhD
Department of Physical Therapy
School of Integrative Health Sciences
University of Nevada, Las Vegas
4505 Maryland Parkway, Box 453029
Las Vegas, NV, 89154
United States
Phone: 1 7028951377
Fax: 1 7028954883
Email: merrill.landers@unlv.edu

Abstract

Background: Many people with Parkinson disease do not have access to exercise programs that are specifically tailored to their needs and capabilities. This mobile app allows people with Parkinson disease to access Parkinson disease–specific exercises that are individually tailored using in-app demographic questions and performance tests which are fed into an algorithm which in turn produces a video-guided exercise program.

Objective: To test the feasibility, safety, and signal of efficacy of a mobile app that facilitates exercise for people with Parkinson disease.

Methods: A prospective, single-cohort design of people with Parkinson disease who had downloaded the 9zest app for exercise was used for this 12-week pilot study. Participants, who were recruited online, were encouraged to exercise with the full automated app for ≥150 minutes each week. The primary endpoints were feasibility (app usage and usability questions) and safety (adverse events and falls). The primary endpoints for signal of efficacy were a comparison of the in-app baseline and 8-week outcomes on the 30-second Sit-To-Stand (STS) test, Timed Up and Go (TUG) test, and the Parkinson's Disease Questionnaire 8 (PDQ8).

Results: For feasibility, of the 28 participants that completed the study, 12 participants averaged >150 minutes of app usage per week (3 averaged 120-150, 4 averaged 90-120, and 9 averaged less than 90 minutes). A majority of participants (>74%) felt the exercise was of value (16/19; 9 nonrespondents), provided adequate instruction (14/19; 9 nonrespondents), and was appropriate for level of function (16/19; 9 nonrespondents). For safety, there were no serious adverse events that occurred during the app-guided exercise. There were 4 reports of strain/sprain injuries while using the app among 3 participants, none of which necessitated medical attention. For signal of efficacy, there was improvement for each of the primary endpoints: STS (P=.01), TUG (P<.001), and PDQ8 (P=.01).

Conclusions: Independent, video-guided exercise using a mobile app designed for exercise in Parkinson disease was safe and feasible though there was variability in app usage. Despite this, the results provide evidence for a signal of efficacy as there were improvements in 3 of the 3 outcomes.

Trial Registration: ClinicalTrials.gov NCT03459586; https://clinicaltrials.gov/ct2/show/NCT03459586

(JMIR Mhealth Uhealth 2020;8(10):e18985) doi:10.2196/18985

KEYWORDS

Parkinson disease; smartphone; mobile phone; telehealth; telerehabilitation; digital health; physical therapy

XSL•FO
RenderX

Introduction

Exercise is an important therapy for people with Parkinson disease. It has a positive effect on physical capacity and physical/cognitive function, including improvements in gait, mobility, posture, and balance [1]. In addition, several meta-analyses and systematic reviews have concluded that exercise and physical therapy improve many Parkinson disease–specific motor and nonmotor symptoms [2-6]. Importantly, there are several lines of evidence, using rodent Parkinson disease models and in humans with Parkinson disease, that suggest a possible disease-modifying effect of exercise on Parkinson disease [7-22].

Despite the evidence revealing the benefits of exercise in Parkinson disease, many people with Parkinson disease do not participate regularly in exercise and the reasons for this are complex [23]. First, people with Parkinson disease have many of the same type of inherent intrinsic and extrinsic barriers that are present in healthy, older adults. These include a lack of motivation, finances, knowledge, skill, accessibility to exercise facilities, and transportation, among other things [24-26]. Many of these issues are compounded in Parkinson disease because of Parkinson disease–specific nonmotor symptoms, including fatigue [27] and depression [28]. In addition, low outcome expectation, lack of time, and fear of falling have been identified as important perceived barriers to exercise in people with Parkinson disease [29].

While exercise is not routinely recommended by neurologists in the early course of Parkinson disease treatment [30], neurologists are frequently the first to introduce the importance of exercise to people with Parkinson disease. Typically, the patient is referred to a physical therapist who will prescribe an exercise program that is scaled to the person's health and functional status. Ideally, people with Parkinson disease will continue to exercise independently or in a community program after they have finished their therapy and then return to their physical therapist every 6-12 months to recalibrate the exercise program to meet the challenges of new and increasing impairments brought on by disease progression. Unfortunately, many have difficulty sustaining regular engagement in an exercise program between physical therapy visits [24,31]. Therefore, a relatively large proportion fall back into previous habits and discontinue their exercise after discharge from supervised therapy. Subsequently, gained benefits are lost and prospective benefits of continuing exercise therapy are unrealized [32].

One solution that addresses many of the aforementioned barriers to sustained exercise participation is a recently developed, commercially available mobile app (9zest Parkinson's Therapy [33], a subsidiary of Moterum Technologies [34]), which was developed by a software engineer with Parkinson disease. This app uses self-report questions and app-guided performance tests to assess status at baseline and progress the exercise program over time. Using these self-report and performance metrics, the app constructs a customized exercise program using a proprietary algorithm. The app selects exercises from a library of exercise videos specifically designed for people with Parkinson disease and is calibrated to one's current level of function. The 9zest app also includes several behavior change techniques, including prompts/cues, goal setting, graded tasks, and performance feedback. Currently, there are a limited number of apps reported in the literature that promote exercise or physical activity for people with Parkinson disease [35,36].

The 9zest app specifically aims to mitigate many of the barriers to exercise participation. First, the app can be used in the privacy of one's home. Because prominent barriers to participation include lack of accessibility and limited time, this app allows those with transportation problems or those in rural areas and "medical/exercise deserts" to have access, at a convenient time, to a semicustomized exercise program for people with Parkinson disease. In addition, because it has periodic assessments, it helps improve participant motivation through goal setting and feedback [29]. Lastly, because the app is relatively low cost, it is ideal for financially disadvantaged populations. While the 9zest app appears promising, it has not been evaluated scientifically to determine its utility for people with Parkinson disease. Therefore, Aim 1 was to test the feasibility (adherence and user feedback) of using the app for independent exercise over 12 weeks for people with Parkinson disease. Aim 2 was to test the safety (adverse events and falls) and Aim 3 was to detect a signal of efficacy for lower extremity strength, functional gait, and quality of life in people with Parkinson disease. Identifying signals of efficacy is important in early studies, such as that described here, as they allow inference about proof of concept and whether continued scientific exploration using more rigorous clinical trials is warranted [37].

Methods

Study Design

A prospective, single-cohort design was implemented for this pilot study wherein participants who had downloaded the commercially available 9zest app were invited to participate in this study via an in-app message. Based on responses to the inclusion and exclusion criteria during in-app consenting, qualified participants were invited to complete the in-app baseline assessment (detailed later). These measures were used by the app to construct an individualized 12-week exercise program. All study-related tasks were conducted without direct contact from a member of the research team. Participants were instructed to be tested in the "on" Parkinson disease medication state for testing and app-guided exercise. For Aim 1 (feasibility), participation data (minutes of use) were recorded by the app and analyzed over the 12-week intervention period. Additionally, participants were asked questions about the usability of the app at the conclusion of the 12-week study. For Aim 2 (safety), adverse events data were tracked via an in-app question every 2 weeks of the 12-week study. Fall data were tracked via an in-app question after every exercise session. Outcomes for Aim 3 (signal of efficacy) were assessed at baseline, 8 weeks, and 12 weeks. However, the primary endpoint for the signal of efficacy was the 8-week measurement point as it was thought that waning adherence or more drop outs would occur over time; therefore, 8 weeks was chosen to optimize the ability to detect a signal of efficacy. The 12-week measurement was secondary.

XSL•FO
RenderX

Participants

People with Parkinson disease who had downloaded the commercially available app were automatically invited to participate in the study. Those meeting all inclusion criteria (English speaking, between 40 and 75 years old, self-report neurologist-diagnosed Parkinson disease, willingness to participate in a 12-week study, able to stand unassisted for 10 minutes, and stable on Parkinson disease medications and deep brain stimulation for 3 months prior to participation) and not having any of the exclusion criteria (diagnosed with dementia; comorbidities that would preclude exercise participation or increase participant risk [eg, severe osteoarthritis/pain, stroke, severe respiratory problems, traumatic brain injury, neuromuscular disease, atrial fibrillation, poorly controlled cardiovascular disease, limb amputation, osteoporosis]; vision or hearing impairment that would interfere with app use; fall that required physician evaluation [emergency visit, urgent care, or hospitalization] within the past year; use of an assistive device for walking; and currently exercising more than 60 minutes per week on average) were invited to advance to consent using an online form. This study was approved by the University of Nevada, Las Vegas Institutional Review Board. Participants were recruited online from a sample of convenience from July 2018 to May 2019, clinical trial registry (clinicaltrials.gov [NCT03459586]) using social media advertisements, health care provider referrals, and snowball recruitment strategies wherein participants were encouraged to recruit other participants among their acquaintances.

Sample Size Calculation

The sample size was estimated using Aim 3 and was calculated with the "paired *t* tests using effect size module" in PASS 19.0 [38] (NCSS, LLC). The sample size estimation was based on an effect size between 0.39 and 0.52 for the 30-second Sit-To-Stand (STS) test from another exercise study for people with Parkinson disease [39]. The analysis was based on a one-sided test at 80% power with the significance level at .05 and a 15% dropout rate; this estimation was conservatively based on the same exercise study for people with Parkinson disease that had an 11.1% dropout rate [39]. With dropouts, a sample size range between 30 (effect size=0.52) and 53 participants (effect size=0.39) was estimated.

Outcomes

Primary and secondary endpoints were all assessed via the app. Principal analyses took place at the 8-week point and exploratory analyses took place at the 12-week point. The primary endpoint of feasibility (Aim 1) was assessed by in-app tracking of app exercise in minutes. Usability was assessed using the following statements at the conclusion of the 12-week study with a 5-point Likert scale (5=strongly agree, 4=agree, 3=neutral, 2=disagree, and 1=strongly disagree): (1) *I believe this activity could be of some value to me*; (2) *I thought the app provided adequate instruction on the exercises*; (3) *I enjoyed using the app to help me exercise*; (4) *I would recommend this app to others with Parkinson disease*; and (5) *The exercises were appropriate for my level of functioning*.

The primary endpoint of safety (Aim 2) was assessed by tracking adverse events (assessed every 2 weeks over 12 weeks with a short in-app question asking if they have experienced any app exercise–related adverse events). Participants were asked questions about the occurrence of the following while doing the app-guided exercises: strains/sprains, chest pain, shortness of breath, and dizziness. If an event was reported, participants were asked follow-up questions to record the details (ie, pain rating, if it required medical attention, outcome). Safety was also measured by asking about the occurrence of a fall/s and other medical issues/events that occurred during the exercise (assessed after every app exercise session using an in-app question). In addition, at the conclusion of the 12-week study participants were asked the following statement with a 5-point Likert scale (strongly agree, agree, neutral, disagree, and strongly disagree): *I felt safe doing the exercises using the app*.

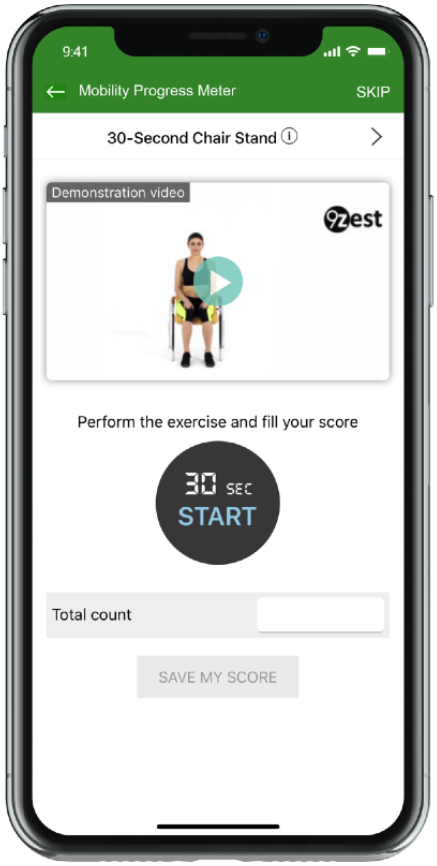
The following endpoints for Aim 3 (signal of efficacy) were assessed at baseline, and at the 8-week (primary) and 12-week (secondary) measurement points as part of the in-app assessment:

- 30-second STS, which is a measure of functional lower extremity strength with good reliability in people with Parkinson disease [40,41] and a minimal detectable change (MDC) of 3 [40];
- Timed Up and Go (TUG), which is a test of mobility and dynamic balance with evidence for good reliability in people with Parkinson disease [42-44] and an MDC of 4.85 [45]; and
- Parkinson's Disease Questionnaire 8 (PDQ8), which is a shortened version of the PDQ39, a Parkinson disease–specific quality of life measure, which has good reliability in people with Parkinson disease [46-49] and an MDC of 5.43 [45].

The STS and TUG tests were preceded by a demonstration video and an explanation of the test prior to the assessment using a built-in timer (Figure 1). In addition to the STS, TUG, and PDQ8, the Global Rating of Change score was asked at the conclusion of the 12-week study.

XSL•FO
RenderX

Figure 1. An example (30-second Sit-To-Stand test) of the in-app performance-based assessment used as a primary endpoint for the study and also used by the 9zest Smart to construct the exercise routine.

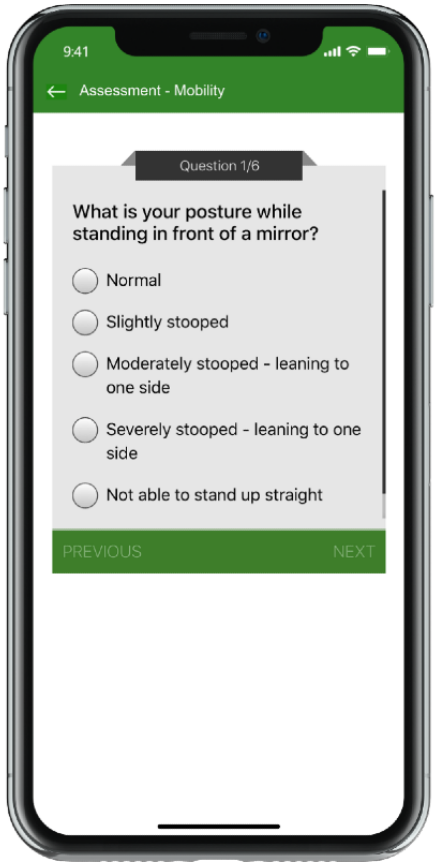


Exercise Intervention

Participants were encouraged to use the app on their phone or tablet for at least 150 minutes per week (eg, 3-5 days per week at 30-60 minutes each session) for 12 weeks. After registering on the app and completing several self-report Likert scale questions (ie, standing posture, tremors, balance, fall history, turning in bed, body movements with activities of daily living; Figure 2) and performance-based assessments (ie, STS, timed 360 degree turn, TUG, timed shirt removal), the 9zest app, using a proprietary algorithm (9zest Smart), constructed an exercise regimen for the participant's level of function. The customized exercise regimen was constructed from a library of over 1000 original exercise videos, developed by physical therapists. The library consisted of exercise in the following categories: aerobic, strengthening, balance, yoga based, range of motion/stretching, meditation based, and speech therapy exercises. However, because mobility was assigned as the in-app goal for each participant, the 9zest Smart algorithm drew exercises from the strengthening, balance, and stretching categories. The 9zest Smart, the app's intelligent engine, determined the regimen and levels of exercises for a user based on the most recent assessment (Figures 3 and 4). Thus, the app chooses the exercise program based on the primary goal (ie, mobility) and then uses the responses from self-report questions and the data from performance-based tests to determine the severity of the Parkinson disease. From this information, the app selects exercises that are consistent with the primary goal and at the appropriate level of function based on one's severity of Parkinson disease. At preset intervals (generally after 2 weeks), the app reassessed functional capacity using the same self-report questions and performance-based measures. The 9zest Smart algorithm adjusted the type, duration, and intensity of each exercise in the constructed exercise regimen. All of the exercises and dosing features are consistent with contemporary, clinical-based physical therapy practice as determined by the physical therapy development team who helped build the exercise library.

XSL•FO
RenderX

Figure 2. An example of an in-app self-report question that is used in the algorithm to construct the exercise program.



Because app-guided exercises are typically performed independently, safety was a primary concern in the development. All exercises in the library were deemed safe for people with Parkinson disease by the physical therapy development team for level of impairment and function. However, because this had not been scientifically vetted, it became a primary aim of this study. An additional safety feature was an audiovisual demonstration of each exercise to ensure that the exercise was performed with the proper technique. After the demonstration, the participant would then follow along with the audiovisual of the exercise in real time (Figures 3 and 4). If participants were deemed a fall risk by the in-app assessment, then all balance exercises would be scaled to level of function. For example, the balance exercise may be performed in sitting or at a counter/chair with hand contact (Figure 5). The exercise library consists of many variations of the same exercise (eg, standing, standing with support, sitting and lying down) and the decision on which exercise is selected for the exercise regimen is determined by the in-app assessment. Thus, the 9zest Smart algorithm would select an appropriate exercise variation with the intent of minimizing the risk for an injury or a fall. Additionally, before the session started, a warning message was displayed asking the participant to skip any exercises that might cause any pain or imbalance. Lastly, the in-app instruction reminded participants to work within their own limits to avoid any injury or fall.

XSL•FO
RenderX

Figure 3. An example of an exercise program.

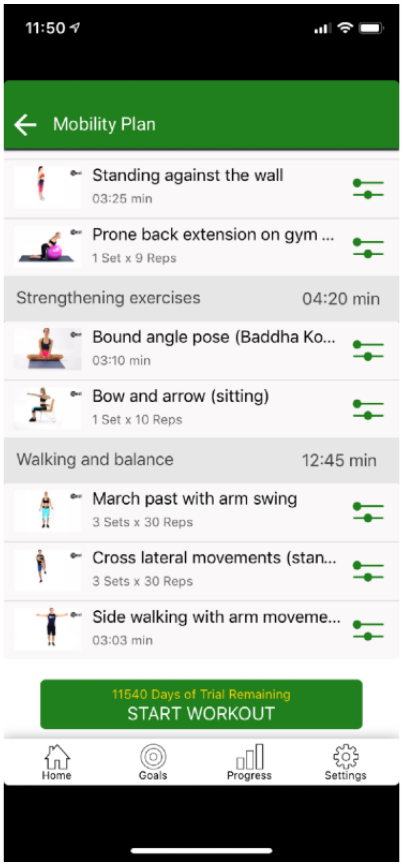
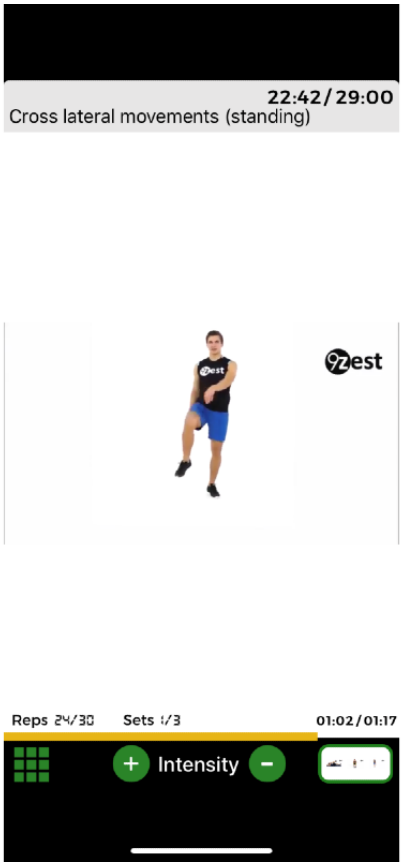
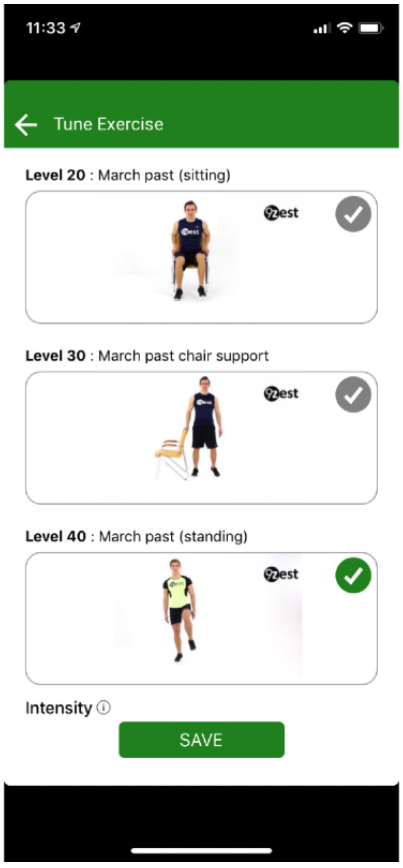


Figure 4. An example of a specific app-guided exercise (b).



XSL•FO
RenderX

Figure 5. An example of graded exercises based on level of function.



Data Analysis

Data were analyzed using SPSS version 24.0 (IBM SPSS Statistics for Windows) and α=.05 (where α is a priori level of significance). For Aim 1 (feasibility), basic descriptive statistics were used for in-app tracked exercise minutes and participants' responses to Likert scale questions. For Aim 2 (safety), basic descriptive statistics were used for the number of adverse events, falls, and a Likert-style question on whether the app was safe. For Aim 3 (signal of efficacy), the primary endpoint of the pre and 8-week outcome measures were analyzed using paired *t* tests to determine if app-guided exercise improved outcomes over the first 8 weeks of the study. In addition, the number of participants who improved beyond the MDC was tabulated for each of the primary endpoints. For the secondary endpoint of the 12-week outcomes and to determine if improvement continued after the 8-week measurement point, a one-way repeated measures ANOVA was used to compare the pre, 8-week, and 12-week outcomes. Descriptive statistics were used to report the results of the Global Rating of Change responses.

To determine if there was a dosing effect, those who used the app as instructed (≥150 minutes per week) were compared with those who did not using a 2 (≥150 minute average per week: yes or no) × 2 (time: premeasurement, 8-week measurement) mixed factorial ANOVA. To explore this further, Pearson correlational coefficient analyses were performed to determine if the minutes of app use was associated with overall improvement (8-week measurement – premeasurement) for each of the 3 primary endpoints (ie, STS, TUG, and PDQ8).

Results

Recruitment

A total of 28 participants (mean age in years 62.1 [SD 9.6], mean years since diagnosis 3.3 [SD 2.5]; males=6; females=14, and unknown=8) completed the 12 weeks of the study (Figure 6). There were no statistically significant differences between the dropouts and participants that completed the trial for age, sex, and years since diagnosis (Table 1).

XSL•FO
RenderX

Figure 6. CONSORT flow diagram of study participants.

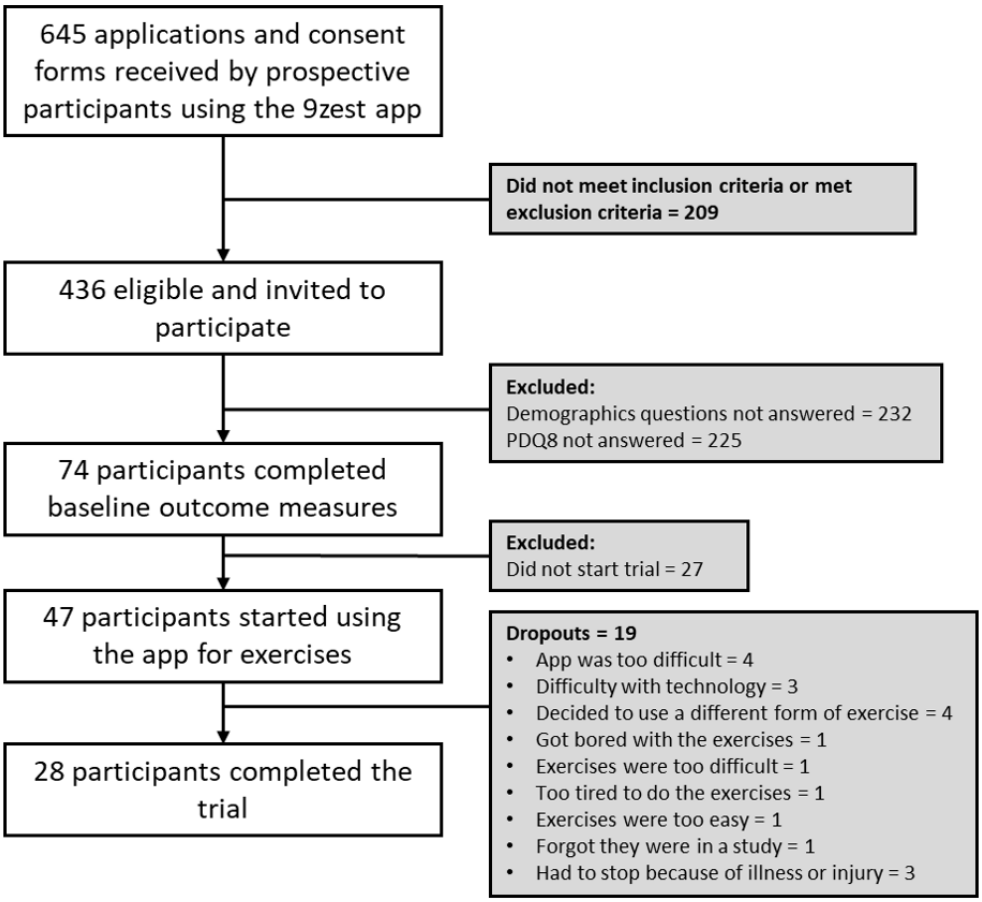


Table 1. Means, proportions, and statistical comparisons for participant demographics for those who completed the trial and dropouts.

Demographics	Participants who completed the trial	Dropouts	*P* value
N	28	19	NA[a]
Age (years), mean (SD)	62.1 (9.6)	59.7 (8.6)	.42
Sex			.13
Male, n	6	2	
Female, n	14	17	
Not reported, n	8	0	
Years since diagnosis, mean (SD)	3.3 (2.5)	4.7 (4.4)	.24

[a]NA: not applicable.

Aim 1 (Feasibility)

As much as 12 of the 28 participants averaged more than 150 minutes of app usage per week. The remaining 16 participants averaged the following: 120-150 minutes (n=3), 90-120 (n=4), and <90 minutes (n=9). Of the 19 respondents (9 did not respond), a majority felt the app exercise was of value (47% [9/19] strongly agreed, 37% [7/19] agreed, 5% [1/19] neutral, 11% [2/19] strongly disagreed), provided adequate instruction (53% [10/19] strongly agreed, 21% [4/19] agreed, 11% [2/19] neutral, 5% [1/19] disagreed, 11% [2/19] strongly disagreed), was enjoyable (47% [9/19] strongly agreed, 26% [5/19] agreed, 16% [3/19] neutral, 11% [2/19] strongly disagreed), would recommend to other people with Parkinson disease (53% [10/19] strongly agreed, 26% [5/19] agreed, 5% [1/19] neutral, 16% [3/19] strongly disagreed), and was appropriate for my level of function (37% [7/19] strongly agreed, 47% [9/19] agreed, 5% [1/19] disagreed, 11% [2/19] strongly disagreed).

Aim 2 (Safety)

There were no reported bouts of dizziness, falls, chest pain, or other medical events/issues during the app-guided exercise but there was 1 report of shortness of breath that resolved without the need for medical attention. There were 4 reports of strain/sprain injuries (3 participants: low back pain and 3 episodes of knee pain), none of which necessitated medical attention. Of the 19 respondents (9 did not respond), a majority felt the app exercise was safe (63% [12/19] strongly agree, 21% [4/19] agree, 5% [1/19] neutral, and 11% [2/19] strongly disagree).

XSL•FO
RenderX

Aim 3 (Signal of Efficacy)

There was a statistically significant improvement for each of the 3 primary endpoints. The STS improved from the premeasurement, mean 11.6 (SD 4.0), to the 8-week measurement, mean 14.3 (SD 4.7; *P*=.01; Hedges g=0.59; 95% CI 0.16-1.04). At the 8-week point, 15/28 (54%) improved beyond the MDC on the STS. The TUG improved from the premeasurement, mean 11.2 (SD 3.9), to the 8-week measurement, mean 8.5 (SD 2.6; *P*<.001; Hedges g=.80; 95% CI 0.46-1.18). At the 8-week point, 8/28 (29%) improved beyond the MDC on the TUG. Lastly, the PDQ8 improved from the premeasurement, mean 6.8 (SD 5.0), to the 8-week measurement, mean 4.1 (SD 5.0; *P*=.01; Hedges g=0.53; 95% CI 0.14-0.94). At the 8-week point, 6/28 (21%) improved beyond the MDC on the PDQ8.

At the conclusion of the 12-week study, 63% (12/19, 9 nonrespondents) felt their condition was better using the Global Rating of Change score. The results of the ANOVAs (premeasurement, 8-week, and 12-week measurements) suggest that there were no additional improvements from the 8-week to the 12-week measurement points for the STS (*P*>.99), TUG (*P*>.99), and PDQ8 (*P*=.94). There were no statistically significant interactions for the factorial ANOVAs to test dosing effect: STS (*P*=.39), TUG (*P*=.41), and PDQ8 (*P*=.86). Likewise, there were no statistically significant correlations for the average time of app exercise usage and change scores on the STS (r=–0.148, *P*=.45), TUG (r=0.113, *P*=.57), and PDQ8 (r=–0.017, *P*=.93).

Discussion

Principal Findings

The main purpose of this study was to test the feasibility, safety, and signal of efficacy of the 9zest app for people with Parkinson disease. Results of this study suggest that independent, video-guided exercise using the 9zest mobile app, designed for exercise for people with Parkinson disease, may be safe and feasible with considerable variability in app usage. A majority of participants felt the app-guided exercise was of value and enjoyable, and would recommend it to other people with Parkinson disease. A majority of participants (16/19, 84%) felt the app-guided exercise was not only safe but also appropriate for their current level of function. Despite the variability in app usage, the results suggest a signal of efficacy as there were statistically significant improvements on all 3 outcome measurements (*P*=.01 for STS; *P*<.001 for TUG; and *P*=.01 for PDQ8). Additionally, a majority of participants (12/19, 63%) felt that their condition had improved over the 12 weeks of the study. Based on these promising data, the 9zest app may be a safe, feasible, and useful technology for people with Parkinson disease as an adjuvant to a formalized physical therapy program or for those who wish to exercise independently or who do not have access to a physical therapist or Parkinson disease–specific exercise instruction because they live in a rural or underserved community. However, caution is warranted as larger, well-controlled trials are needed to draw more definitive conclusions. While the study was not designed to make cause and effect inferences, the results suggest an association between exercising via the app and improvement in lower extremity strength (STS), dynamic balance and mobility (TUG), quality of life (PDQ8), and overall improvement (Global Rating of Change). These results are consistent with a beneficial effect of independent exercise using other smartphone apps in other clinical populations [50].

It is important to note that the design of this study was such that participants did not have any contact with members of the research team and all assessments and training were done with in-app programming. This was certainly advantageous from a research resource perspective and it does mimic real-world use of this commercially available app; however, the lack of direct researcher contact/interaction and the requests to carry out assessments may have contributed to the poor on-boarding rate (Figure 6). The major reason for the poor onboarding rate may be that the request to provide demographic information and to perform assessments may have deterred participation as 436 prospective participants expressed interest in participating in the study but only 74 participants provided demographic data and completed outcome measures. This may be due to participant burden or concerns related to privacy (eg, not wanting to share demographic information). Losing 27 after completing assessments (ie, not starting the intervention) may be due to lack of interest or lack of follow through. The lack of interaction with a health care professional or coach may have contributed to the poor on-boarding and retention/adherence rates.

It is noteworthy that only 43% (12/28) of the participants reached the target of at least 150 minutes of app-guided exercise per week. There was considerable variability in usage of the app for exercise. Because the app was used independently with no supervision, it is possible that some participants may have had the app on and running but were not actually performing the exercise along with the video, thereby inflating participation rates. However, due to the study design, there is no way of knowing whether this occurred or not. Additionally, there is no way of knowing if participants “dropped-in” to another exercise program during the 12-week study. Regardless, 2 analyses (mixed factorial ANOVAs and correlational analyses) were conducted to determine if there was a dose effect of app usage on the outcome measures. The results of these analyses did not indicate that there was a dose effect and the very low correlations support this notion despite the fact that other exercise and smartphone app studies have linked exercise exposure to efficacy [51]. Further investigation into the dosing of the 9zest app-guided exercise using a more rigorously controlled design is warranted.

Based on the results of Aim 3, it is clear that the associations of improvement in the 3 main outcomes measures occurred during the first 8 weeks of the study. There were no statistically significant improvements from the 8-week measurement point to the 12-week measurement point for any of the outcome measures. This suggests that the improvement may have plateaued by the 8-week measurement point. However, the Global Rating of Change question, asked at the 12-week measurement point, indicates that a majority felt their condition had improved over the duration of the study.

XSL•FO
RenderX

Behavior change elements integrated into the app may have helped adherence to the app-guided exercise. For instance, incorporating remote Parkinson disease–specific peer coaching [52] or remote supervision by a physical therapist [53] to promote app use may be a helpful way to encourage adherence and promote accountability [52]. While the app currently has several behavior change techniques, including prompts/cues, goal setting, graded tasks, and outcome feedback, to promote exercise that are consistent with other apps for exercise [54,55], these may have been less personalized compared with approaches implemented by someone trained to address behavior change like a physical therapist. By contrast, it is possible that the behavior change elements in the study were ineffective at promoting adherence to the research study or the exercise program. It is also plausible that the app was not engaging enough or lacked the optimized motivational prompts to promote adherence.

Limitations

The most prominent limitation of the study was the poor onboarding rate entering the study. There was a poor yield from those who were invited to participate (74/436, 16.9%) and a high dropout rate once participants had completed the baseline measurements and started the exercise (19/47, 40%). These challenges are not unique to this study as other app-based studies have also had similar struggles with recruitment/yield and dropout [56]. As much as 7 of the 19 dropouts had problems with the app or technology, which may be an inherent app-related problem or simply a challenge for older adults using unfamiliar technology. This suggests that some human interaction (eg, a remote health care professional/coach) in studies like this, especially in the early stages of research, may be important design elements to promote better onboarding and adherence. The poor yield and high dropout rate resulted in a considerably smaller sample size than was estimated and was also smaller than the a priori sample size estimate. The study participants were also a sample of convenience which may have resulted in a biased sample of those comfortable with the technology and those who were already motivated to exercise. While the sample size was sufficient for the primary endpoints of the study, it was likely underpowered for the Aim 3 factorial ANOVAs. It is also important to note that low levels of app use do not necessarily equate to actual exercise levels as it is possible that participants were exercising without the app. Another limitation of the study was the remote testing of performance-based measures such as the STS and TUG, which have not been psychometrically vetted using a remote test. Two participants strongly disagreed that the app was safe. In fact, these same 2 participants strongly disagreed on every Likert question despite both being regular users of the app and both improving over the course of the study. There were no consistent characteristics, patterns, or themes in these participants' data. It is possible that they may have misinterpreted the direction of the scale. Lastly, adverse events were only asked every 2 weeks and this length of time may have increased the chance for recall bias (ie, under- or over-reporting because of poor memory).

Conclusion

For those participants who completed this study, independent, video-guided exercise using the 9zest mobile exercise app for people with Parkinson disease was safe and feasible and a majority of participants felt the app-guided exercise was enjoyable, provided adequate instruction, and would recommend it to others. While this study was not designed to determine cause and effect, the results provide evidence for a signal of efficacy as there were improvements in lower extremity functional strength, mobility and dynamic balance, and quality of life after 8 weeks of participation which were sustained at 12 weeks. The poor onboarding and adherence may suggest a limited generalizability only to those that are able to interact successfully with the technology.

Conflicts of Interest

Both authors are currently serving on the medical advisory board of 9zest.

References

1. Lauzé M, Daneault J, Duval C. The Effects of Physical Activity in Parkinson's Disease: A Review. J Parkinsons Dis 2016 Oct 19;6(4):685-698 [FREE Full text] [doi: 10.3233/JPD-160790] [Medline: 27567884]
2. Tomlinson CL, Patel S, Meek C, Herd CP, Clarke CE, Stowe R, et al. Physiotherapy versus placebo or no intervention in Parkinson's disease. Cochrane Database Syst Rev 2013 Sep 10(9):CD002817. [doi: 10.1002/14651858.CD002817.pub4] [Medline: 24018704]
3. Goodwin VA, Richards SH, Taylor RS, Taylor AH, Campbell JL. The effectiveness of exercise interventions for people with Parkinson's disease: a systematic review and meta-analysis. Mov Disord 2008 Apr 15;23(5):631-640. [doi: 10.1002/mds.21922] [Medline: 18181210]
4. Mehrholz J, Kugler J, Storch A, Pohl M, Hirsch K, Elsner B. Treadmill training for patients with Parkinson's disease. Cochrane Database Syst Rev 2015 Sep 13(9):CD007830. [doi: 10.1002/14651858.CD007830.pub4] [Medline: 26363646]
5. Tambosco L, Percebois-Macadré L, Rapin A, Nicomette-Bardel J, Boyer F. Effort training in Parkinson's disease: a systematic review. Ann Phys Rehabil Med 2014 Mar;57(2):79-104 [FREE Full text] [doi: 10.1016/j.rehab.2014.01.003] [Medline: 24582335]
6. Fayyaz M, Jaffery SS, Anwer F, Zil-E-Ali A, Anjum I. The Effect of Physical Activity in Parkinson's Disease: A Mini-Review. Cureus 2018 Jul 18;10(7):e2995 [FREE Full text] [doi: 10.7759/cureus.2995] [Medline: 30245949]

XSL•FO
RenderX

7. Hirsch MA, Farley BG. Exercise and neuroplasticity in persons living with Parkinson's disease. Eur J Phys Rehabil Med 2009 Jun;45(2):215-229 [FREE Full text] [Medline: 19532109]
8. Hirsch MA, Iyer SS, Sanjak M. Exercise-induced neuroplasticity in human Parkinson's disease: What is the evidence telling us? Parkinsonism Relat Disord 2016 Jan;22 Suppl 1:S78-S81. [doi: 10.1016/j.parkreldis.2015.09.030] [Medline: 26439945]
9. Faherty CJ, Raviie Shepherd K, Herasimtschuk A, Smeyne RJ. Environmental enrichment in adulthood eliminates neuronal death in experimental Parkinsonism. Brain Res Mol Brain Res 2005 Mar 24;134(1):170-179. [doi: 10.1016/j.molbrainres.2004.08.008] [Medline: 15790541]
10. Cohen AD, Tillerson JL, Smith AD, Schallert T, Zigmond MJ. Neuroprotective effects of prior limb use in 6-hydroxydopamine-treated rats: possible role of GDNF. J Neurochem 2003 Apr;85(2):299-305 [FREE Full text] [doi: 10.1046/j.1471-4159.2003.01657.x] [Medline: 12675906]
11. Tillerson JL, Cohen AD, Philhower J, Miller GW, Zigmond MJ, Schallert T. Forced limb-use effects on the behavioral and neurochemical effects of 6-hydroxydopamine. J Neurosci 2001 Jun 15;21(12):4427-4435 [FREE Full text] [Medline: 11404429]
12. Tillerson JL, Caudle WM, Reverón ME, Miller GW. Exercise induces behavioral recovery and attenuates neurochemical deficits in rodent models of Parkinson's disease. Neuroscience 2003;119(3):899-911. [doi: 10.1016/s0306-4522(03)00096-4] [Medline: 12809709]
13. Tajiri N, Yasuhara T, Shingo T, Kondo A, Yuan W, Kadota T, et al. Exercise exerts neuroprotective effects on Parkinson's disease model of rats. Brain Res 2010 Jan 15;1310:200-207. [doi: 10.1016/j.brainres.2009.10.075] [Medline: 19900418]
14. Yoon M, Shin M, Kim T, Kim B, Ko I, Sung Y, et al. Treadmill exercise suppresses nigrostriatal dopaminergic neuronal loss in 6-hydroxydopamine-induced Parkinson's rats. Neurosci Lett 2007 Aug 09;423(1):12-17. [doi: 10.1016/j.neulet.2007.06.031] [Medline: 17644250]
15. Dutra MF, Jaeger M, Ilha J, Kalil-Gaspar PI, Marcuzzo S, Achaval M. Exercise improves motor deficits and alters striatal GFAP expression in a 6-OHDA-induced rat model of Parkinson's disease. Neurol Sci 2012 Oct;33(5):1137-1144. [doi: 10.1007/s10072-011-0925-5] [Medline: 22231471]
16. Fisher BE, Petzinger GM, Nixon K, Hogg E, Bremmer S, Meshul CK, et al. Exercise-induced behavioral recovery and neuroplasticity in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-lesioned mouse basal ganglia. J Neurosci Res 2004 Aug 01;77(3):378-390. [doi: 10.1002/jnr.20162] [Medline: 15248294]
17. Lau Y, Patki G, Das-Panja K, Le W, Ahmad SO. Neuroprotective effects and mechanisms of exercise in a chronic mouse model of Parkinson's disease with moderate neurodegeneration. Eur J Neurosci 2011 Apr;33(7):1264-1274 [FREE Full text] [doi: 10.1111/j.1460-9568.2011.07626.x] [Medline: 21375602]
18. Pothakos K, Kurz MJ, Lau Y. Restorative effect of endurance exercise on behavioral deficits in the chronic mouse model of Parkinson's disease with severe neurodegeneration. BMC Neurosci 2009 Jan 20;10:6 [FREE Full text] [doi: 10.1186/1471-2202-10-6] [Medline: 19154608]
19. Chen H, Zhang SM, Schwarzschild MA, Hernán MA, Ascherio A. Physical activity and the risk of Parkinson disease. Neurology 2005 Mar 22;64(4):664-669. [doi: 10.1212/01.WNL.0000151960.28687.93] [Medline: 15728289]
20. Sasco AJ, Paffenbarger RS, Gendre I, Wing AL. The role of physical exercise in the occurrence of Parkinson's disease. Arch Neurol 1992 Apr;49(4):360-365. [doi: 10.1001/archneur.1992.00530280040020] [Medline: 1558515]
21. Thacker EL, Chen H, Patel AV, McCullough ML, Calle EE, Thun MJ, et al. Recreational physical activity and risk of Parkinson's disease. Mov Disord 2008 Jan;23(1):69-74 [FREE Full text] [doi: 10.1002/mds.21772] [Medline: 17960818]
22. Sääksjärvi K, Knekt P, Männistö S, Lyytinen J, Jääskeläinen T, Kanerva N, et al. Reduced risk of Parkinson's disease associated with lower body mass index and heavy leisure-time physical activity. Eur J Epidemiol 2014 Apr;29(4):285-292. [doi: 10.1007/s10654-014-9887-2] [Medline: 24633681]
23. van Nimwegen M, Speelman AD, Hofman-van Rossum EJM, Overeem S, Deeg DJH, Borm GF, et al. Physical inactivity in Parkinson's disease. J Neurol 2011 Dec;258(12):2214-2221 [FREE Full text] [doi: 10.1007/s00415-011-6097-7] [Medline: 21614433]
24. Forkan R, Pumper B, Smyth N, Wirkkala H, Ciol MA, Shumway-Cook A. Exercise adherence following physical therapy intervention in older adults with impaired balance. Phys Ther 2006 Mar;86(3):401-410. [Medline: 16506876]
25. Schutzer KA, Graves BS. Barriers and motivations to exercise in older adults. Prev Med 2004 Nov;39(5):1056-1061. [doi: 10.1016/j.ypmed.2004.04.003] [Medline: 15475041]
26. Trost SG, Owen N, Bauman AE, Sallis JF, Brown W. Correlates of adults' participation in physical activity: review and update. Med Sci Sports Exerc 2002 Dec;34(12):1996-2001. [doi: 10.1249/01.MSS.0000038974.76900.92] [Medline: 12471307]
27. Englert D, Hirsch M, Sanjak M, Russo P, Quinlan M, Crum K. Barriers to Physical Activity among Adults with Idiopathic Parkinson's Disease (PD) (P04.151). Neurology 2013;80(7 Supplement):P04.151-P04.151 [FREE Full text]
28. Afshari M, Yang A, Bega D. Motivators and Barriers to Exercise in Parkinson's Disease. J Parkinsons Dis 2017;7(4):703-711. [doi: 10.3233/JPD-171173] [Medline: 29103050]
29. Ellis T, Boudreau JK, DeAngelis TR, Brown LE, Cavanaugh JT, Earhart GM, et al. Barriers to exercise in people with Parkinson disease. Phys Ther 2013 May;93(5):628-636 [FREE Full text] [doi: 10.2522/ptj.20120279] [Medline: 23288910]

XSL•FO
RenderX

30. Keus SHJ, Bloem BR, Verbaan D, de Jonge PA, Hofman M, van Hilten BJ, et al. Physiotherapy in Parkinson's disease: utilisation and patient satisfaction. J Neurol 2004 Jun;251(6):680-687. [doi: 10.1007/s00415-004-0402-7] [Medline: 15311343]
31. Pickering RM, Fitton C, Ballinger C, Fazakarley L, Ashburn A. Self reported adherence to a home-based exercise programme among people with Parkinson's disease. Parkinsonism Relat Disord 2013 Jan;19(1):66-71. [doi: 10.1016/j.parkreldis.2012.07.006] [Medline: 22878222]
32. Schenkman M, Hall DA, Barón AE, Schwartz RS, Mettler P, Kohrt WM. Exercise for people in early- or mid-stage Parkinson disease: a 16-month randomized controlled trial. Phys Ther 2012 Nov;92(11):1395-1410 [FREE Full text] [doi: 10.2522/ptj.20110472] [Medline: 22822237]
33. 9zest. 9zest. URL: https://9zest.com [accessed 2020-08-26]
34. Moterum Technologies. URL: https://moterum.com [accessed 2020-08-26]
35. Ellis TD, Cavanaugh JT, DeAngelis T, Hendron K, Thomas CA, Saint-Hilaire M, et al. Comparative Effectiveness of mHealth-Supported Exercise Compared With Exercise Alone for People With Parkinson Disease: Randomized Controlled Pilot Study. Phys Ther 2019 Feb 01;99(2):203-216. [doi: 10.1093/ptj/pzy131] [Medline: 30715489]
36. Siegert C, Hauptmann B, Jochems N, Schrader A, Deck R. ParkProTrain: an individualized, tablet-based physiotherapy training programme aimed at improving quality of life and participation restrictions in PD patients - a study protocol for a quasi-randomized, longitudinal and sequential multi-method study. BMC Neurol 2019 Jun 25;19(1):143 [FREE Full text] [doi: 10.1186/s12883-019-1355-x] [Medline: 31238908]
37. Brown MJ, Chuang-Stein C, Kirby S. Designing studies to find early signals of efficacy. J Biopharm Stat 2012;22(6):1097-1108. [doi: 10.1080/10543406.2011.570466] [Medline: 23075010]
38. NCSS. PASS 19.0. Kaysville, UT: NCSS, LLC URL: www.ncss.com/software/pass [accessed 2020-08-30]
39. Landers MR, Navalta JW, Murtishaw AS, Kinney JW, Pirio Richardson S. A High-Intensity Exercise Boot Camp for Persons With Parkinson Disease: A Phase II, Pragmatic, Randomized Clinical Trial of Feasibility, Safety, Signal of Efficacy, and Disease Mechanisms. J Neurol Phys Ther 2019 Jan;43(1):12-25. [doi: 10.1097/NPT.0000000000000249] [Medline: 30531382]
40. Petersen C, Steffen T, Paly E, Dvorak L, Nelson R. Reliability and Minimal Detectable Change for Sit-to-Stand Tests and the Functional Gait Assessment for Individuals With Parkinson Disease. J Geriatr Phys Ther 2017;40(4):223-226. [doi: 10.1519/JPT.0000000000000102] [Medline: 27805924]
41. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999 Jun;70(2):113-119. [doi: 10.1080/02701367.1999.10608028] [Medline: 10380242]
42. Morris S, Morris ME, Iansek R. Reliability of measurements obtained with the Timed "Up & Go" test in people with Parkinson disease. Phys Ther 2001 Feb;81(2):810-818. [doi: 10.1093/ptj/81.2.810] [Medline: 11175678]
43. Brusse KJ, Zimdars S, Zalewski KR, Steffen TM. Testing functional performance in people with Parkinson disease. Phys Ther 2005 Mar;85(2):134-141. [Medline: 15679464]
44. Steffen T, Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. Phys Ther 2008 Jun;88(6):733-746. [doi: 10.2522/ptj.20070214] [Medline: 18356292]
45. Dal Bello-Haas V, Klassen L, Sheppard MS, Metcalfe A. Psychometric Properties of Activity, Self-Efficacy, and Quality-of-Life Measures in Individuals with Parkinson Disease. Physiother Can 2011;63(1):47-57 [FREE Full text] [doi: 10.3138/ptc.2009-08] [Medline: 22210979]
46. Fitzpatrick R, Norquist JM, Jenkinson C. Distribution-based criteria for change in health-related quality of life in Parkinson's disease. J Clin Epidemiol 2004 Jan;57(1):40-44. [doi: 10.1016/j.jclinepi.2003.07.003] [Medline: 15019009]
47. Peto V, Jenkinson C, Fitzpatrick R. Determining minimally important differences for the PDQ-39 Parkinson's disease questionnaire. Age Ageing 2001 Jul;30(4):299-302. [doi: 10.1093/ageing/30.4.299] [Medline: 11509307]
48. Hagell P, Nygren C. The 39 item Parkinson's disease questionnaire (PDQ-39) revisited: implications for evidence based medicine. J Neurol Neurosurg Psychiatry 2007 Nov;78(11):1191-1198 [FREE Full text] [doi: 10.1136/jnnp.2006.111161] [Medline: 17442762]
49. Jenkinson C, Fitzpatrick R. Cross-cultural evaluation of the short form 8-item Parkinson's Disease Questionnaire (PDQ-8): results from America, Canada, Japan, Italy and Spain. Parkinsonism Relat Disord 2007 Mar;13(1):22-28. [doi: 10.1016/j.parkreldis.2006.06.006] [Medline: 16931104]
50. Bonato M, Turrini F, DE Zan V, Meloni A, Plebani M, Brambilla E, et al. A Mobile Application for Exercise Intervention in People Living with HIV. Med Sci Sports Exerc 2020 Feb;52(2):425-433. [doi: 10.1249/MSS.0000000000002125] [Medline: 31415448]
51. Edney S, Ryan JC, Olds T, Monroe C, Fraysse F, Vandelanotte C, et al. User Engagement and Attrition in an App-Based Physical Activity Intervention: Secondary Analysis of a Randomized Controlled Trial. J Med Internet Res 2019 Nov 27;21(11):e14645 [FREE Full text] [doi: 10.2196/14645] [Medline: 31774402]
52. Colón-Semenza C, Latham NK, Quintiliani LM, Ellis TD. Peer Coaching Through mHealth Targeting Physical Activity in People With Parkinson Disease: Feasibility Study. JMIR Mhealth Uhealth 2018 Feb 15;6(2):e42. [doi: 10.2196/mhealth.8074] [Medline: 29449201]

XSL•FO
RenderX

53. van der Kolk NM, de Vries NM, Kessels RPC, Joosten H, Zwinderman AH, Post B, et al. Effectiveness of home-based and remotely supervised aerobic exercise in Parkinson's disease: a double-blind, randomised controlled trial. Lancet Neurol 2019 Nov;18(11):998-1008. [doi: 10.1016/S1474-4422(19)30285-6] [Medline: 31521532]
54. Dunn EE, Gainforth HL, Robertson-Wilson JE. Behavior change techniques in mobile applications for sedentary behavior. Digit Health 2018;4:2055207618785798 [FREE Full text] [doi: 10.1177/2055207618785798] [Medline: 31463076]
55. Sancho-Garcia S, Sanz-de Diego S, Medina-Porqueres I. Apps to prescribe therapeutic exercise among rehabilitating adults: a systematic review. J Sports Med Phys Fitness 2020 Mar;60(3):472-478. [doi: 10.23736/S0022-4707.19.09601-4] [Medline: 31565908]
56. Brindal E, Hendrie GA, Freyne J, Noakes M. A Mobile Phone App Designed to Support Weight Loss Maintenance and Well-Being (MotiMate): Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Sep 04;7(9):e12882 [FREE Full text] [doi: 10.2196/12882] [Medline: 31486407]

Abbreviations

MDC: minimal detectable change
PDQ8: Parkinson's Disease Questionnaire 8
STS: 30-second Sit-To-Stand
TUG: Timed Up and Go

Edited by G Eysenbach; submitted 30.03.20; peer-reviewed by M Hirsch, N de Vries, E van der Velde; comments to author 12.06.20; revised version received 10.07.20; accepted 23.07.20; published 05.10.20.

Please cite as:
Landers MR, Ellis TD
A Mobile App Specifically Designed to Facilitate Exercise in Parkinson Disease: Single-Cohort Pilot Study on Feasibility, Safety, and Signal of Efficacy
JMIR Mhealth Uhealth 2020;8(10):e18985
URL: https://mhealth.jmir.org/2020/10/e18985
doi:10.2196/18985
PMID:33016887

©Merrill R Landers, Terry D Ellis. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 05.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patients' Experiences of Using Smartphone Apps to Support Self-Management and Improve Medication Adherence in Hypertension: Qualitative Study

Ciara M McBride[1*], BSc, MSc; Eimear C Morrissey[2*], BA, MSc, PhD; Gerard J Molloy[3*], BSc, PhD

[1]School of Psychology, National Univeristy of Ireland Galway, Galway, Ireland

[2]Health Behaviour Change Research Group, National Univiersity of Ireland Galway, Galway, Ireland

[3]Medication Adherence Across the Lifespan Research Group, National University of Ireland Galway, Galway, Ireland

[*]all authors contributed equally

Corresponding Author:
Ciara M McBride, BSc, MSc
School of Psychology
National Univeristy of Ireland Galway
University Road
Galway, H91 TK33
Ireland
Phone: 353 91524411
Email: mcbrideciara5@gmail.com

Abstract

Background: Worldwide, hypertension control rates remain suboptimal despite clinically effective antihypertensive drug therapy. Patient failure to take medication as prescribed (ie, nonadherence) is the most important factor contributing to poor control. Smartphone apps can facilitate the delivery of evidence-based behavior change techniques to improve adherence and may provide a scalable, usable, and feasible method to deliver self-management support.

Objective: The aim of this study is to explore patients' experiences of the usability and feasibility of smartphone apps to support self-management and improve medication adherence in hypertension.

Methods: A qualitative descriptive study was conducted. A total of 11 people living with hypertension from the West of Ireland were sampled purposively and interviewed about their experience of using a self-management app for a 4-week period, which included two key functionalities: self-monitoring of blood pressure (BP) and medication reminders. Thematic analysis was carried out on the semistructured interview data.

Results: Participants' age ranged from 43 to 74 years (mean 62 years, SD 9.13). Three themes were identified: digital empowerment of self-management, human versus digital systems, and digital sustainability. Although patients' experience of using the technology to self-monitor BP was one of empowerment, characterized by an enhanced insight and understanding into their condition, control, and personal responsibility, the reminder function was only feasible for patients who reported unintentional nonadherence to treatment. Patients experienced the app as a sustainable tool to support self-management and found it easy to use, including those with limited technological competence.

Conclusions: The study's findings provide new insights into the experience of using apps to support medication adherence in hypertension. Overall, the data support apps as a usable and feasible method to aid self-management of hypertension and highlight the need for personalized functionality, particularly with regard to medication adherence reminder strategies. The study's findings challenge the perspective that the use of these technologies to support self-management can inevitably add to the burden of treatment experienced by patients.

(JMIR Mhealth Uhealth 2020;8(10):e17470) doi:10.2196/17470

KEYWORDS

hypertension; self-management; mobile applications; feasibility; usability; medication adherence; qualitative research; digital technology; mobile phone

XSL•FO
RenderX

Introduction

Hypertension has been identified as the leading modifiable risk factor for cardiovascular disease and consequently represents a major cause of premature morbidity and mortality due to adverse cardiovascular and cerebrovascular events [1]. Blood pressure (BP) control and management of hypertension can be achieved through antihypertensive drug treatment, which has proved to be clinically effective [2,3]. However, recent evidence (including control rates across 12 countries) suggests that BP control through antihypertensive treatment is suboptimal, with at least 20% of those prescribed treatment failing to achieve control [4]. Findings from nationally representative community-based studies, such as The Irish Longitudinal Study on Ageing, report control rates as low as 50% [5]. Patient nonadherence (ie, failure to take medication as prescribed) continues to be recognized as the most critical variable in suboptimal BP control through antihypertensive treatment [6].

Some intervention approaches to support self-management of hypertension and enhance adherence have been shown to be effective in improving BP control [7] and increasing adherence [8]. These interventions have included a heterogeneous range of behavior change techniques to improve medication adherence [9]. Systematic and meta-analytic review evidence supports self-monitoring of BP as a successful intervention component to reduce mean systolic and diastolic BP readings [10,11] and increase adherence to antihypertensive therapy [12]. Automated medication reminders (ie, environmental prompts or cues) offer a simple, yet empirically established intervention component associated with enhanced adherence and improved BP control among high-risk patients with hypertension [13]. It is clear that such active intervention ingredients can be delivered using innovative digital technologies, including smartphone apps, which might constitute a scalable, potentially usable, and feasible method to deliver self-management support to patients with hypertension [14,15].

Despite the potential of such devices to improve adherence, a contemporary theoretical perspective proposes that newer technologies that encourage a proactive approach to patient care and can be easily shifted from the clinic to the community may serve to place new demands on patients, causing them to experience an increased burden of treatment [16]. Furthermore, an exploration of patients' perspectives on the possible adoption of smartphone apps to support self-management and improve adherence in hypertension has revealed important concerns, including the potential of such technologies to increase health-related anxiety, and the authors have cautioned that the current patient perspective might be best characterized as one of *ambivalence* [15]. An exploration of the patient experience of using these technologies in an ongoing community context is now warranted to provide timely insights into these concerns. Therefore, the aim of this study is to explore patients' experiences of the usability and feasibility of smartphone apps to support self-management and improve medication adherence in hypertension.

Methods

Design

A qualitative descriptive study was conducted. The Consolidated Criteria for Reporting Qualitative Research Checklist was used to ensure explicit reporting on how this study was carried out (Multimedia Appendix 1) [17].

Participants and Recruitment

Participants were recruited through Croí (the West of Ireland Cardiac Foundation), a heart and stroke charity located in Galway city. Recruitment emails advertising the study were sent to database records of individuals who had used the charity's service until March 2019. Interested patients living with hypertension contacted the primary researcher by telephone or email to obtain further study information and ensure eligibility. To be eligible for participation, patients must have been prescribed at least one form of antihypertensive therapy and own an Android smartphone. Purposive sampling was used to achieve adequate variability in age, sex, and complexity of the medication regime.

The App

The smartphone app used is called *BP Journal*. BP Journal is commercially available (Google Play Store), serves as a companion app to a clinically validated home BP monitor, and was chosen by the research team as it is typical of medication management apps for hypertension in that it includes two key functionalities: self-monitoring of BP and medication reminders. Users of the app create a personal profile and set daily reminders to take BP medication and readings. The app allows for the self-reporting and storage of BP readings, producing feedback of BP measurements in the form of statistics and interactive charts. An export function provides users with an opportunity to print or send via email BP data in comma-separated values or PDF format.

Procedure

Participants attended an initial session during which they downloaded the app and created their personal profile. They were then provided with a home BP monitor (A&D Medical model UA-767S-W) and shown how to navigate the equipment. Participants were requested to use these materials to support hypertension self-management for 4-weeks. To ensure participants became familiar with app functionalities, it was recommended that they took BP readings at least twice per week (unless they had a pre-established self-monitoring routine). The 1-month feasibility duration was based on previous mobile health feasibility studies using 2- to 4-week periods [18,19].

Semistructured, face-to-face interviews were facilitated and audio-recorded by the primary researcher (CM, a female MSc graduate in health psychology with experience and training in qualitative research methods of both collection and analysis) to elicit patients' experiences of using the app. The interview topic guide, which was centered around questions relating to feasibility and usability, was composed by reviewing qualitative research in the area and revised during the data collection process to be responsive to unforeseen issues. Therefore, an

XSL•FO
RenderX

iterative approach was adopted to ensure that the researcher did not restrict the analysis to only issues anticipated as relevant [20]. Feasibility was defined in terms of practicability of using the app to self-manage hypertension and the extent to which participants successfully (or unsuccessfully) used the BP Journal app for the 4-week period (Multimedia Appendix 2). The interviewer had no previous relationship with participants before the study commenced, and the duration of the interviews ranged from 9 to 29 min. All data were collected at Croí House, the charity's dedicated facility. No nonparticipants were present during the data collection process other than the primary researcher.

The System Usability Scale (SUS) [21] was distributed to participants before the commencement of interviews to provide descriptive data on the perceived usability of the app. Scores on the SUS range from 0 to 100, with higher scores indicating greater perceived *ease of use*. A score of 68 or above is considered above average [22].

Data Analysis

Interviews were transcribed verbatim and analyzed using thematic analysis, following the six phases proposed by Braun and Clarke [23]. "Thematic analysis is a method for identifying, analyzing, and reporting patterns (themes) within data" and is celebrated for its accessible and theoretically flexible approach to analyzing qualitative data. A realist inductive approach, a data-driven approach to analysis that reports on the experiences of participants without trying to fit into an existing coding frame or follow prior theoretical conceptions, was taken [23]. During phase one, the familiarization process involved active reading of the transcripts. Phase two involved the generation of an initial list of codes and involved the assignment of sections of the transcripts to descriptive codes. Phase three involved refocusing the analysis by sorting codes into potential themes. This involved the creation of preliminary categories comprising codes that were conceptually linked. Phase four included refining a set of candidate themes produced during phase three, and phase five involved defining and naming the themes. Phase six included the write-up of the analysis in a concise, coherent, and logical manner [23]. The refinement of themes during phase five resulted in the organization of patterns within each theme into subthemes, which also enhanced the coherence and presentation of the results. In an effort to enhance reflexivity, 2 members of the research team (2 health psychologists) joined the primary researcher to review preliminary categories, refine candidate themes and subthemes, and each contributed to the analysis [24]. NVivo version 12 (QSR International) was used to facilitate the analysis.

Ethical Approval

Ethical approval was sought and granted for this study by the School of Psychology Research Ethics Committee at the National University of Ireland, Galway.

Results

Participant Characteristics

Patient characteristics are summarized in Table 1. Three themes were identified in the data: digital empowerment of self-management, human versus digital systems, and digital sustainability. To reflect patients' perceived usability of the technology, the SUS score of each participant complements supporting quotations below.

XSL•FO
RenderX

Table 1. Patient characteristics (N=11).

Characteristics	Values
Demographic characteristics	
Age (years), mean (range)	62 (43-74)
Female, n (%)	6 (54)
Urban, n (%)	8 (73)
Married, n (%)	9 (82)
Single, n (%)	2 (18)
Working full-time, n (%)	4 (36)
Working part-time, n (%)	1 (9)
Unemployed, n (%)	1 (9)
Retired, n (%)	5 (46)
Private health insurance, n (%)	5 (46)
Clinical characteristics	
Number of antihypertensive medications, mean (range)	1.73 (1-3)
>12 months since diagnosis, n (%)	11 (100)
Presence of multimorbid conditions, n (%)	9 (82)
Other	
Previous use of a smartphone to manage hypertension, n (%)	1 (9)
SUS[a] score, mean (range)	89.1 (70-97.5)
Self-rated technology literacy, n (%)	
High	7 (64)
Medium	4 (36)

[a]SUS: System Usability Scale.

Digital Empowerment of Self-Management

The theme of digital empowerment of self-management refers to patients' experiences of using the technology to self-monitor BP and is represented by two subthemes: *insight and understanding* and *control and responsibility* Although none of the participants had a regular self-monitoring routine at the time of study commencement, some reported prior occasional use of monitoring equipment. For others, the study provided their first experience of self-monitoring BP. Many patients who reported prior use of monitoring devices described how using the app had provided them with a better monitoring system, when compared with previous *ad hoc* systems:

> *The thing about [using the app] is that it is more regular...up to this I would only use the monitor when I felt I had high BP...doing it on regular basis is a better monitoring system, whereas mine was ad hoc.* [Male, 65 years; SUS 87.5]

Insight and Understanding

Several patients valued the output information produced by the self-monitoring function and felt it provided them with a greater insight into the state of their BP. The retrieval of visual feedback in the charts offered to participants more meaningful information and a feeling of empowerment:

> *The good thing about the app was I could see it visually; visual things mean a lot more to me.* [Male, 65 years; SUS 87.5]

> *All the information it gave me about the state of the BP [was the most helpful aspect of the app]...I could see the different measurements myself and see the charts up and down.* [Female, 63 years; SUS 97.5]

Many participants expressed how visualizing BP data in the charts made their condition more tangible, whereas the visual qualities in the app helped others make sense of and interpret BP readings. Some acknowledged that using the self-monitoring function had provided them with an enhanced understanding of the asymptomatic nature of hypertension and the hazards of judging their condition by how they felt, thereby reconstructing the way they perceived their condition:

> *You are getting an accurate account of what's happening in your body...normally you are judging yourself by how you feel...I was surprised my BP was higher than I thought...that made me aware sometimes I might feel good and be busy, but I was overlooking.* [Male, 65 years; SUS 87.5]

Visualizing fluctuations in BP readings was perceived as interesting by several and sparked curiosity. Fluctuations in readings despite feelings of accurate medication-taking behaviors evoked a sense of concern in some patients regarding

XSL•FO
RenderX

whether their medication was working to control their BP. For others, the retrieval of consistent readings verified that their BP was under control:

> *It was a nice app...it told me my BP was fluctuating even though it was the same time of day. I am considering going back to my doctor and showing him the results on my follow-up...there might be another tablet that might be better.* [Female, 71 years; SUS 97.5]

Control and Responsibility

Many patients expressed how their newly established self-monitoring regime had provided them with an improved feeling of control over their condition between consultations in primary care. They described how this had offered them reassurance and had promoted an increased feeling of responsibility for one's health. One participant described how she thought this improved feeling of control might function to decrease health anxiety in patients who were worried about their BP:

> *When you go and get your BP taken and you're put on medication if everything is going as it should you may not go back to the doctor, you don't know what's going on...six months can make a big change. I just heard of another person, they are panicked about their BP and ringing the doctor...if you had this at home you might calm down...you have this back up to know that everything is controlled...you're looking after yourself and have your own check.* [Female, 50 years; SUS 95]

Several of the participants described how their improved feeling of responsibility would help them *take action* if their BP readings were consistently out of range. Many patients for whom the study provided them with their first experience of self-monitoring BP discussed how their new sense of responsibility was stronger than what they had felt with the conventional standard office monitoring system in primary care:

> *If there were any unusual stuff happening, I'd pick it up...sometimes you go to a GP to get a check-up, but you wouldn't pick it up there and then...if you are doing this self-checking yourself...if there is something unusual happening you could take action.* [Male, 56 years; SUS 92.5]

Feeling empowered, some reported instances during which the app had encouraged them to play an active role in consultations. This occurred when patients used the app to guide conversation with their general practitioner (GP). A few patients underwent ambulatory monitoring, followed by the detection of uncontrolled BP and, in turn, experienced changes to their medication:

> *I showed it to my doctor, and she did a blood pressure monitor...[It] advised that my blood pressure was very high...I feel glad I found it out.* [Female, 63 years; SUS 97.5]

Human Versus Digital Systems

The theme of *human versus digital systems* focuses on patients' experiences of the technology to support reminder strategies in daily self-management and is represented by two subthemes: *utility of the digital system* and *interference of the digital system.*

Utility of the Digital System

For those patients who reported they might frequently forget to take medication (ie, unintentional nonadherence) and could recognize their current medication-taking system was not perfect, the medication reminder was experienced as useful:

> *Some mornings I was rushing and only for it did beep I probably would have forgotten to take them...it's a tool, it helps because we can forget easily.* [Female, 50 years; SUS 95]

> *I mean the reminder was good because it told you at the time to take your medication and there is a chance with me, I would forget...there was a couple of times I totally forgot.* [Male, 74 years; SUS 70]

One participant described how the medication reminder was becoming integrated into his everyday routine:

> *The reminder was excellent...I'd look at the clock and say it wasn't some message coming in...I knew it was time to take my medication.* [Male, 65 years; SUS 87.5]

Interference of the Digital System

The medication reminder was experienced as less practical among patients who expressed a strong satisfaction with their current *human system*, linking medication taking to events in their daily routines. These patients varied in future intent to engage with the reminder. Although several patients could see the benefit of using the reminder if they were out of routine, others considered the reminder as unnecessary:

> *I didn't need reminding about my tablets in all the years I'm on them I've probably forgotten to take them twice ever...it's part of my routine after breakfast and before I go to bed...having said that I am going on holidays in a few weeks, it will be good for that.* [Female, 43 years; SUS 87.5]

One patient described how she had tried to stick to the time-based reminder but reverted to *old habits* out of fear that she might forget to take her medication. In this way, the *digital system* interfered with the patients pre-established medication routine. This patient expressed how she thought the reminder would be more feasible for patients who did not have strong habits:

> *I thought the reminder was good...if I was somebody starting off that hadn't made all these habits...but for me after 25 or 30 years I couldn't do it, I think it's just old habits die hard.* [Female, 63 years; SUS 97.5]

Digital Sustainability

The final theme to be identified was *digital sustainability.* The theme of digital sustainability refers to participants' thoughts and ideas about the use of the technology in the future to support

XSL•FO
RenderX

hypertension self-management and suggestions to improve app functionality. This theme is further represented by two subthemes: *long-term use of the digital technology* and *improvement suggestions*.

Long-Term Use of the Digital Technology

Most participants reported no concerns about using the technology going forward. Several reported that the app and monitor were easy to use despite limited technological competence:

> *Technology over the last couple of years...what I could do before I have kind of forgotten. I mean I [still] found the app very...once you did one or two testings with it, it became automatic.* [Male, 74 years; SUS 70]

Most participants expressed that they intended to continue using the technology after the study. The self-monitoring function was a major motivation for continued use. Many patients who did not own a monitoring device acknowledged that they intended to buy one. Perceiving the app as being beneficial in consultations encouraged sustainability:

> *I am actually going to buy a monitor of my own and if the app stays there I am going to continue taking readings...I will have it there to show the trend in my blood pressure, so I can pass it on to a doctor.* [Male, 63 years; SUS 92.5]

Improvement Suggestions

Important improvement suggestions included making the reminder more intrusive by adding an audio signal or a *snooze* function. Another suggestion included making the reminder function more comprehensive by allowing patients to confirm in the app that they had taken their medications. The utility for more comprehensive app functionalities, including components to promote adherence to subsequent lifestyle behaviors such as physical activity and diet, was also mentioned. Suggestions to increase the complexity of the app were weighed against a strong emphasis to keep app functionalities simple to ensure that the app was suitable for all users, particularly those who might not be savvy:

> *I find in life, it's called the KISS principle...keep it straight and simple...and the app in my opinion is easy to use...you have to make it easy to use...remember your clients that you are giving it to are not savvy...* [Male, 63 years; SUS 92]

Discussion

Principal Findings

The data from these interviews provide new insights into patients' experiences of the usability and feasibility of smartphone apps to improve adherence to hypertension. The patient experience of using the technology to support self-monitoring of BP was one of empowerment, characterized by an enhanced insight and understanding of hypertension, an improved feeling of control over health, reassurance, and increased patient responsibility. Patients' experiences of using the technology to support reminder strategies were more varied; although the reminder function was useful for patients who reported unintentional nonadherence to treatment, it was less practical for those who reported existing context-based medication-taking habits. Overall, most patients experienced the technology as sustainable, reporting that they intended to continue to engage with the app in the future. Patients were confident in their use of the technology by the end of the feasibility period despite limited digital competence. This was complemented by descriptive data, which demonstrated that all participants scored above average on the SUS, a robust evaluation tool of perceived *ease of use* [25].

Strengths and Limitations

This study provides timely data on the use of technology to support self-management of hypertension and to the researcher's knowledge is the first to use qualitative methods to examine patients' experiences of the usability and feasibility of smartphone apps specifically for medication adherence in hypertension. This extends research in this area [15] that focused on initial impressions of smartphone app use rather than using them over an extended period. This provides the perspective of users beyond an early novelty phase and provides new findings that have enhanced ecological validity. Semistructured, face-to-face interviews were appropriate because they prevented the loss of contextual and nonverbal data, allowing the interviewer to pick up on visual cues, such as how to proceed with dialogue [26]. In addition, individual interviews were appropriate as they allowed patients to express their own individual opinion, avoiding the social desirability bias, which might concern group-based methods of data collection. It is also likely that the research team members with different levels of expertise coming together to each contribute to the analysis heightened reflexivity, which is another potential methodological strength.

One limitation is the potential of sampling bias. Although purposive sampling was used, the sample was from a single geographic location and limited to Android smartphone users. It is also possible that this sample captures a more positive patient experience of self-management technologies, as it is based only on a subsample of patients who volunteered to participate and who had prior involvement with Croí, the Irish heart and stroke charity. This is due to the potential likelihood of the participants reflecting a more motivated, technologically inclined, and adherent group. It must also be recognized that all the patients had a hypertension diagnosis of >12 months, which could have influenced the findings.

Comparison With Prior Research

Our findings support previous research investigating patients' experiences of self-management technologies across other illness contexts. A metaethnography on digital interventions to support self-management covering a range of conditions (eg, asthma, chronic obstructive pulmonary disease, diabetes) has found that patients monitoring their health felt reassured by the insight provided and felt they had more meaningful conversations with health care practitioners [27]. The review also found that patients who could understand self-monitored data in the context of lifestyle behaviors (eg, medication adherence) felt an improved sense of control over their condition, which allowed them to

XSL•FO
RenderX

find meaning in physiological self-monitored data. This is reflected in our theme of *digital empowerment of self-management* wherein patients expressed a sense of reassurance, played a more active role in consultations (or believed self-monitored data would allow them to do so), and ascribed meaning to visual feedback. Similarly, many patients linked readings to medication-taking behavior and felt improved control over their BP. Our theme of *digital empowerment of self-management* is also consistent with the findings of a thematic synthesis by Fletcher et al [28], which focused specifically on self-monitoring of BP in hypertension and found that increased control, autonomy, and self-efficacy enabled patients to move from a passive to an active role in consultations, promoting patient involvement in care.

Our findings support previous literature from a number of areas including information visualization, human-computer interaction, and medicine, which argues that making health data visible creates opportunities for patients to make sense of their illness, which in turn may lead to more effective self-management [29]. In our data, visualizing self-reported BP data made hypertension more tangible for patients and reconstructed the way they perceived their condition. These findings are consistent with the findings of a study by Hallberg et al [30], who found that using a mobile phone–based self-management system that included visual feedback of self-monitored BP data caused patients with hypertension to experience changes to illness perceptions, including timeline and treatment beliefs. These theoretical constructs from the Common-Sense Self-Regulation Model (CS-SRM [31]) have been consistently shown to predict adherence to antihypertensive therapy in quantitative studies [32,33]. These findings have also been supported in qualitative studies [34].

The findings of this study also lend support to emerging research that points to the possible utility of digital technology, especially smartphone apps, to support reminder strategies in some patients who report unintentional nonadherence to treatments (eg, [35]). This is reflected in our theme *human versus digital systems* wherein patients who reported unintentional nonadherence to antihypertensive medication found the digital reminder useful, whereas those who reported contextual-based habits did not.

Implications for Practice, Research, and Design

Ineffective communication with patients with hypertension is correlated with poor adherence to antihypertensive therapy, and effective communication has been identified as an important obstacle for practitioners, largely due to a lack of understanding of hypertension in patients [36]. In our data, use of the technology enhanced patients' understanding of hypertension and many reported uses of the app in consultations to facilitate discussion. The use of the app in consultations helped to engage patients in shared decision making regarding treatment. These findings suggest that the use of the technology can promote more effective interaction in practice, potentially inadvertently increasing adherence and patient involvement in care, and helping to bridge a gap in the traditional doctor-patient relationship. Despite this implication of this finding for practice, previous inquiries have shown that this shift in power balance is not always viewed as a positive thing by practitioners [37,38]. Another important consideration highlighted by a recent study on GPs' perspectives that must be addressed in relation to the doctor-patient relationship and such technologies in practice concerns the issue of who holds the responsibility for self-monitored BP data [37].

The findings of this study have important implications for optimizing technological design. In addition to considering the feasibility of the suggestions made by these patients to promote the sustainability of the app, the authors recommend that future developers of adherence technologies for hypertension design contextual-based reminder functionalities to support habit development in patients who report unintentional nonadherence. Prospective memory is an important component of medication adherence and is supported by context-based cues more effectively than time-based cues [39]. This is reflected in our findings, as patients who were satisfied with their *human system* linked medication taking to events in their everyday lives. The potential capacity of such devices to support habit formation in hypertension may represent a key mechanism whereby adherence can be enhanced, as the extended CS-SRM has emphasized that habit strength is a significant theoretical predictor of unintentional nonadherence, and related evidence indicates that it may be the most potent predictor of long-term adherence to antihypertensive therapy [40].

Finally, it is noteworthy that a large proportion of participants who were living with hypertension in this study reported the presence of multimorbidity (defined as the presence of two or more chronic health conditions coexisting in one individual). Epidemiologic evidence suggests that multimorbidity is now the norm, rather than the exception [41]. Although the prevalence of multimorbidity in the study sample is higher than that reported in other studies, including community-based samples of patients with hypertension [42], the high prevalence reflects a need for medication management technologies to take such findings into account. Although the current focus was on hypertension self-management, it remains undetermined whether patients' experience of self-managing hypertension in this study had an impact on their management of coexisting conditions. It is a potential avenue for future research to develop scalable methods to support the management of polypharmacy (use of multiple medications) to promote adherence in patients with multimorbidity. The authors are aware that preliminary research efforts are being made in this area (eg, [43]). However, a recent Cochrane review [44] examining interventions to manage multiple medications in older adults included limited studies that currently use digital technology to support self-management of multimorbidity. Careful development is required, which integrates behavior change science.

Conclusions

The patients in this qualitative study experienced the technology as a usable and feasible method to support self-monitoring of BP. However, to experience the technology as feasible to support reminder strategies, this study's findings suggest that the patients must first have found a need that their current medication-taking systems could not fill. The technology was generally experienced by these patients as a sustainable tool to support hypertension self-management in the long term. Overall, the

XSL•FO
RenderX

data support apps as a usable and feasible method to support self-management and highlight the need for personalized functionality in relation to medication adherence reminder strategies. This study adds to a growing body of literature that challenges the perspective that the use of such self-management technologies will inevitably add to the patient burden of treatment [16].

Acknowledgments

The authors would like to acknowledge Professor Liam Glynn from the University of Limerick for providing BP monitors. The authors would also like to acknowledge the team at Croí for helping to facilitate recruitment and data collection. Finally, we would like to acknowledge and thank all the participants.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consolidated Criteria for Reporting Qualitative Research Checklist.
[PDF File (Adobe PDF File), 412 KB - mhealth_v8i10e17470_app1.pdf]

Multimedia Appendix 2
Interview Topic Guide.
[PDF File (Adobe PDF File), 106 KB - mhealth_v8i10e17470_app2.pdf]

References

1. Amici A, Cicconetti P, Sagrafoli C, Baratta A, Passador P, Pecci T, et al. Exaggerated morning blood pressure surge and cardiovascular events. A 5-year longitudinal study in normotensive and well-controlled hypertensive elderly. Archiv Gerontol Geriat 2009 Sep;49(2):e105-e109. [doi: 10.1016/j.archger.2008.10.003]
2. Musini V, Tejani A, Bassett K, Wright J. Pharmacotherapy for hypertension in the elderly. Cochrane Database Syst Rev 2009;4. [doi: 10.1002/14651858.cd000028.pub2]
3. Wright J, Musini V, Gill R. First line drugs for hypertension. Cochrane Database Syst Rev 2018;4. [doi: 10.1002/14651858.cd001841.pub3]
4. Zhou B, Danaei G, Stevens GA, Bixby H, Taddei C, Carrillo-Larco RM, et al. Long-term and recent trends in hypertension awareness, treatment, and control in 12 high-income countries: an analysis of 123 nationally representative surveys. Lancet 2019 Aug;394(10199):639-651. [doi: 10.1016/s0140-6736(19)31145-6]
5. Murphy C, Kearney P, Shelley E, Fahey T, Dooley C, Kenny R. Hypertension prevalence, awareness, treatment and control in the over 50s in Ireland: evidence from The Irish Longitudinal Study on Ageing. J Public Health 2015 Apr 28;38(3):450-458. [doi: 10.1093/pubmed/fdv057]
6. Burnier M, Egan BM. Adherence in hypertension. Circ Res 2019 Mar 29;124(7):1124-1140. [doi: 10.1161/circresaha.118.313220]
7. Glynn L, Murphy A, Smith S, Schroeder K, Fahey T. Interventions used to improve control of blood pressure in patients with hypertension. Cochrane Database Syst Rev 2010;3. [doi: 10.1002/14651858.cd005182.pub4]
8. Conn VS, Ruppar TM, Chase JD, Enriquez M, Cooper PS. Interventions to improve medication adherence in hypertensive patients: systematic review and meta-analysis. Curr Hypertens Rep 2015 Nov 11;17(12). [doi: 10.1007/s11906-015-0606-5]
9. Morrissey EC, Durand H, Nieuwlaat R, Navarro T, Haynes RB, Walsh JC, et al. Effectiveness and content analysis of interventions to enhance medication adherence and blood pressure control in hypertension: A systematic review and meta-analysis. Psychol Health 2017 Jan 27;32(10):1195-1232. [doi: 10.1080/08870446.2016.1273356]
10. Bray EP, Holder R, Mant J, McManus RJ. Does self-monitoring reduce blood pressure? Meta-analysis with meta-regression of randomized controlled trials. Annal Medicine 2010 May 27;42(5):371-386. [doi: 10.3109/07853890.2010.489567]
11. Cappuccio FP, Kerry SM, Forbes L, Donald A. Blood pressure control by home monitoring: meta-analysis of randomised trials. Br Med J 2004 Jun 11;329(7458):145. [doi: 10.1136/bmj.38121.684410.ae]
12. Fletcher BR, Hartmann-Boyce J, Hinton L, McManus RJ. The effect of self-monitoring of blood pressure on medication adherence and lifestyle factors: a systematic review and meta-analysis. Am J Hypertens 2015 Oct;28(10):1209-1221. [doi: 10.1093/ajh/hpv008] [Medline: 25725092]
13. Patel S, Jacobus-Kantor L, Marshall L, Ritchie C, Kaplinski M, Khurana PS, et al. Mobilizing your medications: an automated medication reminder application for mobile phones and hypertension medication adherence in a high-risk urban population. J Diabetes Sci Technol 2013 May;7(3):630-639. [doi: 10.1177/193229681300700307]
14. Kumar N, Khunger M, Gupta A, Garg N. A content analysis of smartphone-based applications for hypertension management. J Am Soc Hypertens 2015 Feb;9(2):130-136 [FREE Full text] [doi: 10.1016/j.jash.2014.12.001] [Medline: 25660364]

XSL•FO
RenderX

15. Morrissey EC, Casey M, Glynn LG, Walsh JC, Molloy GJ. Smartphone apps for improving medication adherence in hypertension: patients' perspectives. PPA 2018 May;Volume 12:813-822. [doi: 10.2147/ppa.s145647]
16. May CR, Eton DT, Boehmer K, Gallacher K, Hunt K, MacDonald S, et al. Rethinking the patient: using burden of treatment theory to understand the changing dynamics of illness. BMC Health Serv Res 2014;14:281 [FREE Full text] [doi: 10.1186/1472-6963-14-281] [Medline: 24969758]
17. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
18. Rhee H, Allen J, Mammen J, Swift M. Mobile phone-based asthma self-management aid for adolescents (mASMAA): a feasibility study. Patient Prefer Adherence 2014;8:63-72 [FREE Full text] [doi: 10.2147/PPA.S53504] [Medline: 24470755]
19. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014 Nov;40(6):1244-1253. [doi: 10.1093/schbul/sbu033] [Medline: 24609454]
20. Ziebland S, McPherson A. Making sense of qualitative data analysis: an introduction with illustrations from DIPEx (personal experiences of health and illness). Med Educ 2006 May;40(5):405-414. [doi: 10.1111/j.1365-2929.2006.02467.x]
21. Brooke J. System Usability Scale (SUS): A Quick-and-Dirty Method of System Evaluation User Information. Reading, United Kingdom: Digital Eqipment Co Ltd; 1986.
22. Sauro J. Denver, CO: Measuring Usability LLC; 2011.
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Richards L. Handling Qualitative Data: A Practical Guide. London, UK: SAGE; 2014.
25. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum-Comput Interac 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
26. Novick G. Is there a bias against telephone interviews in qualitative research? Res Nurs Health 2008 Aug;31(4):391-398 [FREE Full text] [doi: 10.1002/nur.20259] [Medline: 18203128]
27. Morton K, Dennison L, May C, Murray E, Little P, McManus RJ, et al. Using digital interventions for self-management of chronic physical health conditions: A meta-ethnography review of published studies. Patient Educ Couns 2017 Apr;100(4):616-635 [FREE Full text] [doi: 10.1016/j.pec.2016.10.019] [Medline: 28029572]
28. Fletcher BR, Hinton L, Hartmann-Boyce J, Roberts NW, Bobrovitz N, McManus RJ. Self-monitoring blood pressure in hypertension, patient and provider perspectives: a systematic review and thematic synthesis. Patient Educ Counsel 2016 Feb;99(2):210-219. [doi: 10.1016/j.pec.2015.08.026]
29. Faisal S, Blandford A, Potts HW. Making sense of personal health information: challenges for information visualization. Health Informatics J 2013 Sep;19(3):198-217. [doi: 10.1177/1460458212465213] [Medline: 23981395]
30. Hallberg I, Ranerup A, Kjellgren K. Supporting the self-management of hypertension: patients' experiences of using a mobile phone-based system. J Hum Hypertens 2016 Feb;30(2):141-146 [FREE Full text] [doi: 10.1038/jhh.2015.37] [Medline: 25903164]
31. Leventhal H, Brissette I, Leventhal E. The common-sense model of self-regulation of healthillness. In: The Self-Regulation of HealthIllness Behaviour. New York, USA: Routledge; 2003:42-65.
32. Hsiao C, Chang C, Chen C. An investigation on illness perception and adherence among hypertensive patients. Kaohsiung J Med Sci 2012 Aug;28(8):442-447. [doi: 10.1016/j.kjms.2012.02.015]
33. Meyer D, Leventhal H, Gutmann M. Common-sense models of illness: the example of hypertension. Health Psychol 1985;4(2):115-135. [doi: 10.1037//0278-6133.4.2.115]
34. Bokhour BG, Cohn ES, Cortés DE, Solomon JL, Fix GM, Elwy AR, et al. The role of patients' explanatory models and daily-lived experience in hypertension self-management. J Gen Intern Med 2012 Jul 21;27(12):1626-1634. [doi: 10.1007/s11606-012-2141-2]
35. Stawarz K, Rodríguez MD, Cox AL, Blandford A. Understanding the use of contextual cues: design implications for medication adherence technologies that support remembering. Digit Health 2016;2:2055207616678707 [FREE Full text] [doi: 10.1177/2055207616678707] [Medline: 29942574]
36. Jolles EP, Clark AM, Braam B. Getting the message across. J Hypertension 2012;30(8):1500-1510. [doi: 10.1097/hjh.0b013e32835476e1]
37. Morrissey EC, Glynn LG, Casey M, Walsh JC, Molloy GJ. New self-management technologies for the treatment of hypertension: general practitioners' perspectives. Fam Pract 2018 May 23;35(3):318-322. [doi: 10.1093/fampra/cmx100] [Medline: 29088438]
38. Mudge S, Kayes N, McPherson K. Who is in control? Clinicians' view on their role in self-management approaches: a qualitative metasynthesis. BMJ Open 2015 May 5;5(5):e007413. [doi: 10.1136/bmjopen-2014-007413]
39. Zogg JB, Woods SP, Sauceda JA, Wiebe JS, Simoni JM. The role of prospective memory in medication adherence: a review of an emerging literature. J Behav Med 2011 Apr 13;35(1):47-62. [doi: 10.1007/s10865-011-9341-9]

XSL•FO
RenderX

40. Alison Phillips L, Leventhal H, Leventhal EA. Assessing theoretical predictors of long-term medication adherence: Patients' treatment-related beliefs, experiential feedback and habit development. Psychol Health 2013 Oct;28(10):1135-1151. [doi: 10.1080/08870446.2013.793798]
41. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012 Jul;380(9836):37-43. [doi: 10.1016/S0140-6736(12)60240-2]
42. Glynn L, Casey M, Walsh J, Hayes PS, Harte RP, Heaney D. Patients' views and experiences of technology based self-management tools for the treatment of hypertension in the community: A qualitative study. BMC Fam Pract 2015 Sep 9;16:119 [FREE Full text] [doi: 10.1186/s12875-015-0333-7] [Medline: 26354752]
43. Galvin M, Murphy E, Doyle J, Hannigan C, Smith S, Dinsmore J. ProACT: Fostering PPI within the design of digital health solutions for people with multimorbidity. 2018 Presented at: PPIT'18; November 15, 2018; Newcastle, United Kingdom.
44. Cross A, Elliott R, Petrie K, Kuruvilla L, George J. Interventions for improving medication-taking ability and adherence in older adults prescribed multiple medications. Cochrane Database Syst Rev 2020;5(5). [doi: 10.1002/14651858.cd012419.pub2]

Abbreviations

BP: blood pressure
CS-SRM: Common-Sense Self-Regulation Model
GP: general practitioner
SUS: System Usability Scale

Edited by G Eysenbach; submitted 14.12.19; peer-reviewed by S Abdi, Y Bartlett, H Long, A Fioravanti; comments to author 03.02.20; revised version received 05.06.20; accepted 21.06.20; published 28.10.20.

Please cite as:
McBride CM, Morrissey EC, Molloy GJ
Patients' Experiences of Using Smartphone Apps to Support Self-Management and Improve Medication Adherence in Hypertension: Qualitative Study
JMIR Mhealth Uhealth 2020;8(10):e17470
URL: http://mhealth.jmir.org/2020/10/e17470/
doi:10.2196/17470
PMID:33112251

©Ciara M McBride, Eimear C Morrissey, Gerard J Molloy. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 28.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Association Between Usage of an App to Redeem Prescribed Food Benefits and Redemption Behaviors Among the Special Supplemental Nutrition Program for Women, Infants, and Children Participants: Cross-Sectional Study

Qi Zhang[1], PhD; Junzhou Zhang[2], PhD; Kayoung Park[3], PhD; Chuanyi Tang[4], PhD

[1]School of Community and Environmental Health, Old Dominion University, Norfolk, VA, United States

[2]Department of Marketing, Feliciano School of Business, Montclair State University, Montclair, NJ, United States

[3]Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, United States

[4]Department of Marketing, Old Dominion University, Norfolk, VA, United States

Corresponding Author:
Qi Zhang, PhD
School of Community and Environmental Health
Old Dominion University
3130 Health Sciences Building
Old Dominion University
Norfolk, VA, 23529
United States
Phone: 1 7576836870
Email: qzhang@odu.edu

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2020/10/e25073/

Abstract

Background: The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) is one of the most important food assistance programs in the United States, serving 6.4 million low-income, eligible women, infants, and children under 5 years of age in 2019. In the program, participants are prescribed a list of food benefits, which can be redeemed in WIC-authorized stores. However, there are multiple behavioral barriers in the program and the stores that prevent participants from redeeming the benefits fully.

Objective: This study aims to examine the relationship between the use of a widely used mobile phone app, WICShopper, and the redemption of the prescribed food packages.

Methods: WIC administrative data were obtained from West Virginia for the period January 2019 to January 2020 and included 30,440 WIC households that had received food benefits in that period. The redemption rates of 18 WIC food benefits were compared between app users and nonapp users, that is, those who never used the app in the study period. The use behaviors were defined for the app users, including the number of active use benefit cycles, active benefit cycle rates, number of active use days in the cycle, and proportion rates of daytime use. Panel linear regressions were applied to examine how the redemption rates were related to these behaviors over time.

Results: App users consistently had higher average redemption rates than nonapp users; the difference ranged from 3.6% (4.8% relative) for infant formula to 14.3% (40.7% relative) for fish. After controlling for sociodemographics, the coefficients of app use were significantly positive for all benefit categories except for WIC-eligible nutritionals. More active cycles and active days in the cycle were significantly related to redemption rates for all categories, except for frozen juice (coefficient=−0.002, P=.09). Daytime app access was positively associated with redemption rates for most food benefits except only a few, such as infant formula (coefficient=−0.03, P<.001).

Conclusions: Use of the WIC app was significantly related to higher redemption rates across food benefits, although the association varied across benefit categories. More active days were positively related to benefit redemptions across food categories,

XSL•FO
RenderX

and the app's daytime use was positively associated with the redemption of most benefit categories. These findings suggest that the WIC app can be an important tool for the promotion of benefit redemption among WIC participants.

(JMIR Mhealth Uhealth 2020;8(10):e20720) doi:10.2196/20720

KEYWORDS

mobile phone app; WIC; EBT; benefit redemption; mobile phone

Introduction

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) is one of the most successful food assistance programs in the United States, serving approximately 6.4 million low-income pregnant, breastfeeding, postpartum women, infants, and children under 5 years of age in the fiscal year 2019 [1]. The WIC program plays an important role in improving the health and developmental outcomes of infants and children during critical developmental stages, promoting healthier eating, better birth outcomes, lower risk of childhood obesity, and improved household food security [2-7].

To achieve its goal, WIC certifies its enrollees for a certain number of months, that is, the certification period. During that period, an enrollee is prescribed a maximum quantity of supplemental foods, such as breakfast cereal and milk, for each monthly benefit issuance cycle, either a calendar month (eg, July 1-31) or a rolling month (eg, July 15-August 14) [8]. The prescribed food benefits can be redeemed in WIC-authorized stores before the benefits expire [9]. A WIC enrollee becomes a participant only after he or she redeems some or all of the prescribed food packages [8].

Although the WIC food benefits are free for participants, there are barriers in the retail and clinic environment that prevent participants from redeeming their benefits [9-11]. For example, it is challenging for participants to remember the specific eligible brand names or product types in each food category and identify them correctly in stores (eg, Great Value 100% apple juice is authorized in West Virginia, but Minute Maid 100% natural flavor fruit punch is not). Incorrect identification of eligible food may result in denial of the redemption in the checkout process, which may create embarrassment or frustration among WIC participants [12]. Other barriers include difficulty in remembering the benefit expiration date or clinic appointment times [13]. These barriers could contribute to the under-redemption of food benefits. As a result, only 12.6% of the WIC households in Kentucky, Michigan, and Nevada redeemed all their benefits in 2012 [14]. It is expected that when participants are not able to fully redeem the benefits they need and/or when they have negative redemption experiences, some of the unsatisfied participants may drop out of the WIC program. The national rate of participation in the WIC program steadily decreased from a peak in 2011 (64%) to the bottom in 2017 (51%) [15]. Various interventions have been implemented at the national and state levels to improve participation rate and participant satisfaction with the WIC program [16].

Electronic benefits transfer (EBT) is one of the most significant changes implemented to improve WIC administration [17]. Compared with the traditional WIC paper voucher, the EBT card, similar to a debit card, is thought to be easier and more convenient for WIC participants, vendors, and the WIC program. The implementation of the EBT system has been completed in most states; the deadline for all states to make the transition is October 1, 2020. Transition to EBT has been associated with a higher redemption of food benefits, reduced stigma at checkout, and improved overall retail experience [12,18]. More importantly, EBT allows integrating the electronic transaction system and other WIC administrative systems, which can provide an eWIC platform that makes it possible for redemption data mining and future technology innovations.

On the basis of the eWIC system, WIC apps have become a prevalent innovation across WIC state agencies. Although a federal agency funds the WIC, the program is operated at the state level in 50 states, among 34 Indian tribal organizations, in the District of Columbia, and in five US territories [19]. A recent review indicated that 17 different apps were used by participants in 37 state WIC agencies [20]. These apps include some or all of 4 main features: facilitating benefit redemption in stores, such as identifying eligible food, facilitating clinical experience, such as scheduling reminders, supplemental information (eg, recipes), and nutrition education required by the WIC program [20]. Although many WIC participants have used these apps, little research has been conducted to rigorously evaluate them, except for 1 small-scale qualitative study that demonstrated a high priority among WIC participants to access information through a WIC app [21].

Evaluating the effectiveness of WIC apps is critical for WIC agencies, participants, and researchers. First, evaluation can assist the decision making of the state agencies that have not adopted the app. Second, evaluation can help app creators identify gaps and improve app functions for WIC participants. Third, the WIC app has the potential to become a health promotion platform. Specifically, evaluation helps researchers understand WIC participants' health-related behaviors. This information can be used to develop future intervention and health promotion programs. Finally, all these improvements can eventually help WIC participants navigate within the program more easily so that they can benefit more from it.

This study aims to evaluate the effectiveness of WIC apps to fill this knowledge gap. To the best of our knowledge, this is the first state-level WIC app evaluation in the United States that compares the redemption outcomes between WIC app users and nonapp users. We selected the WICShopper app as the research context because it has been adopted by 32 state WIC agencies, 3 Indian tribes, and 1 US territory [22], having the largest market share and the highest number of installations among all WIC apps reviewed by Weber et al [20]. In contrast, most of the other WIC apps, such as the California WIC App, are limited to 1 state and have fewer users. The WICShopper

XSL•FO
RenderX

app allows the participants to check real-time benefits, eligibility of food products, WIC vendor and office locations, and recipes. The app was rated 4.5 in the Apple app store and 4.5 on Google Play, one of the highest ratings among the WIC apps, with the ratings ranging between 2 and 5; for example, Alabama WIC was rated 2.3 in the Apple app store and Arizona WIC was rated 4.7 on Google Play. Therefore, the WICShopper app is an appropriate candidate for a large-scale evaluation.

As the main purpose of the WICShopper app is to facilitate benefit redemption via multiple features, such as checking benefit balance and scanning the bar code to check eligibility, this study aims to examine how the WICShopper app usage is associated with redemption in various food benefit categories by comparing the redemption outcomes between WICShopper app users and nonapp users.

Methods

Study Design and Sample Selection

To examine the relationship between app usage and WIC benefit redemption, we adopted a mix of cross-sectional and longitudinal study designs. The data source was the administrative data from the West Virginia WIC agency merged with the app usage data provided by JPMA, Inc, the developer of the WICShopper app. The West Virginia WIC agency was one of the first 3 state agencies to adopt this app (in 2013).

The WIC administrative data included participants' sociodemographics, participation status, food benefit prescriptions, and redemption records. The redemption data set contains detailed transaction information, such as transaction dates, description of the food categories, units redeemed, and the number of units redeemed. The app activity data have a timestamp that records when a user turns on the app. The WIC administrative data were linked with the app data by WIC household identification number. As 1 household normally has 1 EBT card, which can have multiple packages for multiple participants, such as a mother and her children, we used the household as the unit of analysis. The study period was from January 19, 2019, to January 18, 2020, because of the archiving cycles of the app data.

Statistical Analyses

Measures

The main outcome was the redemption rate of a household in a benefit cycle, which is defined as the sum of the redeemed amount for a food category divided by the prescribed amount for that food category per household in a benefit cycle. The benefit cycle is usually a month, but it could be shorter than a month if a participant has a late visit to the WIC clinic for benefit renewal. The redemption rate was calculated using the following 18 food categories: infant meats, frozen juice, legumes, whole-grain bread, infant cereal, fish, adult cereal, WIC-eligible nutritionals, yogurt, shelf-stable juice, low-fat milk, infant fruit and vegetable, eggs, cheese, whole milk, exempt infant formula, infant formula, and cash value benefits. Each participant's prescribed food package may vary based on participation circumstances. For example, the prescribed benefits for pregnant women are different from those prescribed for breastfeeding women [23].

The primary explanatory variable was the indicator of app users versus nonapp users, which was defined by whether a participating household had ever used the app during the entire study period or not. For example, if a WIC household used the app in March but not in April, it was still classified as an app user in the year. Secondary explanatory variables calculated among the app users were: the *number of active app cycles*, defined as the number of benefit cycles in which the participant actively used the app; *active cycle rate*, defined as the percentage of active app cycles out of all benefit cycles for that household in the study period; the *number of active app days*, defined as the number of days in a benefit cycle the participant used the app; and *daytime rate*, defined as the percentage of app usage from 6 AM to 5 PM in a benefit cycle. As each household has a unique ID to log into the app, activity was recorded at the household level, regardless of how many individual users accessed it.

The control variables included the participants' demographic information: race (White or non-White), whether the household had infant, child, or female participants, and the number of WIC participants and overall household size, including members not participating in WIC. According to the US Census, 92% of residents in West Virginia in 2019 were non-Hispanic Whites [24]. In our study sample, 26,630 of the 30,440 households (87.5%) were non-Hispanic Whites. If the non-White households were further broken down, the sample size would be too small to be representative of the specific groups. Therefore, we combined all non-Whites into 1 racial or ethnic group. These control variables were at the household level; in the case of a household with multiple participants, for example, a mixed-race household, we used the race of the eldest WIC participant.

Statistical Methods

We first estimated the descriptive statistics for sociodemographics and app usage. Differences in sociodemographics between app users and nonapp users were then tested using Chi-square or Wilcoxon rank-sum tests. App usage activities were also tested across the sociodemographics. Next, we estimated the average redemption rates across food categories and tested the difference in the redemption rates between app and nonapp users by using the Wilcoxon rank-sum test. The absolute difference in the redemption rates between app and nonapp users was calculated as the app users' average redemption rate minus that of nonapp users, whereas the relative difference was calculated as the absolute difference divided by the average nonapp users' redemption rate. We divided the 18 food groups into low, medium, and high redemption groups with 6 food categories per group. We plotted the average redemption rates by app usage and groups with absolute and relative differences.

To examine the relationship between the redemption rate and app usage behaviors, we applied 3 sets of linear regressions across food categories. First, we regressed the average annual redemption rate on the app usage indicator among all households (app users and nonusers) in the study period. We then focused only on app users with additional analyses. The second set of

XSL•FO
RenderX

regressions examined how the average annual redemption rate was related to the number of active app cycles or the active app cycle rates. The third set of regressions were panel analyses across benefit cycles. Random-effects models with an unbalanced panel were employed in our empirical evaluation of WIC participants' redemption behaviors. Sociodemographics were controlled in all sets of regressions. Statistical significance was set at $P<.05$, and the R software was employed to conduct the statistical analyses [25]. This study was approved by the institutional review board of the Old Dominion University.

Results

Descriptive Statistics

Table 1 compares the descriptive statistics of the WIC participants' sociodemographics between app users and nonapp users. The entire study sample included 30,440 WIC households in West Virginia during the study period. Non-Hispanic White accounted for 87.48% (26,630/30,440) of the participating households. The prevalence rates of infants, children, and women were 29.05% (8844/30,440), 76.17% (23,187/30,440), and 70.85% (21,567/30,440), respectively. Of the 30,440 households, 13,925 (45.75%) had 2 WIC participants; households with 1 WIC participant and 3 or more participants accounted for a similar percentage of the study sample (8015/30,440, 26.33%, and 8500/30,440, 27.92%, respectively). On average, the household size was 3.84. Of the study population, 72.26% (21,996/30,440) of the households used the WICShopper app at least once in the study period. Non-Hispanic White households had a significantly higher percentage of app users (19,496/26,630, 73.21%) than non-White households (2500/3810, 65.62%; $P<.001$). The percentage of app users increased when there were more participants in a household (5070/8015, 63.26%; 10,358/13,925, 74.38%; and 6568/8500, 77.27% for households with 1, 2, and 3 or more participants, respectively; $P<.001$), which is consistent with the result that app users, on average, tended to have a larger household ($P<.001$).

Table 1. Descriptive statistics of the sociodemographics from participants of the Special Supplemental Nutrition Program for Women, Infants, and Children in West Virginia (N=30,440).

Variables	Whole sample, n	Nonapp users (n=8444), n (%)	App users (n=21,996), n (%)	*P* value[a]
Racial group				<.001
White	26,630	7134 (26.79)	19,496 (73.21)	
Non-White	3810	1310 (34.38)	2500 (65.62)	
Did the household have an infant participant?				<.001
Yes	8844	2092 (23.65)	6752 (76.35)	
No	21,596	6352 (29.41)	15,244 (70.59)	
Did the household have a child participant?				.92
Yes	23,187	6436 (27.76)	16,751 (72.24)	
No	7253	2008 (27.69)	5245 (72.31)	
Did the household have a female participant?				<.001
Yes	21,567	5266 (24.42)	16,301 (75.58)	
No	8873	3178 (35.82)	5695 (64.18)	
Number of WIC[b] participants				<.001
1	8015	2945 (36.74)	5070 (63.26)	
2	13,925	3567 (25.62)	10,358 (74.38)	
≥3	8500	1932 (22.73)	6568 (77.27)	
Household size (people), mean (SD)	3.84 (1.60)	3.73 (1.69)	3.88 (1.56)	<.001

[a]Statistical tests performed: Chi-square test of independence and Wilcoxon rank-sum test.

[b]WIC: The Special Supplemental Nutrition Program for Women, Infants, and Children.

App Usage Behaviors

Table 2 presents the pattern of WIC app users' app usage behavior across key demographic variables. At the aggregated level, these app users actively used the WIC app in 7.93 benefit cycles, which was equivalent to 85.8% of all their benefit cycles in the study period. App users activated the app for 4.05 days per monthly benefit issuance cycle, which is, on average, equivalent to once per week. On average, 63.3% of app usage behaviors occurred during the daytime. White households were more active in app usage than non-White households, except for the number of active app days ($P=.70$).

In addition, households with infant or woman participants tended to be more active in terms of using the WIC app in a prescribed benefit cycle. For example, households with infants and households with woman participants used the WIC app about 4.38 days and 4.11 days in 1 benefit cycle, respectively, whereas households without infants and households without woman

XSL•FO
RenderX

participants averaged 3.91 days and 3.90 days, respectively (*P*<.001). In addition, households with children were less active in using the WIC app, averaging 4.01 days of WIC app use in a prescribed benefit cycle compared with 4.19 days for households without children (*P*<.001).

Table 2. The Special Supplemental Nutrition Program for Women, Infants, and Children app users' characteristics at the household level (n=21,996). Statistics are presented as mean (SD), and the tests performed were the Wilcoxon rank-sum test and Kruskal-Wallis test.

Variables	Number of active app cycles	*P* value	Active cycle rate, n (%)	*P* value	Number of active app days	*P* value	Daytime rate, n (%)	*P* value
Whole sample	7.93 (3.35)	N/A[a]	85.8 (23.5)	N/A	4.05 (2.05)	N/A	63.3 (22.8)	N/A
Race group		.005	N/A	.02	N/A	.70	N/A	<.001
White	7.97 (3.34)		85.9 (23.4)		4.05 (2.04)		63.6 (22.8)	
Non-White	7.78 (3.38)		84.6 (24.2)		4.08 (2.09)		61.1 (22.9)	
Did the household have an infant participant?		<.001	N/A	.40	N/A	<.001	N/A	<.001
Yes	7.94 (3.07)		85.7 (23.5)		4.38 (2.08)		62.4 (22.2)	
No	7.95 (3.46)		85.8 (23.5)		3.91 (2.01)		63.7 (23.1)	
Did the household have a child participant?		<.001	N/A	.007	N/A	<.001	N/A	<.001
Yes	8.40 (3.23)		86.1 (23.2)		4.01 (2.03)		64.2 (22.6)	
No	6.50 (3.29)		84.6 (24.2)		4.19 (2.07)		60.3 (23.3)	
Did the household have a woman participant?		.20	N/A	.90	N/A	<.001	N/A	<.001
Yes	7.98 (3.31)		86.0 (23.2)		4.11 (2.04)		62.9 (22.5)	
No	7.87 (3.44)		85.2 (24.4)		3.90 (2.05)		64.5 (23.6)	
Number of WIC[b] participants		<.001	N/A	<.001	N/A	<.001	N/A	<.001
1	6.65 (3.63)		84.4 (24.8)		3.82 (2.07)		62.3 (24.4)	
2	8.13 (3.20)		85.3 (23.6)		3.94 (1.95)		62.5 (22.7)	
≥3	8.66 (3.04)		87.5 (22.1)		4.42 (2.13)		65.3 (21.7)	

[a]N/A: not applicable.

[b]WIC: The Special Supplemental Nutrition Program for Women, Infants, and Children.

Redemption Rates and App Usages

Multimedia Appendices 1-3 present the average redemption rates of 18 WIC food categories among app users and nonapp users. We divided the food benefits into 3 groups (low, medium, and high redemption) according to the rankings of their redemption rates from the lowest to highest (infant meat, 27.4%, to infant formula, 77.9%). In all groups, app users had a higher redemption rate than nonusers, and the difference was statistically significant in all food categories except in WIC-eligible nutritionals (*P*=.12).

Multimedia Appendices 4 and 5 compare the differences in the redemption rates between app users and nonapp users across food benefits. First, the 18 food benefits were divided into 3 groups based on the redemption rates (6 benefits in each group: low, medium, and high redemption rate groups). The figure marked these food benefits with red, yellow, and green bars, respectively. These bars were then ordered based on the magnitude of the differences, from the smallest to the largest difference. In Multimedia Appendix 4, the 3 groups of colored bars were almost evenly distributed. However, in Multimedia Appendix 5, high (green bars) and medium (yellow bars) redemption groups were more concentrated toward the left end of the axis, whereas the low redemption groups (red bars) were more concentrated toward the right end of the axis. For example, the bottom 5 food groups in terms of relative differences were either high or medium redemption groups (green or yellow bars toward the left: infant formula, exempt infant formula, WIC-eligible nutritionals, low-fat milk, and whole-grain bread), whereas the top 5 groups in relative difference were all in low redemption groups (red bars toward the right: fish, frozen juice, infant cereal, infant meats, and legumes). The patterns indicate that the WIC app may be more helpful in improving redemption rates relatively for the food benefits with lower redemption rates, that is, less-popular food benefits.

The regression results for all food categories are presented in Tables 3 and 4. Model 1 examined the association between the annual redemption rate and the app usage indicator for all households, including app users and nonapp users. Models 2 and 3 focused only on app users, with the goal of examining the effect of the number of active app cycles and the percentage of active app cycles in the study period. Models 4 and 5 in Table 4 are panel analyses of the relationship between redemption rates and the number of app active days, and the percentage of app usage in the daytime across benefit cycles. All models controlled sociodemographic variables.

XSL•FO
RenderX

Table 3. Linear regression of the Special Supplemental Nutrition Program for Women, Infants, and Children benefit redemption rates on app activities by food categories in West Virginia participating households (N=30,440 and n=21,996 for app users only).

Variables (food category)	Whole sample[a]			App users only[a]					
	App use (Yes=1; No=0); model 1			Number of active app cycles; model 2			Active cycles rate; model 3		
	Coefficient	SE	*P* value	Coefficient	SE	*P* value	Coefficient	SE	*P* value
Infant meats	0.06	0.03	.045	0.02	0.005	<.001	0.30	0.05	<.001
Frozen juice	0.08	0.007	<.001	−0.002	0.001	.09	0.24	0.01	<.001
Legumes	0.08	0.005	<.001	0.007	0.001	<.001	0.29	0.01	<.001
Whole-grain bread	0.07	0.005	<.001	0.006	0.001	<.001	0.27	0.01	<.001
Infant cereal	0.11	0.008	<.001	0.02	0.001	<.001	0.25	0.02	<.001
Fish	0.14	0.02	<.001	0.007	0.003	.02	0.47	0.04	<.001
Adult cereal	0.07	0.004	<.001	0.005	0.001	<.001	0.29	0.009	<.001
WIC[b]-eligible nutritionals	0.04	0.03	.17	0.01	0.005	.046	0.23	0.06	<.001
Yogurt	0.10	0.005	<.001	0.004	0.001	<.001	0.32	0.01	<.001
Shelf-stable juice	0.08	0.005	<.001	0.01	0.001	<.001	0.30	0.01	<.001
Low-fat milk	0.08	0.005	<.001	0.01	0.001	<.001	0.30	0.009	<.001
Infant fruit and vegetable	0.10	0.007	<.001	0.02	0.001	<.001	0.25	0.01	<.001
Eggs	0.10	0.004	<.001	0.01	0.001	<.001	0.33	0.008	<.001
Cheese	0.09	0.004	<.001	0.01	0.001	<.001	0.34	0.009	<.001
Cash value benefit	0.10	0.004	<.001	0.01	0.001	<.001	0.30	0.008	<.001
Whole milk	0.08	0.007	<.001	0.02	0.001	<.001	0.29	0.01	<.001
Exempt infant formula	0.04	0.01	.001	0.006	0.002	.005	0.12	0.03	<.001
Infant formula	0.03	.005	<.001	0.009	0.001	<.001	0.09	0.01	<.001

[a]Sociodemographics controlled: race, whether the household had an infant, child, female participants, the number of the Special Supplemental Nutrition Program for Women, Infants, and Children participants in a household, and the overall household size (people).

[b]WIC: The Special Supplemental Nutrition Program for Women, Infants, and Children.

XSL•FO
RenderX

Table 4. Panel regression of the Special Supplemental Nutrition Program for Women, Infants, and Children benefit redemption rates on app activities by food categories in West Virginia participating households (n=21,996 for app users only).

Variables (food category)	Number of active app days[a], model 4			Daytime rate[a], model 5		
	Coefficient	SE	*P* value	Coefficient	SE	*P* value
Infant meats	0.02	0.003	<.001	−0.006	0.02	.79
Frozen juice	0.02	0.001	<.001	0.01	0.006	.04
Legumes	0.02	0.001	<.001	0.01	0.004	.003
Whole-grain bread	0.02	<0.001	<.001	0.02	0.004	<.001
Infant cereal	0.03	0.001	<.001	0.02	0.008	.007
Fish	0.03	0.002	<.001	0.02	0.02	.31
Adult cereal	0.03	<0.001	<.001	0.008	0.004	.04
WIC[b]-eligible nutritionals	0.02	0.002	<.001	0.02	0.02	.49
Yogurt	0.02	0.001	<.001	0.01	0.005	.03
Shelf-stable juice	0.03	0.001	<.001	0.01	0.005	.004
Low-fat milk	0.03	<0.001	<.001	0.006	0.003	.10
Infant fruit and vegetable	0.03	0.001	<.001	0.02	0.007	.03
Eggs	0.03	<0.001	<.001	0.02	0.004	<.001
Cheese	0.02	<0.001	<.001	0.02	0.004	<.001
Cash value benefit	0.02	<0.001	<.001	0.02	0.003	<.001
Whole milk	0.03	0.001	<.001	−0.005	0.006	.35
Exempt infant formula	0.01	0.001	<.001	−0.02	0.01	.10
Infant formula	0.008	0.001	<.001	−0.03	0.005	<.001

[a]Sociodemographics controlled: race, whether the household had an infant, child, and female participants, the number of The Special Supplemental Nutrition Program for Women, Infants, and Children participants in a household, and the overall household size (people).

[b]WIC: The Special Supplemental Nutrition Program for Women, Infants, and Children.

In Model 1, we observed that app usage was significantly associated with higher redemption rates for almost all food categories except WIC-eligible nutritionals (coefficient=0.04, *P*=.17). After adjusting for sociodemographics, using an app generated the lowest increase in the average redemption rate, as in the case of 3 percentage points in infant formula, and the highest increase, as in the case of 14 percentage points in fish. In app users (Model 2), a higher number of active app cycles was significantly associated with a higher redemption rate, except in the category of frozen juice (48 oz; coefficient=−0.002, *P*=.09). The largest number of active app cycles was observed for infant fruits and vegetables, infant meats, infant cereals, and whole milk (coefficient=0.02, *P*<.001), whereas the smallest coefficients were observed for yogurt (coefficient=0.004, *P*<.001). These results indicated that having 1 additional active app cycle was associated with a 0.4 to 2 percentage point increase in redemption rates. Similarly, increasing the active cycle rates by 10% was associated with a 0.9 to 4.7 percentage point increase (infant formula vs fish) in redemption rates (Model 3).

In Table 4, Models 4 to 5 present how app activity per cycle was associated with the redemption rates of the food benefits. Model 4 indicates that the number of days of the WICShopper app usage was positively associated with the redemption rates, with statistical significance in all food categories (*P*<.001). Here, again, the smallest and largest coefficients were observed in infant formula and fish. Model 5 presents the coefficients of daytime usage. The percentage of use of the WIC app during the daytime was significantly associated with a higher redemption rate for most food categories, except for infant meats, fish, WIC-eligible nutritionals, low-fat milk, whole milk, and exempt infant formula. Notably, using the app during the daytime was associated with lower redemption rates in infant formula (coefficient=−0.03, *P*<.001).

Overall, the regression results suggest that the use of the WICShopper app was significantly associated with higher redemption rates. However, the positive app effect may diminish with the popularity of the various food benefits; that is, app usage may not have a large effect on the redemption of some popular food benefits (eg, infant formula).

Discussion

Principal Findings

To the best of our knowledge, this is the first state-level study examining the association between WIC app usage and WIC benefit redemption. We found a significant positive association between app usage and redemption rates in all food benefit categories except for WIC-eligible nutritionals. These results are encouraging for WIC policymakers who had concerns about

XSL•FO
RenderX

the prevalent underredemption of benefits among WIC participants [14,26,27]. In general, app user redemption rates were at least 10% higher than those of nonusers in all food categories except for WIC-eligible nutritionals, exempt infant formula, and regular infant formula. Given that WIC-eligible nutritionals and exempt infant formula require medical documentation to be prescribed, infant formula was the only regular food benefit for which redemption was not related to app usage on a large scale, although the association was still statistically significant. Given the positive findings of this study, state WIC agencies may consider increasing access to the app for participants.

Several factors may explain the weak relationship between app usage and infant formula. First, infant formula was the most redeemed food category, with an average redemption rate of 77.9%. Given that the redemption rate was already very high or *near the ceiling*, using an app may not be able to further improve the redemption rate. Thus, this ceiling effect may diminish the scale of the app effect on this redemption. Second, infant formula is the most valuable food benefit for participants, who may treat the redemption of the formula as their highest priority. Usually, each state agency contracts with only one formula brand, such as Similac in West Virginia. On the basis of the findings from behavioral economics, the high priority of obtaining the infant formula, and the ease in identifying the formula product may reduce the need for app assistance to facilitate redemption [12,28]. As a result, app usage is unlikely to change the infant formula redemption rate on a large scale.

In addition to infant formula, app users had much higher redemption rates than nonapp users in the rest of the regular food categories. The relative difference in redemption rate between app users and nonusers was most evident in the least-redeemed food categories, including infant meats, frozen juice, legumes, infant cereal, and fish. There are several potential theoretical explanations for these results. First, the *temptation bundling* theory suggests that combining the less-preferred activities with preferred activities may increase both activities [29]. With the WICShopper app, participants can easily check their remaining benefits for all food categories in their prescribed food package as a bundle. Creating a *bundling effect* within the app may motivate participants to redeem all benefits together, even though some of the bundled food benefits may not be among their most preferred. On the other hand, nonapp users may not have the opportunity to view all the food benefit balances in real time and in a convenient way. As a result, some benefits may remain unredeemed. Second, using the app can help identify the eligible brand and product type more easily for certain food categories, such as infant cereal, which may reduce redemption barriers [26]. However, for other food categories with fewer eligible brands, such as infant formula, the app may have a limited impact on brand choices. Finally, the difference may be caused by factors other than app usage. App users and nonapp users may have different preferences and behavioral patterns extrinsic to the app itself. For example, app users may care more about their welfare than nonusers. Thus, they are more likely to download the app and redeem more benefits than nonusers, even if certain food categories are not popular choices. Although extensive research has studied the self-selection bias in WIC participation, no research has examined the potential self-selection bias in WIC app usage, which should be addressed in future studies using a randomized experiment or statistical methods, such as propensity score or instrumental variable [30,31].

The panel analyses among the app users further examined how the app activities across cycles may be associated with redemption rates. The frequency and timing of app usage per cycle (eg, in the daytime) were associated with redemption rates in most food benefits. Given the app's multiple functions, using the app may not only improve the participants' shopping experience but also help with their food planning, as they can check the remaining balance at any time. The ability to check the benefit balance was ranked as the most useful app feature by WIC participants [32]. About 82% of participants rated it as *very useful*, but lower percentage of participants rated other features as *very useful*, such as checking the WIC shopping guide (73%), checking eligible food products (71%), or locating WIC-authorized stores (62%) [32]. Therefore, the WIC app may help engage participants in the WIC program and help them better utilize the resources the program provides in general.

The transition to EBT is deemed one of the most significant technological changes in the WIC program over the past 4 decades [14]. One of the main purposes of EBT transition is to facilitate the benefit redemption experience. However, a recent state-level study suggests that EBT transition did not significantly change the redemption of produce or infant formula [18]. Although the EBT system itself may not exert a significant impact on every aspect of the WIC program, it provides a necessary base for further technological innovations to improve WIC operations and participant experience [20,33]. For example, we found that WIC app adoption and frequent app usage were positively associated with the redemption of all food benefits, except WIC-eligible nutritionals, in West Virginia. Researchers may develop more EBT-based or app-based interventions to improve participant experience in the WIC program [34].

Limitations

Although this is the first state-level study to provide empirical evidence of positive relationships between app usage and WIC benefit redemptions, the results need to be interpreted carefully because of several limitations. First, the pooled analyses were essentially cross-sectional, so only the association instead of the causality can be concluded. More studies are needed to examine the determinants of app adoption among WIC participants and how self-selection bias can affect study results. More rigorous study designs, such as a field experiment or advanced statistical methods (such as propensity score or instrumental variable), may help address the self-selection bias of app users so that the real app effect can be more accurately estimated. Second, because there are a large number of users from all states adopting the WICShopper app, app developers do not archive detailed app activity data, such as the pages visited or the sequence of the clicks. Therefore, the specific features the app users used are not known. This creates challenges in linking app feature usage directly to redemption behaviors.

XSL•FO
RenderX

Moreover, there is no information about whether the participating households ever download the app or even have smartphones. Finally, WIC administrative data have a limited number of sociodemographic variables that can be controlled in regression models. As applicants can be eligible for WIC if they are participants in other means-tested social welfare programs, such as Medicaid or the Supplemental Nutrition Assistance Program (SNAP) [35], income was not required to be reported in the WIC database, which resulted in 85.3% (25,973/30,440) of the income data being missing. Therefore, we did not control for income in the analyses. As redemption behaviors can be influenced by other individuals, household, and community factors, more insights might be gained by linking our data with other state-level data sets, such as SNAP participant data or Medicaid participant data. However, such a step would require additional administrative reviews within other state agencies, which could be challenging. Despite the above limitations and challenges, this study employed the best available data set to generate important preliminary findings on the impact of technology on WIC participation—app users had greater redemption rates than nonapp users.

Conclusions

Employing WIC administrative data and WIC app usage data from West Virginia, this study provides important evidence about the positive association between WIC app usage and the redemption of WIC food benefits. These results can help the decision-making process in state WIC agencies that have not yet adopted the app in their program operations. Moreover, because more frequent app usage is related to higher redemption rates in most food categories, state agencies may seek to promote more app downloads and usage among WIC participants. Finally, as most state agencies will transition to the EBT system by October 2020 (although some states have received waivers), more app-based innovations may be developed and implemented to improve the participant experience, through which WIC participant retention can be improved. To maximize the benefits of the WIC app, state WIC agencies, app developers, and retail stores need to work together closely to ensure that the latest approved food products and redemption data are integrated, updated, and shared in a timely and efficient manner.

Acknowledgments

QZ and CT conceptualized the research design. QZ secured the data from the West Virginia WIC agency and the app developer. JZ and KP conducted data analyses; QZ and JZ drafted the manuscript, and CT and KP revised it. The authors thank Dr. Denise Ferris, Ms. Kathy Legg, Mr. Ryan Magee, and other colleagues from West Virginia WIC agency and JPMA for data access and technical support. This project was supported by Healthy Eating Research, a national program of the Robert Wood Johnson Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Average annual redemption rates among all subjects, app users, and nonapp users (low redemption food group; *: $P<.001$).
[PNG File , 176 KB - mhealth_v8i10e20720_app1.png]

Multimedia Appendix 2
Average annual redemption rates among all subjects, app users, and nonapp users (medium redemption food group; *$P<.001$; WIC: The Special Supplemental Nutrition Program for Women, Infants, and Children).
[PNG File , 199 KB - mhealth_v8i10e20720_app2.png]

Multimedia Appendix 3
Average annual redemption rates among all subjects, app users, and nonapp users (high redemption food group; *$P<.001$).
[PNG File , 205 KB - mhealth_v8i10e20720_app3.png]

Multimedia Appendix 4
Absolute differences in the redemption rates between app users and nonapp users.
[PNG File , 351 KB - mhealth_v8i10e20720_app4.png]

Multimedia Appendix 5
Relative differences in the redemption rates between app users and nonapp users.
[PNG File , 330 KB - mhealth_v8i10e20720_app5.png]

References

XSL•FO
RenderX

1. USDA-FNS. WIC data tables. USDA Food and Nutrition Service. 2020 Jul 17. URL: https://www.fns.usda.gov/pd/wic-program [accessed 2020-04-11]
2. Bitler MP, Currie J. Does WIC work? The effects of WIC on pregnancy and birth outcomes. J Policy Anal Manage 2005;24(1):73-91. [doi: 10.1002/pam.20070] [Medline: 15584177]
3. Hoynes H, Page M, Stevens AH. Can targeted transfers improve birth outcomes?: evidence from the introduction of the WIC program. J Public Econ 2011 Aug;95(7-8):813-827. [doi: 10.1016/j.jpubeco.2010.12.006]
4. Metallinos-Katsaras E, Gorman KS, Wilde P, Kallio J. A longitudinal study of WIC participation on household food insecurity. Matern Child Health J 2011 Jul;15(5):627-633. [doi: 10.1007/s10995-010-0616-5] [Medline: 20455015]
5. Whaley SE, Ritchie LD, Spector P, Gomez J. Revised WIC food package improves diets of WIC families. J Nutr Educ Behav 2012;44(3):204-209. [doi: 10.1016/j.jneb.2011.09.011] [Medline: 22406013]
6. Chiasson MA, Findley SE, Sekhobo JP, Scheinmann R, Edmunds LS, Faly AS, et al. Changing WIC changes what children eat. Obesity (Silver Spring) 2013 Jul;21(7):1423-1429 [FREE Full text] [doi: 10.1002/oby.20295] [Medline: 23703806]
7. Koleilat M, Whaley SE, Esguerra KB, Sekhobo JP. The role of WIC in obesity prevention. Curr Pediatr Rep 2017 Jul 1;5(3):132-141. [doi: 10.1007/s40124-017-0135-6]
8. Daniels PN. Certification and monthly food benefits issuance cycles and reporting monthly participation on the FNS-798. USDA Food and Nutrition Service. 2007. URL: https://www.fns.usda.gov/wic/certification-and-monthly-food-benefits-issuance-cycles-and-reporting-monthly-participation-fns-798 [accessed 2020-07-23]
9. National Academies of Sciences, Engineering, and Medicine. Review of WIC Food Packages Improving Balance And Choice Final Report. Washington, DC: The National Academies Press; 2017.
10. Woelfel ML, Abusabha R, Pruzek R, Stratton H, Chen SG, Edmunds LS. Barriers to the use of WIC services. J Am Diet Assoc 2004 May;104(5):736-743. [doi: 10.1016/j.jada.2004.02.028] [Medline: 15127057]
11. Ritchie LD, Whaley SE, Crocker NJ. Satisfaction of California WIC participants with food package changes. J Nutr Educ Behav 2014;46(3 Suppl):S71-S78. [doi: 10.1016/j.jneb.2014.01.009] [Medline: 24810000]
12. Chauvenet C, De Marco M, Barnes C, Ammerman AS. WIC recipients in the retail environment: a qualitative study assessing customer experience and satisfaction. J Acad Nutr Diet 2019 Mar;119(3):416-24.e2. [doi: 10.1016/j.jand.2018.09.003] [Medline: 30502034]
13. Summary of WIC state agency strategies for increasing child retention, July 2014. WIC Works Resource System - USDA. 2014 Jul. URL: https://wicworks.fns.usda.gov/sites/default/files/media/document/ChildRetentionStrategiesReport.pdf [accessed 2020-04-11]
14. Phillips D, Bell L, Morgan R, Pooler J. Transition to EBT in WIC: review of impact and examination of participant redemption patterns. Final report. Altarum Institute. 2014. URL: https://altarum.org/sites/default/files/uploaded-publication-files/Altarum_Transition%20to%20WIC%20EBT_Final%20Report_071614.pdf [accessed 2020-04-11]
15. Neuberger Z. Streamlining and modernizing WIC enrollment with current rules and funding, WIC agencies can simplify certification practices to reach more eligible families. Center on Budget and Policy Priorities. 2020 Feb 20. URL: https://www.cbpp.org/sites/default/files/atoms/files/8-30-19fa.pdf [accessed 2020-04-11]
16. Food Research & Action Center. Making WIC work better: strategies to reach more women and children and strengthen benefits use. 2019 May. URL: https://frac.org/wp-content/uploads/Making-WIC-Work-Better-Full-Report.pdf [accessed 2020-04-11]
17. WIC EBT activities. USDA Food and Nutrition Service. 2020 Sep. URL: https://www.fns.usda.gov/wic/wic-ebt-activities [accessed 2020-04-11]
18. Hanks AS, Gunther C, Lillard D, Scharff RL. From paper to plastic: understanding the impact of eWIC on WIC recipient behavior. Food Policy 2019 Feb;83:83-91. [doi: 10.1016/j.foodpol.2018.12.002]
19. About WIC- WIC at a glance. USDA Food and Nutrition Service. 2013 Oct 10. URL: https://www.fns.usda.gov/wic/about-wic-wic-glance [accessed 2020-04-12]
20. Weber SJ, Dawson D, Greene H, Hull PC. Mobile phone apps for low-income participants in a public health nutrition program for women, infants, and children (WIC): review and analysis of features. JMIR Mhealth Uhealth 2018 Nov 19;6(11):e12261 [FREE Full text] [doi: 10.2196/12261] [Medline: 30455172]
21. Friedman LB, Silva M, Smith K. A focus group study observing maternal intention to use a WIC education app. Am J Health Behav 2018 Nov 1;42(6):110-123. [doi: 10.5993/AJHB.42.6.11] [Medline: 30158006]
22. JPMA, Inc. WICShopper: the mobile app for WIC participants. WICShopper. URL: https://ebtshopper.com/ [accessed 2020-04-12]
23. WIC Approved Shopping Guide. West Virginia Department of Health and Human Resources. 2018 Oct 1. URL: https://dhhr.wv.gov/WIC/Pages/WIC-Approved-Shopping-Guide.aspx [accessed 2020-04-15]
24. QuickFacts: West Virginia. United States Census Bureau. 2019 Jul 1. URL: https://www.census.gov/quickfacts/WV [accessed 2020-07-23]
25. The R Project for Statistical Computing. R Foundation for Statistical Computing. 2012. URL: http://www.r-project.org/ [accessed 2020-04-11]
26. Gleason S, Pooler J. The effects of changes in WIC food packages on redemptions. United States Department of Agriculture. 2011 Dec. URL: https://naldc.nal.usda.gov/download/50613/PDF [accessed 2020-04-11]

XSL•FO
RenderX

27. Kim LP, Whaley SE, Gradziel PH, Crocker NJ, Ritchie LD, Harrison GG. Mothers prefer fresh fruits and vegetables over jarred baby fruits and vegetables in the new Special Supplemental Nutrition Program for Women, Infants, and Children food package. J Nutr Educ Behav 2013;45(6):723-727. [doi: 10.1016/j.jneb.2013.01.022] [Medline: 23591317]
28. Zhang Q, Tang C, McLaughlin PW, Diggs L. Individual and store characteristics associated with brand choices in select food category redemptions among WIC participants in Virginia. Int J Environ Res Public Health 2017 Mar 31;14(4):364 [FREE Full text] [doi: 10.3390/ijerph14040364] [Medline: 28362350]
29. Milkman KL, Minson JA, Volpp KG. Holding the hunger games hostage at the gym: an evaluation of temptation bundling. Manage Sci 2014 Feb;60(2):283-299 [FREE Full text] [doi: 10.1287/mnsc.2013.1784] [Medline: 25843979]
30. Guthrie JF, Catellier DJ, Jacquier EF, Eldridge AL, Johnson WL, Lutes AC, et al. WIC and non-WIC infants and children differ in usage of some WIC-provided foods. J Nutr 2018 Sep 1;148(suppl_3):1547S-1556S [FREE Full text] [doi: 10.1093/jn/nxy157] [Medline: 30247584]
31. Fang D, Thomsen MR, Nayga Jr RM, Novotny AM. WIC participation and relative quality of household food purchases: evidence from FoodAPS. South Econ J 2019;86(1):83-105. [doi: 10.3386/w25291]
32. Bensley RJ, Hovis A, Horton KD, Loyo JJ, Bensley KM, Phillips D, et al. Accessibility and preferred use of online web applications among WIC participants with internet access. J Nutr Educ Behav 2014;46(3 Suppl):S87-S92. [doi: 10.1016/j.jneb.2014.02.007] [Medline: 24810002]
33. Hull P, Emerson JS, Quirk ME, Canedo JR, Jones JL, Vylegzhanina V, et al. A smartphone app for families with preschool-aged children in a public nutrition program: prototype development and beta-testing. JMIR Mhealth Uhealth 2017 Aug 2;5(8):e102 [FREE Full text] [doi: 10.2196/mhealth.7477] [Medline: 28768611]
34. Stephan LS, Almeida ED, Guimaraes RB, Ley AG, Mathias RG, Assis MV, et al. Processes and recommendations for creating mHealth apps for low-income populations. JMIR Mhealth Uhealth 2017 Apr 3;5(4):e41 [FREE Full text] [doi: 10.2196/mhealth.6510] [Medline: 28373155]
35. Swann CA. WIC eligibility and participation: the roles of changing policies, economic conditions, and demographics. B E J Econom Anal Policy 2010;10(1):21. [doi: 10.2202/1935-1682.2352]

Abbreviations

EBT: electronic benefits transfer
SNAP: Supplemental Nutrition Assistance Program
WIC: The Special Supplemental Nutrition Program for Women, Infants, and Children

Edited by G Eysenbach; submitted 28.05.20; peer-reviewed by A Panzera, M Zimmer; comments to author 04.07.20; revised version received 11.08.20; accepted 21.08.20; published 14.10.20.

Please cite as:
Zhang Q, Zhang J, Park K, Tang C
Association Between Usage of an App to Redeem Prescribed Food Benefits and Redemption Behaviors Among the Special Supplemental Nutrition Program for Women, Infants, and Children Participants: Cross-Sectional Study
JMIR Mhealth Uhealth 2020;8(10):e20720
URL: http://mhealth.jmir.org/2020/10/e20720/
doi:10.2196/20720
PMID:33052133

©Qi Zhang, Junzhou Zhang, Kayoung Park, Chuanyi Tang. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 14.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

eHealth Literacy of German Physicians in the Pre–COVID-19 Era: Questionnaire Study

Johanna Kirchberg[1], MD; Johannes Fritzmann[1], MD; Jürgen Weitz[1], MD; Ulrich Bork[1], MD

Department of Visceral, Thoracic, and Vascular Surgery, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

Corresponding Author:
Ulrich Bork, MD
Department of Visceral, Thoracic, and Vascular Surgery
University Hospital Carl Gustav Carus
Technische Universität Dresden
Fetscherstr 74
Dresden, 01307
Germany
Phone: 49 3514585400
Email: ulrich.bork@ukdd.de

Abstract

Background: Digitalization is a disruptive technology that changes the way we deliver diagnostic procedures and treatments in medicine. Different stakeholders have varying interests in and expectations of the digitalization of modern medicine. Many recent digital advances in the medical field, such as the implementation of electronic health records, telemedical services, and mobile health apps, are increasingly used by medical professionals and patients. During the current pandemic outbreak of a novel coronavirus-caused respiratory disease (COVID-19), many modern information and communication technologies (ICT) have been used to overcome the physical barriers and limitations caused by government-issued curfews and workforce shortages. Therefore, the COVID-19 pandemic has led to a surge in the usage of modern ICT in medicine. At the same time, the eHealth literacy of physicians working with these technologies has probably not improved since our study.

Objective: This paper describes a representative cohort of German physicians before the COVID-19 pandemic and their eHealth literacy and attitude towards modern ICT.

Methods: A structured, self-developed questionnaire about user behavior and attitudes towards eHealth applications was administered to a representative cohort of 93 German physicians.

Results: Of the 93 German physicians who participated in the study, 97% (90/93) use a mobile phone. Medical apps are used by 42% (39/93). Half of the surveyed physicians (47/93, 50%) use their private mobile phones for official purposes on a daily basis. Telemedicine is part of the daily routine for more than one-third (31/93, 33%) of all participants. More than 80% (76/93, 82%) of the trial participants state that their knowledge regarding the legal aspects and data safety of medical apps and cloud computing is insufficient.

Conclusions: Modern ICT is frequently used and mostly welcomed by German physicians. However, there is a tremendous lack of eHealth literacy and knowledge about the safe and secure implementation of these technologies in routine clinical practice.

(JMIR Mhealth Uhealth 2020;8(10):e20099) doi:10.2196/20099

KEYWORDS

eHealth; electronic health; mobile health; health apps; mobile health apps; eHealth literacy

Introduction

In the health care sector, there are 2 sides to digitalization: on the one side, it offers opportunities for significant improvements; on the other side, it comes along with new and additional challenges for today's health care system and those who work within this system [1].

Economic constraints cause health care professionals to offer their medical services at a low cost [2]. At the same time, medical treatment is increasingly personalized, individualized, and data-based [3-5]. There is also a shortage of qualified personnel in the medical field, whereas, by contrast, demographic development has led to an aging, multi-morbid society with increased medical and nursing needs and efforts. To make matters worse, the health care system is still divided

XSL•FO
RenderX

into inpatient and outpatient sectors that are not adequately interconnected. The tremendous amount of medical data that is generated every day is still stored in different data silos in incompatible systems [6]. A major objective of digitalization is, therefore, to economize scarce capabilities in times of increasing workload and skill shortage in the health care business [7]. At the same time, a skill shortage of qualified and trained information and communication technologies (ICT) experts exist in the health care field [8].

Technologies with the potential to accomplish this are eHealth applications [9-11], artificial intelligence (AI) [12,13], and cloud computing [14-17]. In reality, this aim is still delayed by frequent technical problems and interface difficulties between different hardware and software systems [9]. Unfortunately, this leads to considerable frustration among health care workers [18-20]. The future aim of reducing workload by digitalization has not yet been achieved in the present intermediate stage of digitalization in the German health care system. However, several funded programs, such as the Medical Informatics Initiative, exist to overcome these problems [21].

The aim of this study is to analyze the usage of modern information and communication technologies in everyday medical practice in the pre–COVID-19 era. Furthermore, this study examines which of the possible applications are considered useful and which are considered less meaningful by physicians, and what knowledge exists regarding the technical and legal frameworks of these technologies. To our knowledge, this is the first study on this topic among German physicians. While data already exist about the specific situation of digital technology use among German physicians, specific questions of usage behavior and eHealth literacy have not yet been addressed [22]. Other international trials have surveyed the benefits and challenges arising with the usage of smartphone applications, such as medical apps or consumer messaging apps in a medical context [23-26].

All data in this trial were obtained in the pre-COVID-19 era. In our opinion, the sudden need to physically isolate doctors (for example, in tumor boards, laboratory meetings, and conferences) has massively accelerated the daily spread of digital devices and applications within the medical system [27-30]. However, our study shows that German physicians have limited or even insufficient knowledge about the safe and efficient use of eHealth technologies. Improvements in eHealth literacy are therefore urgently needed.

Methods

A structured questionnaire with 6 categories was developed to evaluate the participants' patterns of use and level of knowledge of eHealth applications (Multimedia Appendix 1). Furthermore, physicians' attitudes towards the potential benefits and disadvantages of eHealth applications were analyzed. The demographic and academic characteristics of the participants were also collected anonymously. The questionnaire was developed with the help of a statistician and an expert in data management, and was tested and validated with a group of 10 persons working at the surgical trial unit in the Department of Surgery, University Hospital Dresden.

The 6 categories of the questionnaire were as follows: (1) personal characteristics (age/sex/work experience/academic degree/specialization); (2) usage profile of eHealth applications; (3) level of knowledge of eHealth applications; (4) critical evaluation of medical versus nonmedical apps; (5) critical evaluation of medical apps for patient use; and (6) critical evaluation of medical apps for physician use.

Physicians were asked to anonymously answer the questionnaire on 4 different occasions. Consent for using the anonymous data was obtained when the participants answered the questionnaire. Formal approval to conduct the surveys and to use the questionnaire was obtained from the corresponding heads of departments or conference organizers before the surveys were conducted.

The first survey was conducted during grand rounds of the Department of Visceral, Thoracic, and Vascular Surgery at the University Hospital Carl Gustav Carus, Technical University Dresden, in Germany, on August 27, 2018; of the 35 questionnaires distributed, 24 were completed. The second survey was conducted at an interdisciplinary course on pancreatic surgery for physicians from the ambulant and clinical sectors at University Hospital Carl Gustav Carus, Technical University Dresden, on August 29, 2018; of the 74 questionnaires distributed, 27 were completed. The third survey was conducted at a gastric cancer course from the Dresden School of Surgical Oncology (DSSO) on August 30, 2018; of the 8 questionnaires distributed, all 8 were completed. The fourth survey was conducted during an interdisciplinary meeting of the physicians' association Siegen/Olpe in Kreuztal Krombach, Germany, on September 5, 2018; of the 80 questionnaires distributed, 34 were completed. The total response rate of completed questionnaires was 42% (93/197).

All data were pooled, transferred, and analyzed using Excel (version 14.0; Microsoft). Figures were created with GraphPad Prism (version 6.07; GraphPad Software Inc). Due to the limited number of questionnaires, data analysis was descriptive only.

Results

Participant Characteristics

Table 1 shows the participants' demographic data. Most participants (37/93, 40%) were 30-45 years old and were men (45/93, 49 %); 29% (27/93) were women, and 22.6% (21/93) of participants did not answer the question regarding their sex. The length of working experience was 5-15 years for 31% (29/93) of participants, more than 25 years for 30% (28/93) of participants, and less than 5 years for 25% (23/93) of participants. The highest academic degrees held by the participants were state examination for 42% (39/93), a doctorate of medicine (MD) for 43% (40/93), and board certification for 63% (59/93).

XSL•FO
RenderX

Table 1. Demographic, professional, and academic characteristics of the study participants (n=93).

Participant characteristics	Values (%)
Age (years)	
<30	17.2
30-45	39.8
45-60	32.3
>60	10.8
Gender	
Male	48.4
Female	29.0
Missing	22.6
Working experience (years)	
<5	24.7
5-15	31.2
15-25	14.0
>25	30.1
Highest academic degree	
State examination	41.9
Medical doctor (MD)	43.0
Habilitation (lecturer)	4.3
Professorship	4.3
Missing	6.5
Board-certified specialist	
Yes	63.4
No	33.3
Missing	3.2

Usage Profile of eHealth Applications

The participants used eHealth applications in daily life via their mobile phones. Among the 93 participants, 90 (96.8%) had their own mobile phone, which was a smartphone in more than 90% of cases. The system software was Android for 48% (45/93) of participants and Apple/iOs for 44% (41/93), while 3% (3/93) used other software. Of the 93 physicians, 67 (72%) carried their private mobile phone with them during work; only 30% (28/93) had an official mobile phone provided by their employer (Table 2).

Table 2. Profile of professional mobile phone usage among study participants (n=93).

Questions regarding mobile phone usage	Responses (%)		
	Yes	No	Missing data
Do you use and own a mobile phone?	96.8	3.2	0.0
Do you use your own mobile phone during work?	72.0	28.0	0.0
Do you use a separate professional mobile phone?	30.1	65.6	4.3
Is your mobile phone a smartphone?	92.5	6.5	1.1

In general, the participants' mobile phones were used (professionally and privately) for phone calls (89/93, 96%), messaging (eg, WhatsApp; 84/93, 90%), surfing the internet (83/93, 89%), navigation (77/93, 83%), online banking and financial issues (33/93, 36%), social media (29/93, 31%), and gaming (16/93, 17%). These results are displayed in Table 3.

XSL•FO
RenderX

Table 3. Purpose of mobile phone use among study participants (n=93).

Purpose of mobile phone use	Values (%)
As a telephone and for SMS text-messaging services	95.7
Messaging apps	90.3
Social media	31.2
Surfing the internet	89.2
Route planner/navigation	82.8
Online banking/finance	35.5
Medical apps	41.9
Gaming	17.2
Other	7.5

Medical apps were used by 42% (39/93) of participants. Of the 93 surveyed physicians, 50% (47/93) used their private mobile phone for official purposes on a daily basis, 24% (22/93) did so once per week, 10% (9/93) did so once per month, and 15% (14/93) never did so. Telemedicine was part of the daily routine for 33% (31/93) of all participants, while 65% (60/93) did not use it. A fitness bracelet was used by 12% (11/93) of participants.

Critical Evaluation of Medical Versus Nonmedical Apps

To obtain further insights into the interviewees' attitudes, the participants were asked to rank their personal levels of importance for safety, quality, fun, and design factors in medical versus nonmedical apps. For professionally used medical apps, 63% (59/93) of participants ranked data safety as most important, while 66% (61/93) ranked quality of content as most important. The fun factor was of minor importance (ie, unimportant) to 67% (62/93) of the physicians, as was the fact that family or friends used the same app (47/93, 51%). Saving time was seen as important for 66% (61/93) of participants (Table 4). For nonmedical privately used apps, 65% (60/93) of the participating physicians ranked data safety, and 62% (58/93) ranked the quality of content, as most important. Additionally, saving time (56/93, 60%) and design factors (49/93, 53%) were very important in the participants' perceptions. The fun factor was the least important for 42% (39/93) of participants, while 41% (38/93) stated that family or friends using the same app was important (Table 4).

Table 4. Evaluations of the importance of professionally used medical apps and privately used nonmedical apps among study participants (n=93).

Factor	Evaluation of importance (%)				
	Not important	Important	Very important	Not applicable	Missing data
Professionally used medical apps					
Design and operation	11.8	26.9	44.1	12.9	4.3
Time saving	1.1	17.2	65.6	12.9	3.2
Data safety	2.2	15.1	63.4	15.1	4.3
Friends/family use the same app	50.5	11.8	9.7	21.5	6.5
Image of provider	43.0	25.8	11.8	15.1	4.3
Quality and topicality of contents	3.2	12.9	65.6	14.0	4.3
Fun factor	66.7	6.5	3.2	18.3	5.4
Privately used nonmedical apps					
Design and operation	5.4	35.5	52.7	5.4	1.1
Time saving	0.0	31.2	60.2	5.4	3.2
Data safety	3.2	24.7	64.5	5.4	2.2
Friends/family use the same app	24.7	40.9	23.7	8.6	2.2
Image of provider	58.1	24.7	4.3	9.7	3.2
Quality and topicality of contents	0.0	29.0	62.4	6.5	2.2
Fun factor	32.3	41.9	12.9	8.6	4.3

XSL•FO
RenderX

Level of Knowledge of eHealth Applications and Data Safety

Of the 93 participants, 66% (61/93) believed that less than 30% of all apps in established app stores conformed with basic standards for data safety and safe communication; 22% (20/93) thought that at least 30-60% of apps met these standards. Furthermore, 39% (36/93) appraised messaging apps (eg, WhatsApp) as appropriate for professional communication while 53% (49/93) did not, and 89% (83/93) did not know of alternative safe messaging apps (eg, Siilo, Careflow Connect, MedCrowd). The participants who rated email communication as sufficiently safe for professional communication in the health care business were 28% (26/93), while 62% (58/93) did not agree. More than 80% of the respondents had never been asked by patients about medical apps. The findings suggest that insecurity with regard to issues of data safety is common: 82% (76/93) admitted that their knowledge was insufficient regarding the legal aspects and data safety of medical apps and cloud computing in clinical life, while 65% (60/93) held this belief about their knowledge regarding technical aspects. Furthermore, 85% (79/93) of participants thought it was necessary to perform a legally bound certification of medical apps regarding data safety, and 79% (73/93) felt this was necessary with regard to the quality of content (Table 5).

Table 5. Questions on the application, state of knowledge, data safety, and legal obligations of medical eHealth apps (n=93).

Question	Responses (%)			
	No	Yes	Unknown	Missing data
Is it appropriate to use common messaging apps for professional communication (eg, WhatsApp)?	52.7	38.7	8.6	0.0
Do you know of safe messaging apps for professional communication (eg, Siilo, Careflow Connect, MedCrowd)?	89.2	10.8	0.0	0.0
Is it appropriate to use common email for professional communication in health systems?	62.4	28.0	8.6	1.1
Do you have sufficient knowledge of the technical aspects of medical apps and cloud computing to evaluate their application in clinical daily work life?	64.5	21.5	12.9	1.1
Do you have sufficient knowledge of the technical aspects of medical apps and cloud computing regarding legal aspects and data safety to evaluate their application in clinical daily work life?	81.7	4.3	14.0	0.0
Do you use telemedical applications in your clinical daily work life?	64.5	33.3	2.2	0.0
Have you been asked about medical apps by patients yet?	82.8	16.1	1.1	0.0
Do you think a legal obligation for external certification of medical apps is required ?				
Regarding safe communication and data storage?	4.3	84.9	9.7	1.1
Regarding medical effectiveness and quality of contents?	5.4	78.5	14.0	2.2

Critical Evaluation of Medical Apps for Patient Use

The use of medication apps, coaching apps for medical issues, online scheduling apps, and emergency apps for patients was ranked as a possible idea or good idea by 86% (80/93), 82% (76/93), 81% (75/93), and 76% (71/93) of participants, respectively. Video consultation apps and follow-up apps were considered a bad idea by 39% (36/93) and 20% (19/93) of all participants, respectively, whereas 55% (51/93) and 73% (68/93) considered both to be possible and good ideas, respectively (Table 6).

XSL•FO
RenderX

Table 6. Physician evaluation of medical apps for patients and physicians among study participants (n=93).

Medical apps	Physician evaluation (%)			
	Bad idea	Feasible	Good idea	Missing data
Medical apps for patients				
Emergency app	15.1	36.6	39.8	8.6
Online appointment app	12.9	50.5	30.1	6.5
Coaching app on illnesses	9.7	37.6	44.1	8.6
Video consultation app/digital health assistance app	38.7	12.9	41.9	6.5
Follow-up app (eg, postoperative, malignancy aftercare)	20.4	40.9	32.3	6.5
Medication app	8.6	58.1	28.0	5.4
Medical apps for physicians				
Diagnostic/differential diagnosis app	3.2	53.8	39.8	3.2
Guideline app	1.1	76.3	19.4	3.2
Drug app	1.1	80.6	15.1	3.2
Documentation/ward round assistance app	9.7	36.6	50.5	3.2
Digital patient record	18.3	38.7	40.9	2.2

Critical Evaluation of Medical Apps for Physician Use

The attitude of the participants towards medical apps for physicians differed. Large medical apps that might facilitate the physicians' daily life ranked positively, such as guideline apps, medication apps, diagnostic help apps, and documentation help apps, which rated as possible ideas or good ideas by 96% (89/93), 96% (89/93), 94% (87/93), and 87% (81/93) of the physicians, respectively. Digital patient records were viewed more critically: 80% (74/93) of the physicians believed they were a possible idea (36/93, 39%) or a good idea (38/93,41%), whereas 18% (16/93) thought they were a bad idea (Table 6).

Discussion

Principal Findings

The market for health apps is confusing and difficult for individuals to grasp. There is a risk of misuse of health data related to patients. Consequently, applications in the eHealth sector must be subject to the same regulations as those otherwise provided in the health sector. Several initiatives have published evaluation guidelines for the systematic assessment of the quality of apps. Thus far, however, no uniform seal or proof of safety and quality exists [31-34]. In fact, only 30% of all commercially distributed health apps have privacy policies. Of these privacy policies, two-thirds are unrelated to the health app itself but only address commercial rights, distribution rights, or the rights of third parties. Health apps still frequently share data with third parties without the user's knowledge, often without encryption [35].

This situation is also related to the third trend in the telehealth sector: medical care for chronically ill patients is increasingly shifting from hospitals to the outpatient sector. With the help of portable diagnostic technologies coupled with smartphones and a telemedical connection of the patient to hospitals, diseases can be treated at home. Offering health services in nursing homes or at home via smartphones would follow a trend that has existed for years in other areas (eg, online shopping, online banking).

The trial population of this study was a representative cohort of German physicians in terms of age and academic degrees. There was an underrepresentation of 29% of female physicians in this study (46.8% of physicians in Germany are female), and the field "sex" was left unanswered by 22.6 % of participants due to an unfavorable layout in the questionnaire [36,37]. Certain facts reduce the general transferability of the study. The cohort mainly consisted of surgeons in the hospital sector because 3 of the 4 time points of inquiry were surgical meetings. Nonsurgical physicians of the ambulant, nonacademic sector are underrepresented. Physicians with a positive attitude towards digitalization in public health might be overrepresented. The questionnaire was self-designed, and to our knowledge, no standardized questionnaire on the topic of digitalization and eHealth exists. The sample size of 93 physicians is too small to guarantee generalizability. However, the answer rate of 47% was in the expected range for this type of survey.

All data in this trial were obtained in the pre–COVID-19 era. In our opinion, the sudden need for physical distancing (for example, in tumor boards, laboratory meetings, and conferences) has massively accelerated the daily spread of digitalization within the medical system [30]. This phenomenon may have a lasting positive impact on doctors' and patients' attitudes towards eHealth applications and should be the subject of further study. In addition, further studies with larger sample sizes are needed.

The participants had contact with eHealth applications, mainly via their private mobile phones. In contrast to the general population in Germany, they favored Apple software (41/93, 44%) at a relevantly higher percentage. The German average of Apple users is only 20% [38,39]. The participants who used telemedicine in their professional lives were 33% (31/93), while 65% (60/93) did not use it. Most physicians (77/93, 83%) had

XSL•FO
RenderX

never received questions from patients about health apps as part of their professional activities, but 16% (15/93) said they had received such questions. There is an increasing desire and need for advice among patients [40]. Furthermore, there is little knowledge among medical staff within health care facilities regarding information technology (IT) security. Great harm can be caused by unsuspecting, reckless behavior. For example, most hacker attacks in the health sector can be traced to the inadvertent installation of malware, for example, by opening file attachments in emails or using external USB sticks [41].

Almost 40% (37/93) of those surveyed considered it generally acceptable to use normal messaging apps (eg, WhatsApp) for professional communication. Only slightly more than half of the respondents considered this to be unjustifiable, although, in a previous question, 80% (74/93) of the participants considered data protection and security to be important or very important. Secure messaging apps that have been specially developed for professional communication in the medical sector, such as Siilo, Careflow connect, or Medcrowd, were only known to approximately 10% (9/93) of those surveyed. At least 30% (28/93) of those surveyed considered professional communication in the health sector via normal (unencrypted) email to be sufficiently secure and reliable, while approximately 60% (56/93) did not consider this to be sufficiently safe and reliable. This means that despite the sensitive data, unsuitable and insecure communication channels such as normal messaging apps and unencrypted email communication are still regarded by many physicians as acceptable means of communication [42,43].

The self-assessment of knowledge was remarkable. With regard to apps and cloud computing, only 20% (19/93) of the surveyed participants stated that they were sufficiently familiar with the technical aspects. With regard to legal aspects and data safety, only 4% (4/93) believed that they were sufficiently informed to assess their use. In the case of health apps for patients to intervene in diagnostics and therapy (eg, medication app), almost half of the physicians surveyed believed that these should be classified as medical devices. The survey results also show that there is little specialist knowledge among the survey participants in this area, since apps for intervention in diagnostics and therapy must be approved as medical devices in Germany.

The evaluation of medical apps for patients is also interesting. It is noticeable that apps for patients were rated more critically by the physicians than potential apps for physicians. One reason for this skepticism among physicians could be the changed doctor-patient relationship or a new way of evaluating and using medical services that would result from these applications [44,45].

The results show that there is no uniform opinion among the physicians surveyed. There seems to be a non-negligible group of up to 40% (37/93) of the surveyed physicians who are critical of the new application options. On the other hand, a large group of the surveyed physicians considered many of these applications to be feasible or potentially good ideas.

The evaluation of medical apps for physicians, however, shows a slightly different picture. Only 3% (3/93) considered an app for assistance with diagnostics and differential diagnoses to be a bad idea, guidelines were considered good by almost all of the respondents, and a drug app and a documentation and visit assistance app were considered favorable by more than 90% of the respondents. Apps for help with diagnostics, differential diagnoses, guideline display, medication, and documentation ward round assistance were rated as good ideas by up to 50% of respondents, and as feasible by up to 50% [43,46,47].

It is striking that an app for electronic medical records (EMR) was considered by almost 20% to be a bad idea, by nearly 40% to be feasible, and by approximately 40% to be a good idea. Again, there is a difference of opinion within the medical profession [48]. This is also in contrast to data from other countries, which are much more in favor of EMR [1].

Similar to the negative attitude towards some patient apps, this may also be an indicator that there are reservations in the medical profession regarding a health system that is changing through eHealth applications, which will also affect the professional profile of physicians and in which the digital patient file is one key player.

Conclusion

Modern information and communication technologies, such as smartphones, are already often used. A large proportion of physicians are open to such new technology-based applications and are in favor of their introduction and use in practice.

The possibility of using these technologies to avoid errors in treatment, to reduce the administrative burden, and to improve supply in the area is apparent. However, a non-negligible proportion of physicians are skeptical about these applications and critically assess or reject their use, especially if the application affects the doctor-patient relationship.

There is also a considerable lack of knowledge about the application of these technologies, which leads to incorrect assessments on the part of medical professionals regarding data security and data protection. It can be assumed that modern information and communication technologies will be used in the German health care system in the near future. Due to the increasing importance of these technologies, at least basic knowledge of eHealth applications in the health sector should be provided by further training measures and by including this content in medical curricula. Digitalization in health care is advancing. The speed of its development and its acceptance by those involved will largely depend on the extent to which those affected participate in this process. Active participation is, therefore, essential for the medical profession and is the only way to advance development in patient-centered, holistic, personalized medicine. Care must also be taken to ensure that the special doctor-patient relationship is not lost.

The rapidly accelerated spread of eHealth applications within the medical system by COVID-19 protection measures since March 2020 may induce a massive and long-lasting rethinking of digitalization in the health care sector by medical professionals and trigger a boom in eHealth applications. This current development should be further studied.

XSL•FO
RenderX

Acknowledgments

The article publishing fee of this paper was funded with the help of open-access funding by the publication fund of the Technische Universität Dresden, Germany.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionaire - Supplementary material.
[PDF File (Adobe PDF File), 211 KB - mhealth_v8i10e20099_app1.pdf]

References

1. Kernebeck S, Busse TS, Böttcher MD, Weitz J, Ehlers J, Bork U. Impact of mobile health and medical applications on clinical practice in gastroenterology. WJG 2020 Aug 7;26(29):4182-4197. [doi: 10.3748/wjg.v26.i29.4182]
2. Sommers BD, Gawande AA, Baicker K. Health Insurance Coverage and Health - What the Recent Evidence Tells Us. N Engl J Med 2017 Aug 10;377(6):586-593. [doi: 10.1056/NEJMsb1706645] [Medline: 28636831]
3. Haendel MA, Chute CG, Robinson PN. Classification, Ontology, and Precision Medicine. N Engl J Med 2018 Dec 11;379(15):1452-1462. [doi: 10.1056/NEJMra1615014] [Medline: 30304648]
4. All of Us Research Program Investigators, Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, et al. The. N Engl J Med 2019 Aug 15;381(7):668-676. [doi: 10.1056/NEJMsr1809937] [Medline: 31412182]
5. Hunter DJ, Longo DL. The Precision of Evidence Needed to Practice. N Engl J Med 2019 Jun 20;380(25):2472-2474. [doi: 10.1056/NEJMe1906088] [Medline: 31157961]
6. Lehne M, Sass J, Essenwanger A, Schepers J, Thun S. Why digital medicine depends on interoperability. NPJ Digit Med 2019;2:79 [FREE Full text] [doi: 10.1038/s41746-019-0158-1] [Medline: 31453374]
7. Penter V, Arnold C, Friedrich S, Eichhorst S. Zukunft deutsches Krankenhaus 2020: Thesen, Analysen, Potenziale. 2. völlig überarbeitete Auflage. Kulmbach: Mediengruppe Oberfranken; 2014:978-973.
8. 2020. Medizininformatik Initiative. URL: https://www.medizininformatik-initiative.de/en [accessed 2020-10-06]
9. Dorsey ER, Topol EJ. State of Telehealth. N Engl J Med 2016 Jul 14;375(2):154-161. [doi: 10.1056/NEJMra1601705] [Medline: 27410924]
10. Kummervold PE, Chronaki CE, Lausen B, Prokosch H, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
11. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
12. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Dec;2(4):230-243 [FREE Full text] [doi: 10.1136/svn-2017-000101] [Medline: 29507784]
13. Yeung S, Downing NL, Fei-Fei L, Milstein A. Bedside Computer Vision - Moving Artificial Intelligence from Driver Assistance to Patient Safety. N Engl J Med 2018 Apr 05;378(14):1271-1273. [doi: 10.1056/NEJMp1716891] [Medline: 29617592]
14. Bork U, Weitz J, Penter V. Apps und Mobile Health: Viele Potenziale noch nicht ausgeschöpft. Ärzteblatt Internet. 2018 Jan 19. URL: https://www.aerzteblatt.de/archiv/195817/Apps-und-Mobile-Health-Viele-Potenziale-noch-nicht-ausgeschoepft [accessed 2018-03-07]
15. Sim I. Mobile Devices and Health. N Engl J Med 2019 Sep 05;381(10):956-968. [doi: 10.1056/NEJMra1806949] [Medline: 31483966]
16. Cloud SCC. Impact of Cloud Computing on Healthcare, Version 2. OMG Cloud Working Group 2017.
17. Council EFT. How The Cloud Is Transforming Healthcare Internet. Forbes. URL: https://www.forbes.com/sites/forbestechcouncil/2017/06/13/how-the-cloud-is-transforming-healthcare/ [accessed 2018-02-18]
18. De Leeuw JA, Woltjer H, Kool RB. Identification of Factors Influencing the Adoption of Health Information Technology by Nurses Who Are Digitally Lagging: In-Depth Interview Study. J Med Internet Res 2020 Aug 14;22(8):e15630 [FREE Full text] [doi: 10.2196/15630] [Medline: 32663142]
19. Sturesson L, Groth K. Effects of the Digital Transformation: Qualitative Study on the Disturbances and Limitations of Using Video Visits in Outpatient Care. J Med Internet Res 2018 Jun 27;20(6):e221 [FREE Full text] [doi: 10.2196/jmir.9866] [Medline: 29950290]
20. Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, Hanson P, et al. Advantages and limitations of virtual online consultations in a NHS acute trust: the VOCAL mixed-methods study Internet. Southampton (UK): NIHR Journals Library; 2018.

XSL•FO
RenderX

21. Adler G, Seufferlein T, Bischoff SC, Brambs H, Feuerbach S, Grabenbauer G, et al. [S3-Guidelines. Z Gastroenterol 2007 Jun;45(6):487-523. [doi: 10.1055/s-2007-963224] [Medline: 17607616]
22. Wie Ärzte über die Digitalisierung des Gesundheitswesens denken Internet. DAK Digitalisierungsreport. URL: https://www.dak.de/dak/gesundheit/e-health-loesungen-so-denken-die-aerzte-2220224.html [accessed 2020-10-06]
23. Abolfotouh MA, BaniMustafa A, Salam M, Al-Assiri M, Aldebasi B, Bushnak I. Use of smartphone and perception towards the usefulness and practicality of its medical applications among healthcare workers in Saudi Arabia. BMC Health Serv Res 2019 Nov 12;19(1):826 [FREE Full text] [doi: 10.1186/s12913-019-4523-1] [Medline: 31718639]
24. 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]
25. Dimond R, Bullock A, Lovatt J, Stacey M. Mobile learning devices in the workplace: 'as much a part of the junior doctors' kit as a stethoscope'? BMC Med Educ 2016 Aug 17;16(1):207 [FREE Full text] [doi: 10.1186/s12909-016-0732-z] [Medline: 27530343]
26. INSTANT MESSAGING IN THE NHS. CommonTime Healthcare Report. URL: http://allcatsrgrey.org.uk/wp/download/informatics/Report-Instant-Messaging-in-the-NHS.pdf [accessed 2020-10-06]
27. Hollander JE, Carr BG. Virtually Perfect? Telemedicine for Covid-19. N Engl J Med 2020 Mar 11. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
28. Dharmarajan H, Anderson JL, Kim S, Sridharan S, Duvvuri U, Ferris RL, et al. Transition to a virtual multidisciplinary tumor board during the COVID-19 pandemic: University of Pittsburgh experience. Head Neck 2020 Apr 24. [doi: 10.1002/hed.26195] [Medline: 32329958]
29. Serper M, Cubell AW, Deleener ME, Casher TK, Rosenberg DJ, Whitebloom D, et al. Telemedicine in Liver Disease and Beyond: Can the COVID-19 Crisis Lead to Action? Hepatology 2020 Apr 10. [doi: 10.1002/hep.31276] [Medline: 32275784]
30. Ting DSW, Carin L, Dzau V, Wong TY. Digital technology and COVID-19. Nat Med 2020 Apr;26(4):459-461 [FREE Full text] [doi: 10.1038/s41591-020-0824-5] [Medline: 32284618]
31. CHARISMA - Chancen und Risiken von Gesundheits-Apps Internet. Albrecht U-V. 2016. URL: http://www.charismha.de/ [accessed 2017-03-03]
32. Albrecht U. Transparency of Health-Apps for Trust and Decision Making. J Med Internet Res 2013 Dec 30;15(12):e277. [doi: 10.2196/jmir.2981]
33. Müthing J, Jäschke T, Friedrich CM. Client-Focused Security Assessment of mHealth Apps and Recommended Practices to Prevent or Mitigate Transport Security Issues. JMIR Mhealth Uhealth 2017 Oct 18;5(10):e147 [FREE Full text] [doi: 10.2196/mhealth.7791] [Medline: 29046271]
34. Yasini M, Beranger J, Desmarais P, Perez L, Marchand G. mHealth Quality: A Process to Seal the Qualified Mobile Health Apps. Stud Health Technol Inform 2016;228:205-209. [Medline: 27577372]
35. Kao C, Liebovitz DM. Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
36. Bundesärztekammer. Ergebnisse der Ärztestatistik zum 31. 12. 2017. URL: https://www.bundesaerztekammer.de/ueber-uns/aerztestatistik/aerztestatistik-2017/ [accessed 2018-09-24]
37. Sueddeutsche Z. Dr. med. URL: https://www.sueddeutsche.de/bildung/medizinstudium-dr-med-zu-leichtgemacht-1.3473555-2 [accessed 2018-09-24]
38. Spehr M. Bis der Arzt kommt. FAZ 2018 2018 Sep 18.
39. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
40. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. NPJ Digit Med 2020;3:14 [FREE Full text] [doi: 10.1038/s41746-019-0212-z] [Medline: 32047860]
41. IBM X-Force Threat Intelligence Index 2017 Internet. IBM. 2017. URL: https://public.dhe.ibm.com/common/ssi/ecm/se/en/se912361usen/SE912361USEN.PDF [accessed 2018-03-07]
42. Beinke JH, Fitte C, Teuteberg F. Towards a Stakeholder-Oriented Blockchain-Based Architecture for Electronic Health Records: Design Science Research Study. J Med Internet Res 2019 Oct 07;21(10):e13585 [FREE Full text] [doi: 10.2196/13585] [Medline: 31593548]
43. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
44. 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 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
45. Magrabi F, Habli I, Sujan M, Wong D, Thimbleby H, Baker M, et al. Why is it so difficult to govern mobile apps in healthcare? BMJ Health Care Inform 2019 Nov;26(1) [FREE Full text] [doi: 10.1136/bmjhci-2019-100006] [Medline: 31744843]

XSL•FO
RenderX

46. Santana S, Lausen B, Bujnowska-Fedak M, Chronaki C, Kummervold PE, Rasmussen J, et al. Online communication between doctors and patients in Europe: status and perspectives. J Med Internet Res 2010;12(2):e20 [FREE Full text] [doi: 10.2196/jmir.1281] [Medline: 20551011]
47. Greaves F, Joshi I, Campbell M, Roberts S, Patel N, Powell J. What is an appropriate level of evidence for a digital health intervention? Lancet 2019 Dec 22;392(10165):2665-2667. [doi: 10.1016/S0140-6736(18)33129-5] [Medline: 30545779]
48. Roehrs A, da CCA, Righi RDR, de OKSF. Personal Health Records: A Systematic Literature Review. J Med Internet Res 2017 Jan 06;19(1):e13 [FREE Full text] [doi: 10.2196/jmir.5876] [Medline: 28062391]

Abbreviations

EMR: electronic medical record
ICT: information and communication technologies

Edited by G Eysenbach; submitted 11.05.20; peer-reviewed by S Mees, C Friedrich; comments to author 30.06.20; revised version received 07.09.20; accepted 20.09.20; published 16.10.20.

Please cite as:
Kirchberg J, Fritzmann J, Weitz J, Bork U
eHealth Literacy of German Physicians in the Pre–COVID-19 Era: Questionnaire Study
JMIR Mhealth Uhealth 2020;8(10):e20099
URL: http://mhealth.jmir.org/2020/10/e20099/
doi:10.2196/20099
PMID:33064102

©Johanna Kirchberg, Johannes Fritzmann, Jürgen Weitz, Ulrich Bork. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 16.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Solving Community SARS-CoV-2 Testing With Telehealth: Development and Implementation for Screening, Evaluation and Testing

Aditi U Joshi[1], MD, MSc; Resa E Lewiss[1], MD; Maria Aini[1], MD; Bracken Babula[1], MD; Patricia C Henwood[1], MD

Thomas Jefferson University Hospital, Philadelphia, PA, United States

Corresponding Author:
Aditi U Joshi, MD, MSc
Thomas Jefferson University Hospital
1020 Samson Street
239 Thompson
Philadelphia, PA, 19103
United States
Phone: 1 2159556844
Email: aditi.joshi@jefferson.edu

Abstract

Background: Telehealth has emerged as a crucial component of the SARS-CoV-2 pandemic emergency response. Simply stated, telehealth is a tool to provide health care from a distance. Jefferson Health has leveraged its acute care telehealth platform to screen, order testing, and manage patients with COVID-19–related concerns.

Objective: This study aims to describe the expansion and results of using a telehealth program to increase access to care while minimizing additional potential exposures during the early period of the COVID-19 pandemic.

Methods: Screening algorithms for patients with SARS-CoV-2–related complaints were created, and 150 new clinicians were trained within 72 hours to address increased patient demand. Simultaneously, Jefferson Health created mobile testing sites throughout eastern Pennsylvania and the southern New Jersey region. Visit volume, the number of SARS-CoV-2 tests ordered, and the number of positive tests were evaluated, and the volume was compared with preceding time periods.

Results: From March 8, 2020, to April 11, 2020, 4663 patients were screened using telehealth, representing a surge in visit volume. There were 1521 patients sent to mobile testing sites, and they received a telephone call from a centralized call center for results. Of the patients who were tested, nearly 20% (n=301) had a positive result.

Conclusions: Our model demonstrates how using telehealth for a referral to central testing sites can increase access to community-based care, decrease clinician exposure, and minimize the demand for personal protective equipment. The scaling of this innovation may allow health care systems to focus on preparing for and delivering hospital-based care needs.

(JMIR Mhealth Uhealth 2020;8(10):e20419) doi:10.2196/20419

KEYWORDS

telehealth; telemedicine; disaster planning; pandemic; COVID-19; SARS-CoV-2; emergency medicine; testing

Introduction

The SARS-CoV-2 (COVID-19) pandemic presented the need to rapidly evolve the traditional in-person care model to distance treatment. Due to infection control considerations for clinicians and patients in higher-risk health care settings as well as worldwide shortages of personal protective equipment (PPE), telehealth became a crucial component to emergency health care [1-3]. Telehealth involves a remote medical encounter between two clinicians or between a patient and a clinician, and can be used for direct patient care at home, remote consults to specialists, and chat visits asynchronously [4-9]. Due to the current pandemic, telehealth has been quickly leveraged to safely keep people at home, disseminate information, allow decisions around testing to be made, coordinate testing when appropriate, and risk stratify patients for evidence-based and resource efficient care. In the time of COVID-19, the benefit of telehealth seems obvious: a patient can be seen remotely, assessed, and evaluated for need of testing, all without a visit to the doctor's office or to the emergency department (ED). A remote visit potentially prevents an in-person patient care visit

XSL•FO
RenderX

and facilitates timely and responsive patient-centered care [10-12].

Jefferson Health, based in Philadelphia, Pennsylvania, created a system-wide program, JeffConnect, in 2015 that included an on-demand program, remote consults, scheduled visits for postoperative or primary care, and virtual rounds. When cases of SARS-CoV-2 abruptly increased around the world, Jefferson Health prepared. The health care system's organizing plan was similar to that of other hospital settings: source more PPE; create further surge capacity; procure surge ventilators, critical equipment, and medications; streamline evaluation and management protocols; and cross-train staff. Essential to this plan was the additional incorporation of the on-demand acute care telehealth program, hereby referred to as JeffConnect, staffed by emergency medicine (EM) physicians already licensed in Pennsylvania, New Jersey, and Delaware. The institution was uniquely positioned to rapidly leverage this as an access point for care and link to SARS-CoV-2 testing. Although national telehealth companies also offer expanded services, they generally do not have local community integration and are often unable to offer testing and follow-up that is geographically linked to patient location. Jefferson Health, with its JeffConnect program and expanded services within its geographic region, did not have this limitation.

The first COVID-19 cases in Pennsylvania were reported on March 6, 2020. A state of emergency was declared in New Jersey on March 9, 2020, and a public outreach campaign regarding the ability to use telehealth for screening and an emphasis on the need for people to socially distance was launched at the same time. Jefferson Health (located in southeastern Pennsylvania and southern New Jersey) saw the increase in telehealth patient calls starting simultaneously early March.

Herein we aim to outline our real-time telehealth intervention at the start of the COVID-19 case surge in the region. We delineate how to integrate telehealth and testing sites to decrease clinician and patient exposure in a health care setting, maximize efficient use of PPE, and increase safe access to COVID-19–related testing and management in the community setting.

Methods

Program Overview

With the intention of preventing unnecessary COVID-19 exposures, many primary care providers, onsite clinics, and urgent cares in the region began to direct patients to consider telehealth first for noncritical complaints. JeffConnect is staffed 24 hours a day by physicians and sees a variety of acute care complaints, with visit times averaging less than 10 minutes. Patients require an internet connection and a smartphone, tablet, or computer, or access to one to use the platform. They pay the one-time out-of-pocket fee, or a reduced copay if they are a Jefferson employee. Due to this convenience, patients were directed to JeffConnect as the first viable solution for COVID-19 complaint screening and evaluation.

At the same time, a team composed of infectious disease specialists, health system clinical leadership, and occupational medicine leadership created screening protocols to determine who should be tested from the community, which groups of employees should be tested if they first contacted JeffConnect, and who could stay at home with guidance on self-isolation (due to limited testing capacity). Initial screening protocol developed during implementation considered limited testing capacity and prioritized COVID-19 infections in patients who presented with fever or symptoms of acute respiratory illness, were at the highest risk of exposure, were at a higher risk of developing severe disease, or were at a high risk of exposing others. Testing was recommended for patients who were symptomatic with one of the following: close contact with confirmed case, international travel within 14 days, at risk for severe disease including chronic medical conditions, residents of congregate settings, health care workers, patients who were hospitalized, and those presenting with acute respiratory illness without apparent cause. During the study period, patients who were asymptomatic were not recommended for testing.

Once mobile testing sites were operationalized and linked to JeffConnect on March 13, 2020, a SARS-CoV-2 lab order was placed in the electronic health record order entry system. After the order was placed by a clinician, the JeffConnect visit ended. Patients were instructed to contact a centralized call center, named Seamless Access. Seamless Access was a hub for patients to coordinate their on-site testing. The call center group completed the patient registration, confirmed that an order was placed, provided directions and times to present to a testing site, and provided patient education on what to expect upon arrival and in the follow-up.

All test results from the electronic health record were routed to a dedicated results follow-up team and separated into positive and negative result pools. Data was extracted from the electronic medical record and stored on secure systems consistent with medical center policies. Patient data was deidentified for aggregate analysis. Every patient tested was contacted via telephone and received counseling on results, a re-evaluation of symptoms, and advice on when to return to work or seek additional care. For positive results, patients were contacted by a designated primary care physician or nurse practitioner. For negative results, patients were contacted by primary care nurse care coordinators. Home isolation for patients who were symptomatic was emphasized regardless of the result, given the possibility for false-negative results. Patients with access to the online patient portal were able to view their results immediately and received additional written instructions based on a positive or negative result. Each week, the number of patients, number of SARS-CoV-2 tests performed, and positive results were tracked for quality improvement processes and to inform considerations on testing capacity.

Staffing Telehealth

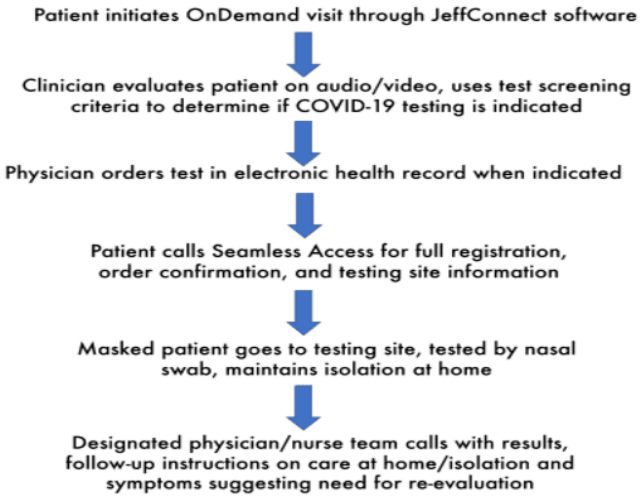
Considering more limited access to ambulatory visits, patients rapidly sought care via JeffConnect for COVID-19–related complaints as well as for other acute complaints that they were not able to see their primary care physician about. The goal was for patients who were generally well to be evaluated and

XSL•FO
RenderX

determine if they required testing or were safe to stay home (Figure 1).

Due to the opening of the testing sites and the aforementioned goals, the prior staffing model of one EM physician managing JeffConnect visits, even with the expectation of seeing four patients an hour, was not sufficient to keep up with the abrupt increase in patient volume. JeffConnect's leadership had to scale their process quickly to remotely evaluate a large volume of patients. To allow patients to be seen in a timely manner, during the weekend of March 13-15, 2020, 150 new physicians from internal medicine (IM), family medicine (FM), anesthesia, and EM advanced practice providers (APPs) were trained remotely using Zoom (Zoom Video Communications, Inc) and myJeffHub, an internal video education platform accessible to all employees. New clinician accounts were created by the JeffConnect Medical Director and all clinicians were given the training videos, tip sheets, screening protocols, and Jefferson policies.

Figure 1. JeffConnect screening and testing process.



The IM and FM physicians added additional coverage from the hours of 7 AM to 10 PM starting on March 15, 2020. IM and FM physicians, due to licensing constraints, only covered patients in Pennsylvania, while the EM staff would cover Pennsylvania, New Jersey, and Delaware. Other clinicians (anesthesia and APPs) were kept in reserve as needed for surge capacity. For EM faculty, contingency planning included ensuring those in high-risk groups such as those 60 years and older or who were immunocompromised were also trained for JeffConnect to convert in-person shifts to telehealth. The IM and FM practices converted a number of their patient visits to telehealth or rescheduled routine visits for later. At that time, there was no emergency state reciprocity for licensure, so most of the additional physicians covered patients from Pennsylvania only. The newly trained clinicians and the existing telehealth EM physician were able to meet the increased patient demand.

Over the next week, the ED redistributed staff physicians and APPs to account for the fluctuating patient volumes. Clinicians deemed high risk for COVID-19 complications were restaffed to telehealth. JeffConnect leaders added a swing shift to increase the IM and FM staffing from 10 AM to 7 PM. New Jersey amended its licensing process, allowing for emergency temporary licensure. This allowed more physicians to provide telehealth to New Jersey patient callers.

During the testing period, we evaluated the visit volume, number of SARS-CoV-2 tests ordered, and number of positive tests. JeffConnect visit volumes from the same period in 2019 and the 5-week time interval preceding this 5-week study period were evaluated for comparison. Descriptive statistics were used in the analysis.

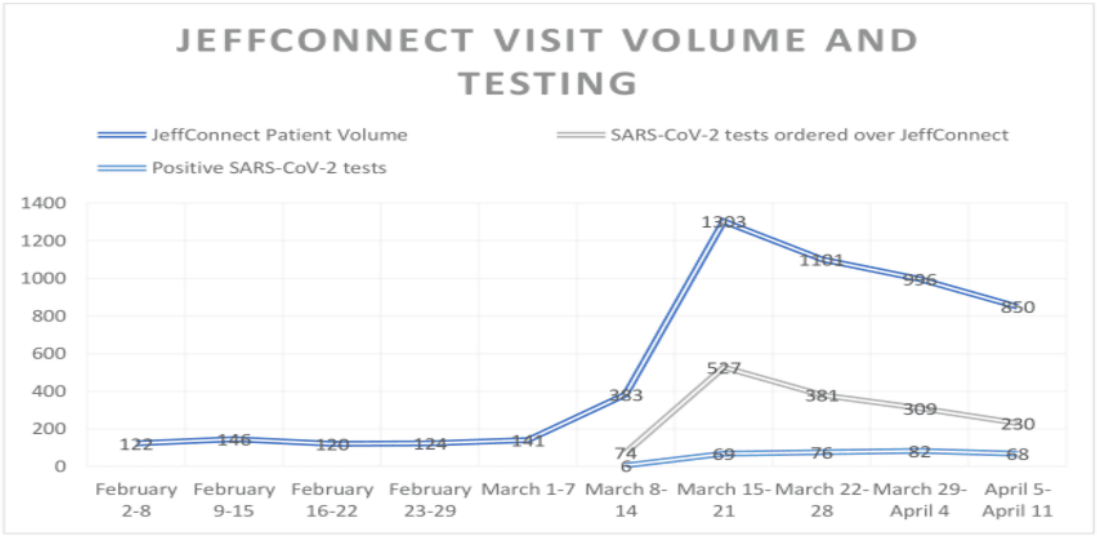
Results

In the telehealth screening and testing study period (March 8, 2020, to April 11, 2020), 4663 patients were evaluated and 1521 (33%) were recommended for SARS-CoV-2 testing. Peak volumes for visits and testing occurred the week of March 15-21 with 1303 visits and 40% (n=527) of those patients sent for testing (Figure 2).

There was an increase in telehealth patient volume from 321 visits from March 8 to April 11 in 2019 to 4663 visits during the same dates in 2020. There were 653 total visits in the 5 weeks preceding the study period (February 2 to March 7, 2020), compared to 4663 during the study period. This represents a greater than seven-fold increase in visit volume. Of the 1521 patients that were tested, nearly 20% (n=301) had a positive result.

XSL•FO
RenderX

Figure 2. JeffConnect visit volumes and testing.



Discussion

Principal Findings

The abrupt rise in telehealth visit volume and COVID-19–related testing need, as well as the approximately 20% (301/1521) positive rate during our study period mirrors the numbers in Pennsylvania, with a 18% positive rate in the state during the same time period (129,207 SARS-CoV-2 tests with 23,340 positive cases in Pennsylvania) [13]. This coincides with the initial surge of COVID-19 cases in the northeast of the United States during this time period [14]. This suggests our screening algorithms and indications for testing, considering constrained testing resources, were similar to those being used in the region during the study period [13].

It is unclear how many in-person evaluations were prevented, and many of the new patients to JeffConnect were seen for evaluation of COVID-19 symptoms or exposures. We believe our surge in volume was also due to patients seeking care for other urgent complaints via telehealth to avoid in-person contact with the health care clinicians. The goal of quickly scaling telehealth capacity was to keep viral exposure at a minimum, and we suspect that receiving at-home screening kept patients from exposing others in person and potentially spared exposure for concerns not related to COVID-19. Considering the ongoing nature of the COVID-19 pandemic, it is not yet clear whether our region has reached its ultimate peak surge, and time will tell if this intervention significantly impacts access to care, testing, and appropriate isolation.

Process Replication

The JeffConnect telehealth COVID-19 innovation came together by the alignment of multiple factors. The region had more time than some other regions: time to write, organize, and execute health system–wide policies, and to evolve previous emergency management preparations for this specific context. As already stated, Jefferson had a pre-existing telehealth structure including and not limited to a platform, a workflow process, training guides, courses, and a quality assurance process. This allowed the efficient and rapid scale-up to meet the need for more clinicians to accommodate the increased number of patient visits. We were able to use existing staffing and electronic health record infrastructure to create multiple outpatient testing sites near these hospitals and urgent cares.

The commitment to resourcing this service to scale rapidly led to a successfully evolved program within just a few days. None of the challenges have been insurmountable. To date, the telehealth clinicians' group can see patients in an efficient and timely manner, and direct them for immediate testing at one of seven sites. For systems seeking to replicate this process, we recommend the following:

- Establish patient care and testing algorithms that clearly identify who needs to be tested, considering local testing resources
- Implement a telehealth software based on a clearly identified use case. Due to the pandemic, many companies have created quick and discounted implementations, and clear definition of need will determine which solution is the right one.
- Train clinicians in telehealth: on technology, screening processes, and converting in-person encounters to virtual care
- Identify physical sites that can be used for walk-up or drive-thru testing, which are close to existing health system infrastructure and convenient for patients to access while maintaining infection prevention practices
- Identify and train staff at the testing sites for clear and specific roles
- Identify the appropriate testing equipment and keep count of testing swabs being used relative to supply chain considerations
- Create clear workflows for testing sites to direct swabs for in-house lab versus send-out testing
- Create team and process for follow-up on test results, distancing, and disease process counseling via telephone, chat, or phone app
- Establish a quality assurance process for testing site staff, telehealth clinicians, results reviewers, and any other groups involved

XSL•FO
RenderX

- Create ongoing process improvements to address obstacles, questions, and changes
- Ensure that each step has specific team members and each team has a specific role and defined leadership. This ensures improved flow and efficiency, and allows challenges to be quickly identified and fixed.

Limitations

We acknowledge certain limitations to this innovation. One challenge appeared on Monday, March 16, 2020, when the telehealth platform software functionality slowed in processing speed. We quickly identified the cause, a large increase in the number of clinicians seeing patients. This was rectified by Teladoc by scaling the cloud-based resources. We considered this a learning lesson; we needed to better anticipate technology limitations and associated rapid growth in users.

Prior to centralization of testing and registration by Seamless Access, patients were required to call separate telephone numbers for different testing sites. Understandably, this caused confusion for both patients and for clinicians. The addition of Seamless Access and a centralized telephone number to call for registration and testing site information streamlined the workflow.

There is a cost to the service, which can also be a barrier to employing this model of care delivery, in addition to the requirement for a smartphone, tablet, or computer and internet connection among users that may limit this as an access point to care.

Conclusion

Telehealth has arrived as a crucial component of testing and treatment during the COVID-19 crisis. In a matter of days, our use of telehealth transformed the way Jefferson Health was providing patient care during a pandemic. Developing a successful and scalable telehealth solution is allowing patients to stay home and preserve hospital resources as much as possible, receive testing for COVID-19 when appropriate, and get directed to additional care when necessary. Additional research will be required to determine the effectiveness of telehealth in reducing exposures for health care workers and the community. However, in this time of uncertainty, patients are seeking evidence-based advice and the benefits of virtual care have become clear to clinicians, patients, and the health care system.

Acknowledgments

Publication made possible in part by support from the Thomas Jefferson University Open Access Fund.

Conflicts of Interest

None declared.

References

1. Ranney ML, Griffeth V, Jha AK. Critical supply shortages — the need for ventilators and personal protective equipment during the COVID-19 pandemic. N Engl J Med 2020 Apr 30;382(18):e41. [doi: 10.1056/nejmp2006141]
2. Rational use of personal protective equipment (PPE) for coronavirus disease (COVID-19) : interim guidance. World Health Organization. 2020 Mar 19. URL: https://apps.who.int/iris/handle/10665/331498
3. Livingston E, Desai A, Berkwits M. Sourcing personal protective equipment during the COVID-19 pandemic. JAMA 2020 May 19;323(19):1912-1914. [doi: 10.1001/jama.2020.5317] [Medline: 32221579]
4. Lew SQ, Sikka N. Operationalizing telehealth for home dialysis patients in the United States. Am J Kidney Dis 2019 Jul;74(1):95-100. [doi: 10.1053/j.ajkd.2019.01.023] [Medline: 30898363]
5. du Toit M, Malau-Aduli B, Vangaveti V, Sabesan S, Ray RA. Use of telehealth in the management of non-critical emergencies in rural or remote emergency departments: A systematic review. J Telemed Telecare 2017 Oct 05;25(1):3-16. [doi: 10.1177/1357633x17734239]
6. Mullen-Fortino M, Rising KL, Duckworth J, Gwynn V, Sites FD, Hollander JE. Presurgical assessment using telemedicine technology: impact on efficiency, effectiveness, and patient experience of care. Telemed J E Health 2019 Feb;25(2):137-142. [doi: 10.1089/tmj.2017.0133] [Medline: 30048210]
7. Izzo J, Bhat R, Blumenthal J, Hoffman D, Watson J, Booker E. Diagnostic accuracy of a rapid telemedicine encounter in the emergency department. Ann Emerg Med 2017 Oct;70(4):S127-S128. [doi: 10.1016/j.annemergmed.2017.07.394]
8. Slivinski A, Jones R, Whitehead H, Hooper V. Improving access to stroke care in the rural setting: the journey to acute stroke ready designation. J Emerg Nurs 2017 Jan;43(1):24-32. [doi: 10.1016/j.jen.2016.10.006] [Medline: 28131346]
9. Creating a framework to support measure development for telehealth. National Quality Forum. 2017 Aug. URL: https://www.qualityforum.org/publications/2017/08/creating_a_framework_to_support_measure_development_for_telehealth.aspx
10. Hong Y, Lawrence J, Williams D, Mainous III A. Population-level interest and telehealth capacity of US hospitals in response to COVID-19: cross-sectional analysis of Google search and national hospital survey data. JMIR Public Health Surveill 2020 Apr 07;6(2):e18961 [FREE Full text] [doi: 10.2196/18961] [Medline: 32250963]

XSL•FO
RenderX

11. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Mar 20;26(5):309-313. [doi: 10.1177/1357633x20916567]
12. Hollander JE, Carr BG. Virtually perfect? Telemedicine for COVID-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
13. COVID-19 data for Pennsylvania. Department of Health of Pennsylvania. 2020. URL: https://www.health.pa.gov/topics/disease/coronavirus/Pages/Coronavirus.aspx [accessed 2020-07-01]
14. Northeastern United States census region. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/surveillance/nrevss/coronavirus/region.html#north [accessed 2020-09-01]

Abbreviations

APP: advanced practice provider
ED: emergency department
EM: emergency medicine
FM: family medicine
IM: internal medicine
PPE: personal protective equipment

Edited by G Eysenbach; submitted 18.05.20; peer-reviewed by P Chai, C Wu, S Omboni; comments to author 13.07.20; revised version received 01.08.20; accepted 22.09.20; published 16.10.20.

Please cite as:
Joshi AU, Lewiss RE, Aini M, Babula B, Henwood PC
Solving Community SARS-CoV-2 Testing With Telehealth: Development and Implementation for Screening, Evaluation and Testing
JMIR Mhealth Uhealth 2020;8(10):e20419
URL: http://mhealth.jmir.org/2020/10/e20419/
doi:10.2196/20419
PMID:33006942

©Aditi U Joshi, Resa E Lewiss, Maria Aini, Bracken Babula, Patricia C Henwood. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 16.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementation and Application of Telemedicine in China: Cross-Sectional Study

Fangfang Cui[1,2*], MSc; Qianqian Ma[1,2*], MSc; Xianying He[1,2], MPhil; Yunkai Zhai[1,2,3], PhD; Jie Zhao[1,2], PhD; Baozhan Chen[1,2], MSc; Dongxu Sun[1,2], MSc; Jinming Shi[1,2], MSc; Mingbo Cao[1,2], BSc; Zhenbo Wang[1,2], MSc

[1]National Engineering Laboratory for Internet Medical Systems and Applications, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

[2]National Telemedicine Center of China, Zhengzhou, Henan, China

[3]School of Management Engineering, Zhengzhou University, Zhengzhou, Henan, China

[*]these authors contributed equally

Corresponding Author:
Jie Zhao, PhD
National Engineering Laboratory for Internet Medical Systems and Applications
The First Affiliated Hospital of Zhengzhou University
1 Jianshe Road, Erqi District
Zhengzhou, Henan, 450052
China
Phone: 86 371 67966215
Fax: 86 371 67966215
Email: zhaojie@zzu.edu.cn

Abstract

Background: Telemedicine has been used widely in China and has benefited a large number of patients, but little is known about the overall development of telemedicine.

Objective: The aim of this study was to perform a national survey to identify the overall implementation and application of telemedicine in Chinese tertiary hospitals and provide a scientific basis for the successful expansion of telemedicine in the future.

Methods: The method of probability proportionate to size sampling was adopted to collect data from 161 tertiary hospitals in 29 provinces, autonomous regions, and municipalities. Charts and statistical tests were applied to compare the development of telemedicine, including management, network, data storage, software and hardware equipment, and application of telemedicine. Ordinal logistic regression was used to analyze the relationship between these factors and telemedicine service effect.

Results: Approximately 93.8% (151/161) of the tertiary hospitals carried out telemedicine services in business-to-business mode. The most widely used type of telemedicine network was the virtual private network with a usage rate of 55.3% (89/161). Only a few tertiary hospitals did not establish data security and cybersecurity measures. Of the 161 hospitals that took part in the survey, 100 (62.1%) conducted remote videoconferencing supported by hardware instead of software. The top 5 telemedicine services implemented in the hospitals were teleconsultation, remote education, telediagnosis of medical images, tele-electrocardiography, and telepathology, with coverage rates of 86.3% (139/161), 57.1% (92/161), 49.7% (80/161), 37.9% (61/161), and 33.5% (54/161), respectively. The average annual service volume of teleconsultation reached 714 cases per hospital. Teleconsultation and telediagnosis were the core charging services. Multivariate analysis indicated that the adoption of direct-to-consumer mode (P=.003), support from scientific research funds (P=.01), charging for services (P<.001), number of medical professionals (P=.04), network type (P=.02), sharing data with other hospitals (P=.04), and expertise level (P=.03) were related to the effect of teleconsultation. Direct-to-consumer mode (P=.01), research funding (P=.01), charging for services (P=.01), establishment of professional management departments (P=.04), and 15 or more instances of remote education every month (P=.01) were found to significantly influence the effect of remote education.

Conclusions: A variety of telemedicine services have been implemented in tertiary hospitals in China with a promising prospect, but the sustainability and further standardization of telemedicine in China are still far from accomplished.

(JMIR Mhealth Uhealth 2020;8(10):e18426) doi:10.2196/18426

XSL•FO
RenderX

KEYWORDS

telemedicine; Chinese hospital; implementation; application; influencing factors

Introduction

Telemedicine refers to the remote delivery of health care services with the use of modern communication technology, electronic technology, and multimedia computer technology to realize the remote collection, transmission, processing, storage, and inquiry of medical information and to further provide the examination, surveillance and diagnosis of disease, remote education, and information management [1,2]. The common types of telemedicine services include teleconsultation, telediagnosis, and remote surgery teaching [3-5]. With the rapid development of telemedicine equipment and information communication technology, telemedicine has developed rapidly and is used widely around the world as a new mode of medical service [6,7]. In particular, in China, telemedicine has been used as a crucial method by the government to address the inequality of medical resources between urban and rural areas.

Medical care is an important issue related to the national economy and people's livelihood in China. However, there is a serious imbalance of distribution in medical resources, and patients in rural and remote areas do not have easy access to high-quality medical services [8]. In 2018, there were 1047 tertiary hospitals in 11 provinces in the east of China and 1216 tertiary hospitals in 21 provinces in the central and western regions [9]. High-quality medical resources are concentrated in the economically developed eastern region of China, while the central and western regions have a serious lack of health resources [10]. A health report showed there were 10.91 health care technicians per 1000 people in China’s urban areas in 2018, while there were only 4.63 health care technicians per 1000 people in rural areas. Moreover, the number of beds per 1000 people in urban medical institutions was 8.70, while it was only 4.56 in rural areas, and there were less than 2 medical beds in township health centers per 1000 rural people [9]. The imbalanced distribution of medical resources in urban and rural areas greatly increases the difficulty of medical treatment in rural areas and reduces the efficiency of treatment [11,12]. Medical experts from large hospitals in developed cities can provide telediagnosis for patients in rural and remote areas through telemedicine, which may shorten the spatial distance between doctors and patients and enable patients to be treated locally by off-site medical experts. Numerous studies have shown that telemedicine can effectively reduce expenses and save time for medical treatment [13-16]. Therefore, telemedicine has been used as an important means for medical reform, and many policies have been issued to encourage the development of telemedicine in China.

The development of telemedicine in China began in the 1980s. In 1986, the Guangzhou Ocean Shipping Company conducted a cross-sea consultation for the emergency patients on the ocean-going freighter through telegraph, which was considered to be the earliest telemedicine activity in China. In 1997, the Jinwei Medical Network in China was officially opened to provide remote, off-site, real-time, and dynamic live television consultations for patients with severe illness [17]. Subsequently, the medical institutions at all levels in China began to explore and develop telemedicine [18,19]. After years of efforts, the development of telemedicine entered its golden age. In 2017, 22 provinces in China established telemedicine platforms covering 13,000 medical institutions, providing teleconsultation, telediagnosis, and remote medical education [20].

Many scholars have carried out research on telemedicine, as telemedicine has been implemented in full swing in China. Several studies have analyzed the application of telemedicine in the treatment of different diseases such as diabetes and burns [21-23]. Further, Cai et al [19] investigated the experience of doctors and patients in the implementation of telemedicine in the Gansu Province, China. He et al [24] analyzed the patient satisfaction and compliance with telemedicine implementation in rural areas in the Guangdong Province, China. However, most of the researches focused on the application of telemedicine in certain diseases or were limited to a certain region (such as a province) or small aspects such as patient satisfaction. Research on the overall development of telemedicine in China from a national and comprehensive perspective has not been elucidated. In order to explore the characteristics of telemedicine in a multidimensional and in-depth manner, this study investigated the development of telemedicine in 161 tertiary hospitals in 29 provinces, autonomous regions, and municipalities, and conducted an overall analysis of the implementation, applications, and key factors of telemedicine in China in 2017. This study is the first to summarize the development of telemedicine in China with comprehensive information rather than a small aspect, including human resources, network construction, hardware and software facilities, and so on, thereby showing how the telemedicine was built and operated in China. Besides, this study carried out a nationwide survey, not limited to a certain area, which is unprecedented. Furthermore, the overview of implementation and application of telemedicine would be valuable for formulating policies and improving existing deficiencies.

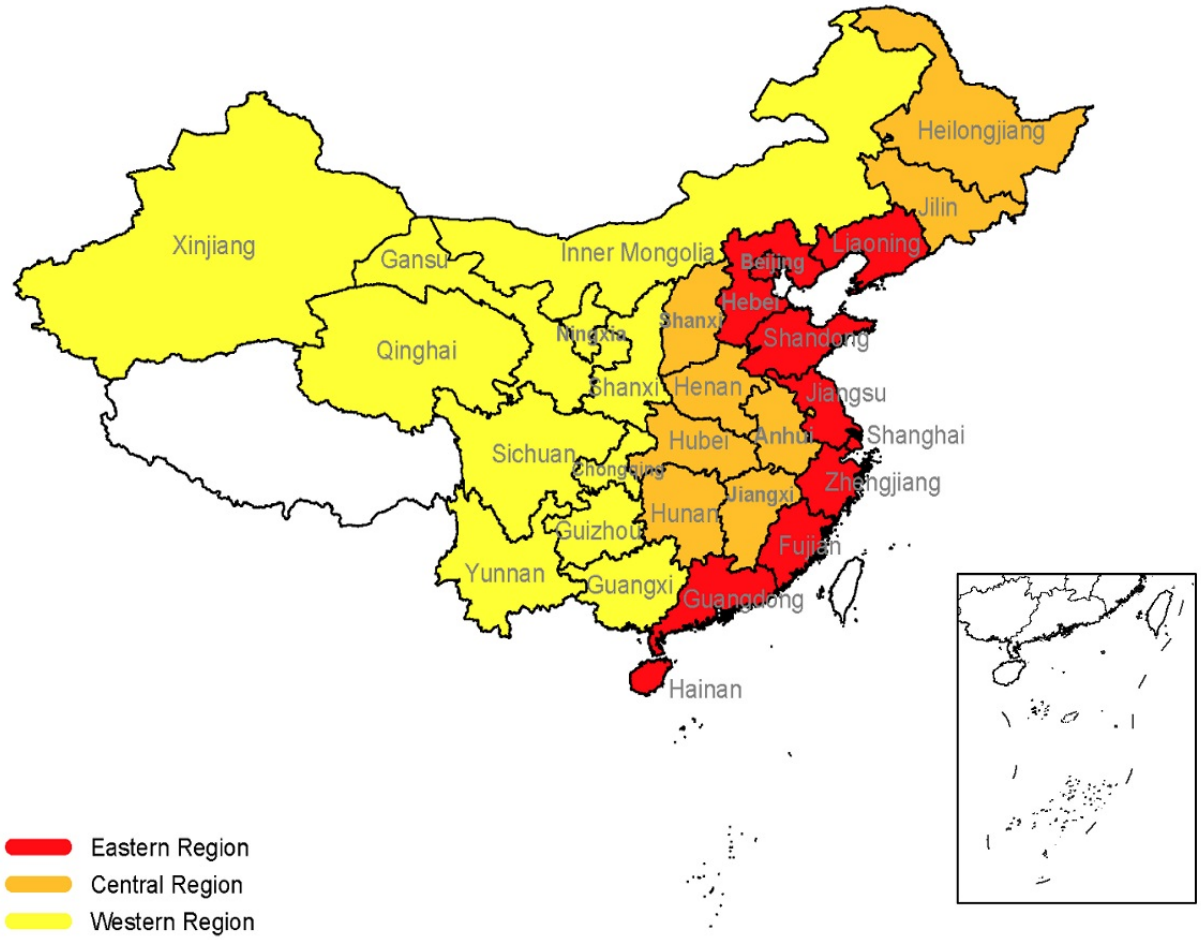
Methods

Study Design and Participants

The Chinese hospitals are divided into 3 levels: tertiary hospitals, secondary hospitals, and primary hospitals. The tertiary hospitals are the main providers of telemedicine services. Therefore, investigation of the implementation and application of telemedicine in tertiary hospitals can help in understanding the development of telemedicine in China. A web-based questionnaire survey (Multimedia Appendix 1) was conducted from August to October 2018 in Chinese tertiary hospitals that carried out telemedicine services through the Telemedicine Information Professional Committee of China (TIPC), and the survey content was with regard to telemedicine services in tertiary hospitals in 2017. TIPC is a professional organization in the telemedicine industry, gathering telemedicine practitioners and hospital administrators from different provinces in China. The survey coverage areas are shown in Figure 1.

XSL•FO
RenderX

Figure 1. Distribution of the survey coverage areas in China.



The preliminary questionnaire was developed based on a literature review [3,14,19,25] and expert consultation combined with telemedicine practice. Ten experts with more than 5 years of working experience in telemedicine were consulted on the questionnaire design. The questionnaire was modified and optimized later according to the presurvey conducted in 18 tertiary hospitals in the Henan province. The questionnaire mainly consisted of 3 parts: telemedicine implementation, telemedicine application, and key factors related to telemedicine development. The human resource allocation, management, funding, network types, data storage, software, and hardware equipment were the main components of telemedicine implementation, and telemedicine application focused on service types, service quantity, charges, and service effects. The influencing factors in the survey mainly explored the key factors affecting the development of telemedicine. We conducted a sampling survey nationwide with the probability proportionate to size sampling technique. First, we calculated the total sample size by using the following formula.

$$n = Z_{\alpha/2}^2 p(1-p)/\delta^2,\ n_c = n/(1+n/N)$$

where $Z_{\alpha/2}$ is 1.96 at the significance level of α=.05, and p is assumed to be 0.5 with the principle of maximum population variance, indicating a more conservative sample size. δ is 0.1, representing that the margin of error is within 10%. N is 2060, denoting the total number of Chinese tertiary public hospitals in 2016. n_c is the number of the samples. According to the value above, we obtained n_c to be 92, and the final sample size was 103 with the waste rate being 10%. Second, on the basis of the number of tertiary hospitals in the eastern, central, and western regions of China in 2016, we distributed the sample sizes in different regions proportionally, that is, 49 hospitals in the eastern region, 27 hospitals in the central region, and 28 hospitals in the western region. At the same time, we identified the person in charge of the survey in each region and organized the questionnaire survey by using snowball sampling.

Statistical Analysis

The total number of tertiary hospitals investigated was 185. However, the questionnaires from 24 hospitals were incomplete with high percentages of missing data, and only 161 questionnaires were valid, with an effective rate of 87.0%. The number of hospitals located in the eastern, central, and western region was 59, 54, and 48, respectively, accounting for 36.7%, 33.5%, and 29.8%, respectively, of the total number of the tertiary hospitals investigated in this study. Among the 161 hospitals, 137 hospitals provided telemedicine services and 111 hospitals obtained telemedicine services from other hospitals. The results of some items were invalid or missing. Thus, in the analysis of the corresponding content, the unqualified answers were processed as missing values. The quantitative data were described by mean values, while the qualitative data were

XSL•FO
RenderX

described by count and percentages. The column chart, bar chart, pie chart, and radar chart were adopted to analyze the implementation and application of telemedicine using the Excel software (Microsoft Corp). Furthermore, the study applied the methods of chi-square test, two-sided *t* test, variance analysis, and nonparametric test to compare the development of telemedicine in different regions. Ordinal regression was adopted to model the dependence of the telemedicine service effect on other factors in the multivariate analysis by using the SPSS 23.0 software (IBM Corp), with the significance test level of α=.05.

Results

Human Resources and Funding

The number of telemedicine staff in 75.8% (122/161) of the tertiary hospitals ranged from 1 to 6 (Figure 2). Overall, the average number of telemedicine staff in each hospital was 6.8 in China, with 7.4 in the eastern region, 6.7 in the central region, and 6.2 in the western region, respectively, but there was no statistically significant difference (*P*=.83). Telemedicine staff were mainly composed of those with master's and bachelor's degrees. Those with a bachelor's degree accounted for 49.58% (536/1081) of the telemedicine staff, followed by masters and junior college and below at 38.85% (420/1081) and 11.56% (125/1081), respectively. The majority of the telemedicine staff were in the fields of medicine, computer science and communication, and management (Table 1). There were statistically significant differences in the number of employees majoring in computer science and communication and medicine in different regions with *P*=.03 and *P*=.01, respectively. By contrast, no significant difference was found in the number of telemedicine staff with management majors in the different regions (*P*=.24).

Figure 2. Distribution of the telemedicine staff in tertiary hospitals of China.

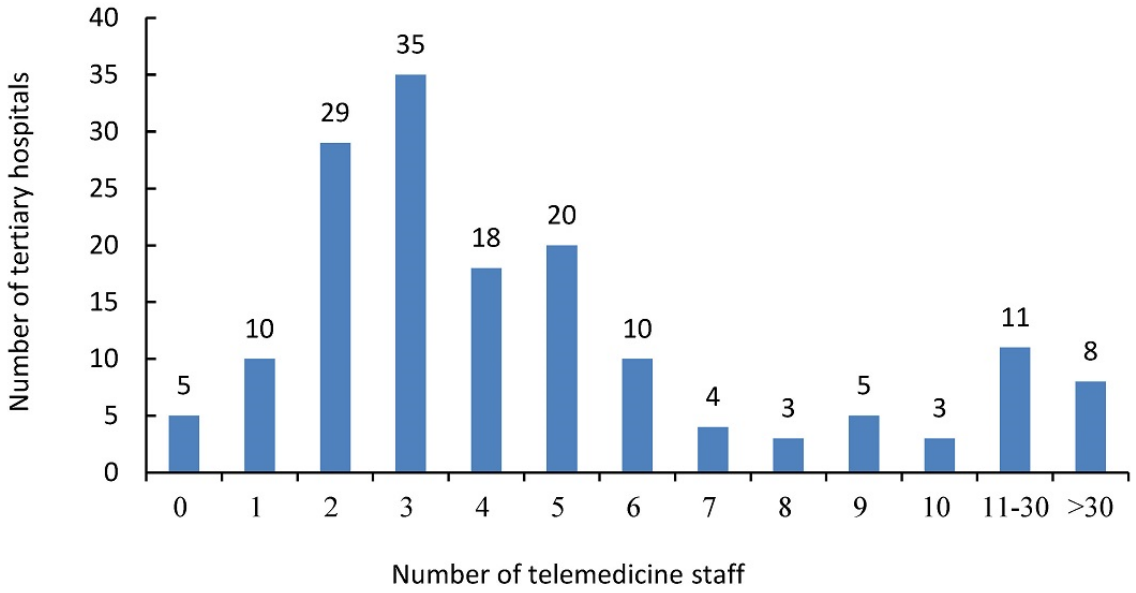


Table 1. Educational background of the telemedicine staff in the tertiary hospitals of different regions in China (person per hospital).

Educational background	Total (n)	Eastern region (n)	Central region (n)	Western region (n)
Computer science and communication	1.8	1.4	1.6	2.5
Medicine	2.7	4.1	1.9	1.7
Management	1.3	1.3	0.9	1.6

The sources of funding for telemedicine implementation mainly include government financial support, hospital self-raising, research funding, and corporate sponsorship. According to the survey results, 83.2% (134/161) of the tertiary hospitals implemented telemedicine by self-fundraising. By comparison, only 8.1% (13/161) of the hospitals received financial support through research funding. In terms of capital investment, the proportion of tertiary hospitals with an investment of less than 500,000 RMB (approximately US $71,218) was the highest at 64.6% (104/161). Approximately 19.3% (31/161) of the hospitals invested more than 1 million RMB (approximately US $142,500) into telemedicine implementation (Table 2).

XSL•FO
RenderX

Table 2. Source and investment amount of telemedicine implementation in tertiary hospitals in different regions of China (N=161).

Funds	Total, N=161, n (%)	Eastern region, n=59, n (%)	Central region, n=54, n (%)	Western region, n=48, n (%)
Sources of funds				
Government finance	47 (29.2)	12 (20)	18 (33)	17 (35)
Hospital self-raising	134 (83.2)	51 (86)	44 (82)	39 (81)
Research funding	13 (8.1)	7 (12)	3 (6)	3 (6)
Corporate sponsorship	14 (8.7)	7 (12)	5 (9)	2 (4)
Investment amount (RMB)[a]				
<100,000	59 (36.7)	21 (36)	21 (39)	17 (35)
100,000-500,000	45 (28.0)	17 (29)	17 (32)	11 (23)
500,000-1 million	26 (16.2)	10 (17)	6 (11)	10 (21)
1-5 million	21 (13.0)	6 (10)	7 (13)	8 (17)
>5 million	10 (6.2)	5 (9)	3 (6)	2 (4)

[a]1 RMB=US $0.14.

Management and Service Modes

As shown in Table 3, 59.1% (95/161) of the hospitals adopted a complete self-management operation pattern, while others delegated part or all of the telemedicine service to a third party. With respect to the management department, obviously, the proportions of hospitals with telemedicine administrative departments in the eastern (48/59, 81%) and western regions (34/48, 71%) were significantly higher than that in the central regions (28/54, 52%), with a significant regional difference (P<.001).

Telemedicine service modes include business-to-business (B2B), direct-to-consumer (DTC), and business-to-business-to-customer (B2B2C). B2B is a mode in which a medical institution provides telemedicine services to doctors in another medical institution. DTC means that medical institutions provide remote services directly to patients, while B2B2C refers to medical institutions providing telemedicine services to patients through other intermediaries such as telemedicine companies. The results suggested that 93.8% (151/161) of the tertiary hospitals carried out telemedicine services in B2B mode, which indicated the dominance of B2B mode in China. The proportion of the DTC mode (28/161, 17.4%) was closely matched to that of the B2B2C mode (32/161, 19.9%).

Table 3. Telemedicine management and service mode of tertiary hospitals in different regions of China (N=161).

Mode	Total, N=161, n (%)	Eastern region, n=59, n (%)	Central region, n=54, n (%)	Western region, n=48, n (%)
Management mode				
Self-management mode	95 (59.1)	40 (68)	29 (54)	26 (54)
Partial entrustment mode	51 (31.7)	16 (27)	22 (41)	13 (27)
Complete entrustment mode	13 (8.1)	3 (5)	3 (6)	7 (15)
Other	2 (1.2)	0 (0)	0 (0)	2 (4)
Management department				
Established	110 (68.3)	48 (81)	28 (52)	34 (71)
Being established	22 (13.7)	5 (9)	10 (19)	7 (15)
Not established	29 (18.0)	6 (10)	16 (30)	7 (15)
Service mode				
B2B[a] mode	151 (93.8)	58 (98)	50 (93)	43 (90)
DTC[b] mode	28 (17.4)	13 (22)	6 (11)	9 (19)
B2B2C[c] mode	32 (19.9)	14 (24)	12 (22)	6 (13)

[a]B2B: business-to-business.

[b]DTC: direct-to-consumer.

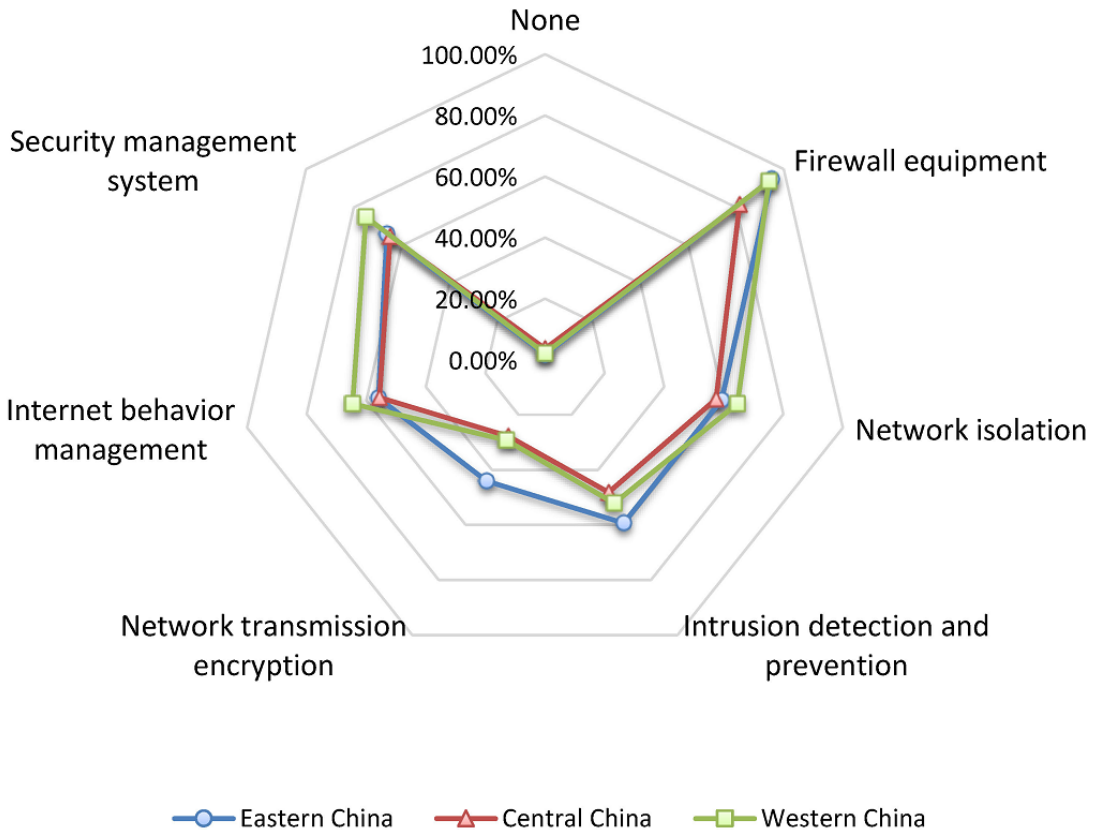
[c]B2B2C: business-to-business-to-customer.

XSL•FO
RenderX

Network Types and Security

Overall, the most widely used type of telemedicine network was virtual private network (VPN, a special telemedicine network constructed in China based on wired network) in 55.3% (89/161) of the hospitals, followed by internet (excluding VPN and wireless networks) in 43.5% (70/161) of the hospitals; both these networks are far ahead of 3G/4G (2/161, 1.2%) in the hospitals. In terms of regions, VPN and internet constituted the same percentage of the telemedicine networks in the central and eastern regions. Most of the hospitals (33/48, 69%) in the western region adopted VPN. With regard to the security of the telemedicine networks, overall, 97.5% (157/161) of the tertiary hospitals developed different cybersecurity measures. The proportion of the hospitals that utilized firewall equipment to ensure network security was the highest at 90.1% (145/161). The other 2 popular network security measures for telemedicine were the formulation of security management systems and the implementation of network isolation. Figure 3 illustrates the network security measures for telemedicine in each region.

Figure 3. Telemedicine network security measures of tertiary hospitals in different regions of China.



Data Storage and Security

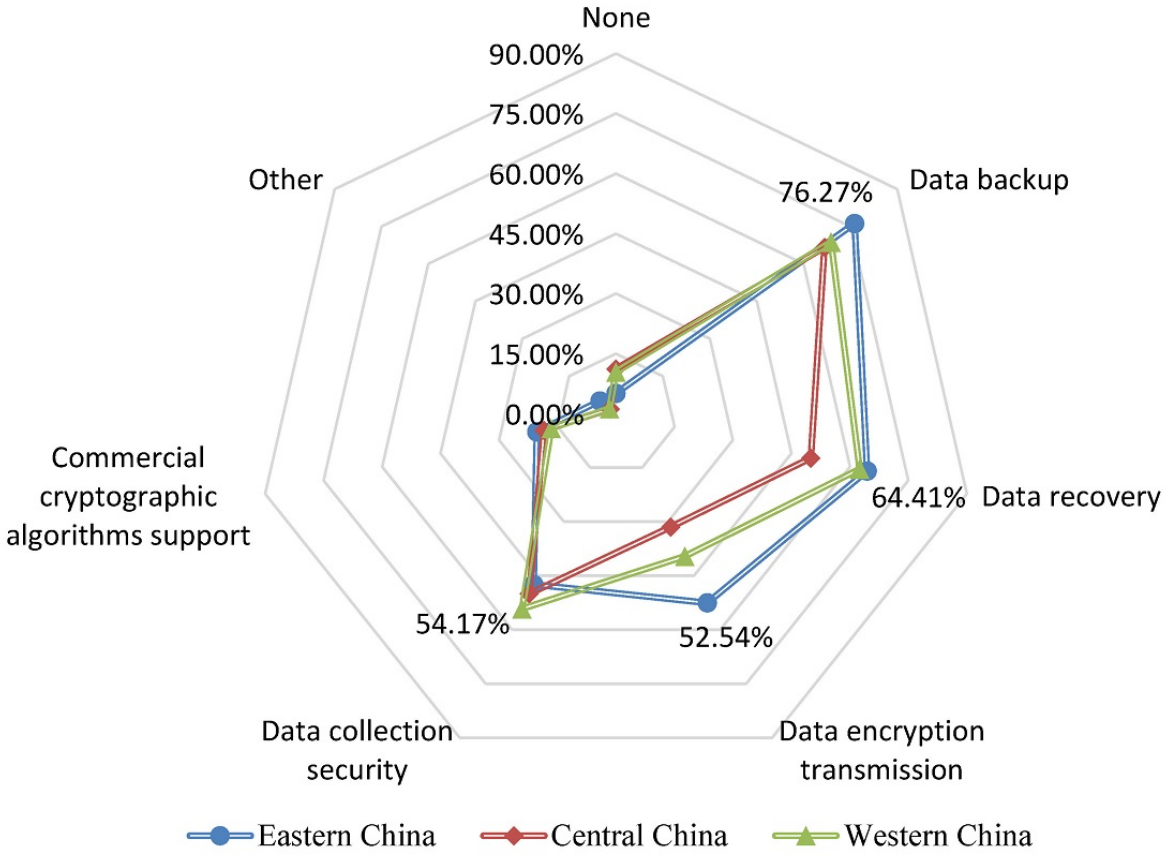
Table 4 indicates how tertiary hospitals store data generated by telemedicine services. Independent storage is a relatively common method for telemedicine data storage in 36.0% (58/161) of the tertiary hospitals. For telemedicine data security measures, only a few tertiary hospitals (14/161, 8.7%) have not established security measures. Among the security measures taken by tertiary hospitals, data backup is the foremost as 70.8% (114/161) of the tertiary hospitals carried out data backup. Followed by that, 59.0% (95/161) of the hospitals implemented data recovery measures after a failure. The results in Figure 4 suggest that data backup, data recovery, and data transmission encryption are the first 3 data security measures adopted by tertiary hospitals in the eastern region. In contrast, the top 3 data security methods in the central and western hospitals are data backup, data recovery, and data collection security.

XSL•FO
RenderX

Table 4. Telemedicine data storage methods of tertiary hospitals in different regions of China (N=161).

Data storage methods	Total, N=161, n (%)	Eastern region, n=59, n (%)	Central region, n=54, n (%)	Western region, n=48, n (%)
Independent storage	58 (36.0)	18 (31)	23 (43)	17 (35)
Sharing with other departments	45 (28.0)	19 (32)	14 (26)	12 (25)
Sharing with other hospitals	13 (8.1)	6 (10)	4 (7)	3 (6)
No storage	31 (19.3)	10 (17)	11 (20)	10 (21)
Other	14 (8.7)	6 (10)	2 (4)	6 (13)

Figure 4. Telemedicine data security measures for tertiary hospitals in different regions of China.



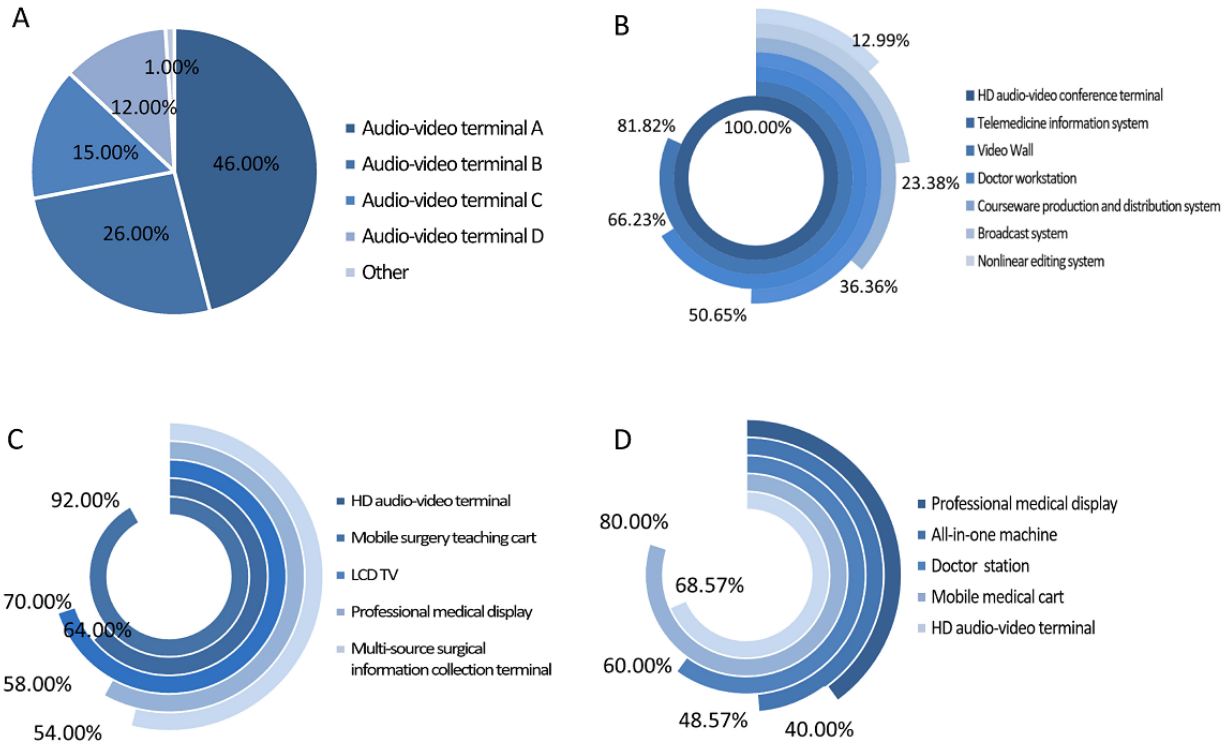
Hardware and Software Equipment

Of the 161 hospitals in the survey, 100 (62.1%) conducted remote videoconferencing supported by hardware for telemedicine, while other hospitals chose remote videoconferencing supported by software. The high-definition audio and video terminals were the most critical devices for teleconsultation with hardware videoconferencing, which were classified into 4 categories, that is, audio-video terminals type A, type B, type C, and type D, according to the Technical Guide for the Construction of Telemedicine Information System (2014 edition) [25]. As shown in Figure 5A, terminal type A (large multiscreens with the highest level of audio and video quality) was the mainstream, which was employed by 46.0% (46/100) of the hospitals, such as the Huawei telepresence conference system. Terminal type B (separate screen and camera) was the second most used at a usage rate of 26.0% (26/100), followed by integrated terminal type C (integrated codec and multiple audio and video interfaces) with a usage rate of 15.0% (15/100). Terminal type D (integrated camera and microphone) was used by 12.0% (12/100) of the hospitals.

Seventy-seven hospitals built remote education systems and were equipped with related hardware and software devices. The results showed that all of these hospitals were equipped with high-definition audio and videoconference terminals. Video walls and doctor workstations were also favored by many hospitals, which were present in 66% (51/77) and 51% (39/77) of the hospitals, respectively (Figure 5B). Figure 5C shows the equipment configuration for remote surgery teaching in 50 tertiary hospitals, in which 92% (46/50) of the hospitals were equipped with high-definition audio and video terminals. The usage rate of liquid-crystal display televisions (35/50, 70%) was higher than that of mobile surgery teaching carts (32/50, 64%). Only 35 tertiary hospitals developed remote ward rounds systems. The usage rate of mobile medical carts (28/35, 80%) overtook that of other devices for remote ward rounds.

XSL•FO
RenderX

Figure 5. Hardware and software equipment for telemedicine services in the tertiary hospitals of China. A. Types of high-definition (HD) audio and video terminals for teleconsultation in tertiary hospitals; B. Hardware and software device configuration for remote education in tertiary hospitals; C. Hardware device configuration for remote surgery teaching in tertiary hospitals; D. Hardware device configuration for remote ward rounds in tertiary hospitals.



Telemedicine Services

Different types of telemedicine services have been more or less implemented nationwide. The top 5 service types are teleconsultation, remote education, telediagnosis of medical images, tele-electrocardiography, and telepathology, with coverage rates of 86.3% (139/161), 57.1% (92/161), 49.7% (80/161), 37.9% (61/161), and 33.5% (54/161), respectively. Other telemedicine services such as remote intensive care unit care (18/161, 11.2%), remote nursing (13/161, 8.1%), and remote emergency care (16/161, 9.9%) were implemented relatively rarely. Figure 6 presents the types of telemedicine services in tertiary hospitals in different regions. The proportions of hospitals that conducted teleconsultations and remote ward rounds in the eastern and western regions were significantly higher than those in the central region (P<.001 and P=.02 respectively). However, there were no statistically significant differences in the proportions of the tertiary hospitals that carried out other types of telemedicine services in different regions (P>.05).

XSL•FO
RenderX

Figure 6. Development of various telemedicine services in the tertiary hospitals in different regions of China.

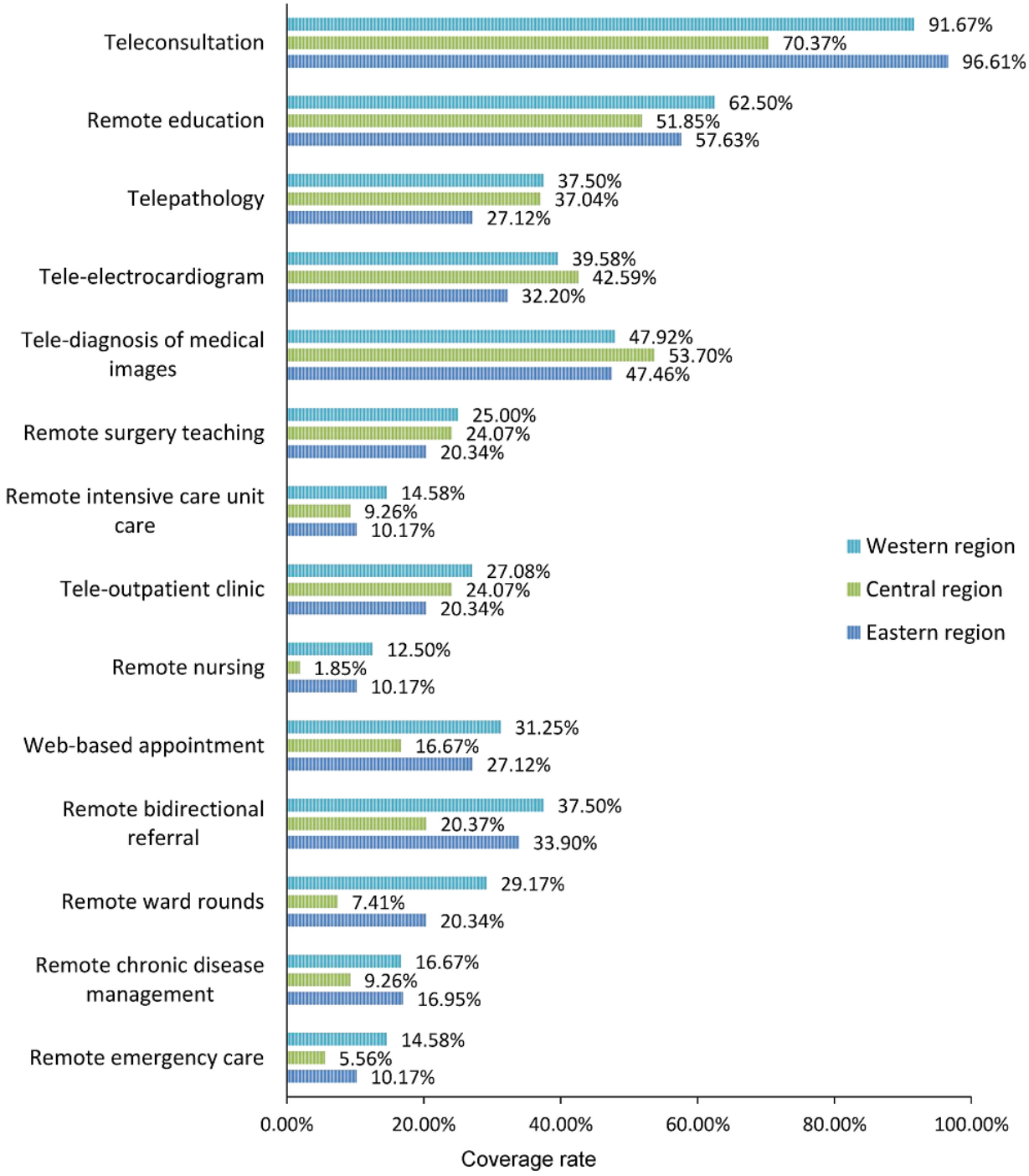


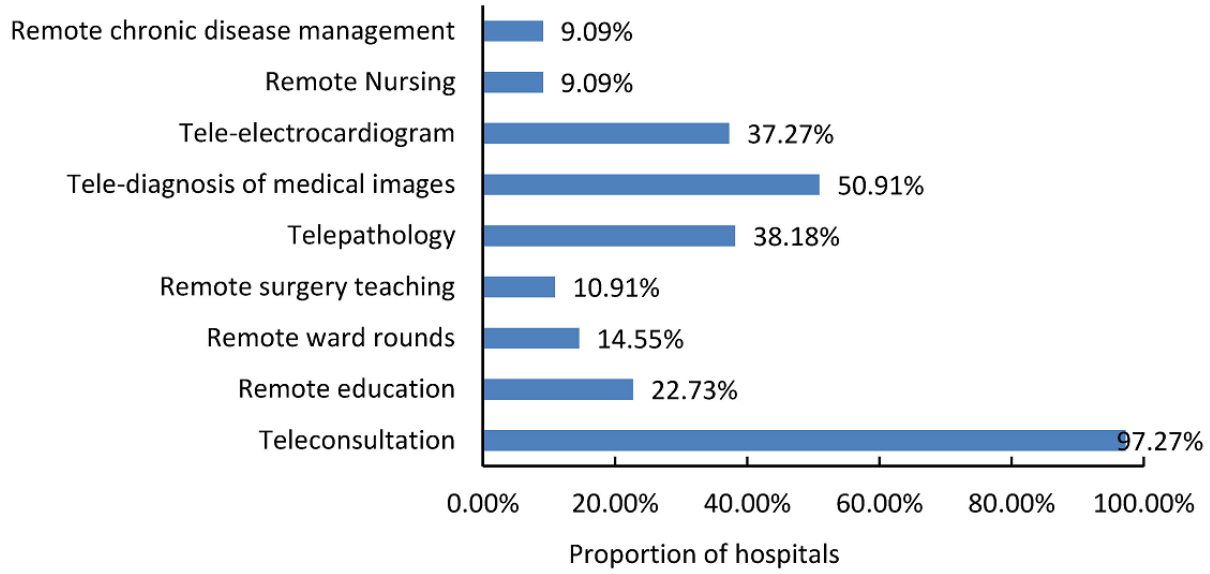
Table 5 shows the annual service quantity of 5 major telemedicine services in tertiary hospitals, in which the annual business volume of tele-electrocardiography ranked first. The annual service volumes of teleconsultation, remote education, and telediagnosis of medical images in the western region substantially outnumbered that in eastern and central regions. As for the telemedicine charge, overall, 71.7% (114/159) of the hospitals charged for telemedicine services in the survey. However, according to the medical insurance policies in different regions, only 22.8% (26/114) of the hospitals have included telemedicine charges into medical insurance. Teleconsultation and telediagnosis (telepathology, telediagnosis of medical images, and tele-electrocardiography) were the main chargeable items for telemedicine as indicated in Figure 7.

XSL•FO
RenderX

Table 5. Business volume of the major telemedicine services in tertiary hospitals (cases per hospital) in different regions of China.

Telemedicine service	Total (n)	Eastern region (n)	Central region (n)	Western region (n)
Teleconsultation	714	604	532	1045
Remote education	44	27	35	76
Telepathology	139	48	266	107
Tele-electrocardiography	3342	2645	4123	3292
Telediagnosis of medical images	1107	593	405	2443

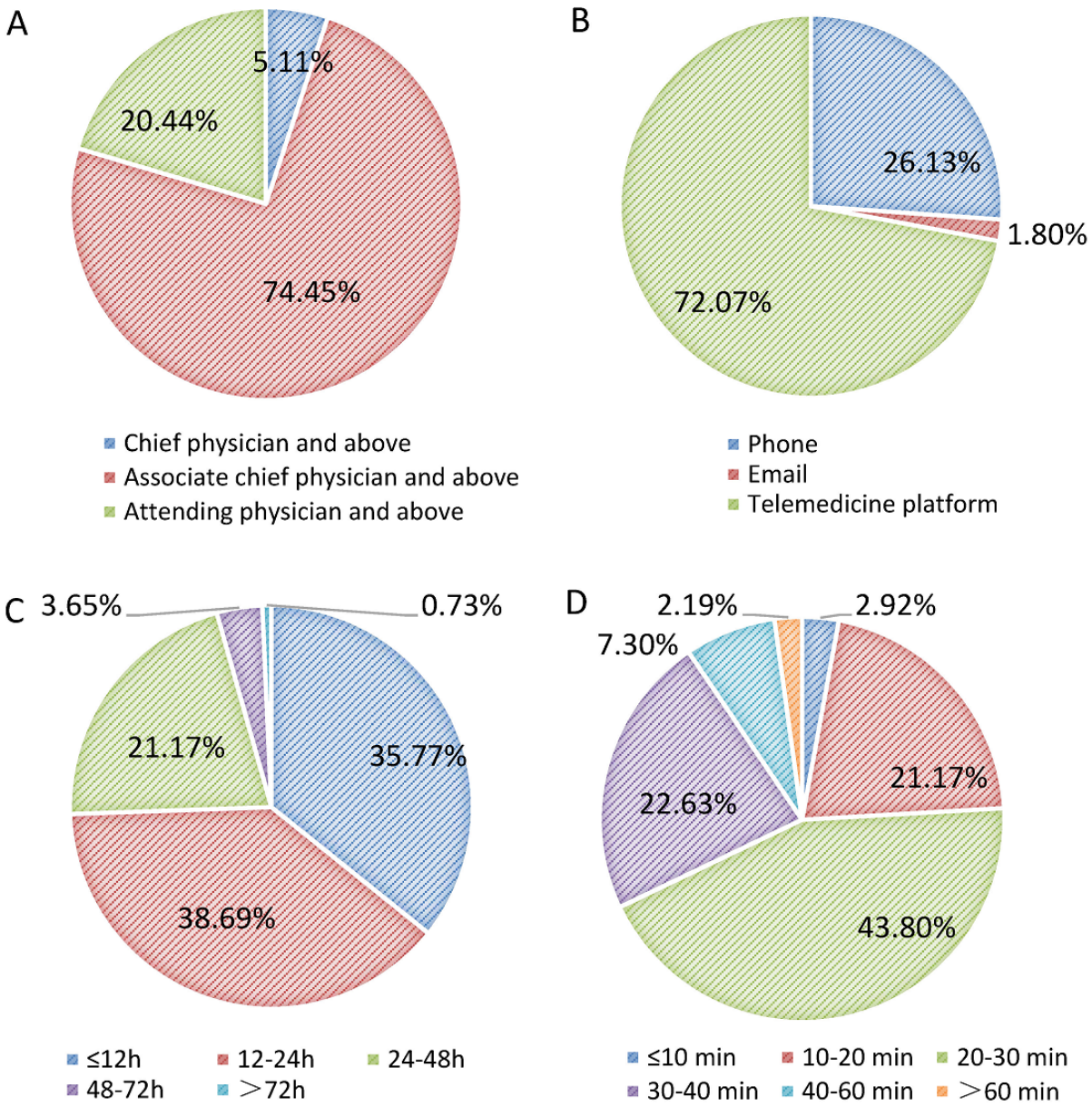
Figure 7. Proportion of hospitals charging for different telemedicine services.



Teleconsultation, as the core service of telemedicine, needs to be explored in depth. Figure 8A depicts that the level of medical experts providing teleconsultation services was deputy chief physician and above in most tertiary hospitals (109/137, 79.6%). Figure 8B illustrates that 72.1% (80/111) of the tertiary hospitals applied for teleconsultation through the telemedicine platform, which was the most important way. In terms of the response duration of teleconsultation, that is, the time interval from the consultation application to the start of the consultation, 74.5% (92/137) of the hospitals had a waiting time within 24 hours. Figure 8D shows that the average duration of each teleconsultation in most hospitals ranged from 10 to 40 minutes. Among them, the percentage of teleconsultations lasting 20-30 minutes was the highest in 43.8% (60/137) of the hospitals. As far as the effect of teleconsultation was concerned, the proportion of hospitals that considered the effect as good and excellent was 51.4% (57/111) and 31.5% (35/111), respectively, whereas 17.1% (19/111) of the hospitals believed the consultation effect to be fair or poor. Remote education is another core business of telemedicine. The overall participation frequency of remote education in tertiary hospitals was relatively low; 83.8% (93/111) of the hospitals performed 0-6 instances of remote education every month. The proportion of hospitals participating in remote education with a frequency of 7-10 cases, 11-14 cases, and 15 cases and above every month was 6.3% (7/111), 1.8% (2/111), and 8.1% (9/111), respectively. As for the effect of remote education, 73.0% (81/111) of the tertiary hospitals believed that it had a certain effect on improving the medical service level, and 26.1% (29/111) of the hospitals thought it had a great effect. However, 0.9% (1/111) of the hospitals considered that it had no effect.

XSL•FO
RenderX

Figure 8. Teleconsultation status in tertiary hospitals: A. Levels of teleconsultation specialists; B. Ways to apply for teleconsultation; C. Average waiting time for teleconsultation; D. Average duration of teleconsultation cases.



Key Factors Influencing Telemedicine Development

Standard formulation was the most crucial factor, as 68.3% (110/161) of the tertiary hospitals believed that the lack of uniform standards hindered the promotion of telemedicine in hospitals. The other influencing factors of the development of telemedicine in tertiary hospitals are listed in Figure 9. In order to explore the key factors influencing the effects of telemedicine further, we considered the core telemedicine services (ie, teleconsultation and remote education) as examples and applied ordinal logistic regression to analyze the relationship between the effect of telemedicine and the aforementioned factors such as human resources, funding, region, management, and service modes. The effect of teleconsultation refers to how useful it is for the treatment and health of the patient. Thus, the teleconsultation effect was divided into 3 categories (1=poor or fair, 2=good, and 3=excellent). The effect of remote education was expressed in terms of improvement in the medical service level (1=no improvement, 2=certain improvement, 3=great improvement), in which “certain improvement” means that remote education helps to improve the medical service level of the hospital, but the improvement is not significant. Taking the high-level effect as a reference, as many factors as possible were included in the models. Some results are presented in Table 6 and Table 7. The complete parameter estimation results are shown in the Multimedia Appendix 2 and Multimedia Appendix 3.

As shown in Table 6, an increase in the number of medical professionals may have a positive impact on the teleconsultation effect (P=.04). The adoption of the DTC mode, the support of scientific research funds, and the charging for consultation

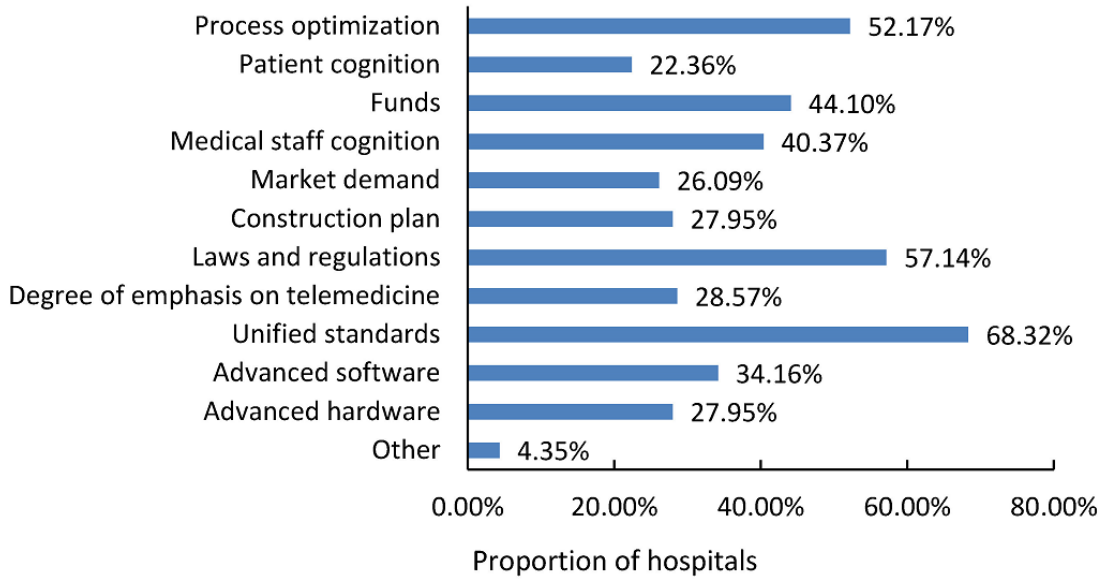
XSL•FO
RenderX

services could improve the teleconsultation effect, with *P*=.003, *P*=.01, and *P*<.001, respectively. Compared with 3G/4G, VPN and internet could significantly improve the consultation effect (*P*=.02 and *P*=.02, respectively). No data storage would reduce the consultation effect compared with other data storage methods, as no statistical significance was found (*P*=.07). Sharing data with other hospitals could significantly improve the consultation effect (*P*=.04). As the duration of the consultation increased, the consultation effect decreased to varying degrees. High-level experts were helpful in improving the consultation effect (*P*=.03).

Similarly, the adoption of the DTC mode, the provision of research funding, and the charging for service had a positive impact on remote education by improving the medical service level. However, regional differences had no impact on the effect of remote education. Compared with hospitals that did not establish professional management departments, the remote education effects of hospitals that were establishing departments or with already established departments were better, with *P*=.04 and *P*=.25, respectively. Compared with the 0-3 instances of remote education every month, the frequency of 15 instances and above every month could significantly improve the effect of remote education (*P*=.01).

For this study, the Checklist for Reporting Results of Internet e-Surveys has been uploaded in Multimedia Appendix 4.

Figure 9. Key factors affecting the development of telemedicine in tertiary hospitals of China.



XSL•FO
RenderX

Table 6. Ordinal logistic regression results of the influencing factors of the consultation effect.

Variable	Regression coefficient	Standard error	Wald statistic	*P* value
Educational background of the professionals				
Computer science and communication professionals	–0.31	0.19	2.63	.11
Medical professionals	0.09	0.04	4.09	.04
Management professionals	–0.27	0.14	3.87	.049
Service mode				
B2B[a]mode				
Yes	1.25	2.45	0.26	.61
No[b]	N/A[c]	N/A	N/A	N/A
DTC[d]mode				
Yes	3.54	1.20	8.68	.003
No	N/A	N/A	N/A	N/A
B2B2C[e]mode				
Yes	–0.57	1.15	0.25	.62
No	N/A	N/A	N/A	N/A
Investment amount (RMB)[f] (reference=less than 100,000 RMB)				
>5 million	–1.26	1.57	0.65	.42
1-5 million	0.01	1.04	0.00	.99
500,000-1 million	–3.26	1.25	6.79	.01
100,000-500,000	–0.61	0.85	0.52	.47
Funds				
Government financial support				
Yes	–0.70	0.92	0.58	.45
No	N/A	N/A	N/A	N/A
Hospital self-raising				
Yes	–1.47	1.17	1.57	.21
No	N/A	N/A	N/A	N/A
Research funding				
Yes	3.72	1.51	6.09	.01
No	N/A	N/A	N/A	N/A
Corporate sponsorship				
Yes	–0.18	1.34	0.02	.89
No	N/A	N/A	N/A	N/A
Network types (reference=3G/4G)				
Virtual private network	4.55	2.02	5.08	.02
Internet	4.75	2.00	5.67	.02
Data storage (reference=other)				
Independent storage	1.41	1.22	1.34	.25
Sharing with other departments	–0.16	1.29	0.02	.90
Sharing with other hospitals	3.66	1.79	4.19	.04
No storage	–2.43	1.32	3.40	.07
Expertise level (reference=attending physician and above)				

XSL•FO
RenderX

Variable	Regression coefficient	Standard error	Wald statistic	*P* value
Chief physician and above	5.29	2.37	4.99	.03
Associate chief physician and above	1.01	0.82	1.51	.22
Duration of consultation per case (reference≤10 min)				
60 min	–11.10	3.50	10.06	.002
40-60 min	–5.50	2.96	3.45	.06
30-40 min	–7.83	3.05	6.61	.01
20-30 min	–6.08	2.89	4.42	.04
10-20 min	–3.37	2.66	1.61	.21
Charge				
Yes	4.26	1.02	17.30	<.001
No	N/A	N/A	N/A	N/A

[a]B2B: business-to-business.

[b]Reference.

[c]N/A: not applicable.

[d]DTC: direct-to-consumer.

[e]B2B2C: business-to-business-to-customer.

[f]1 RMB=US $0.14.

XSL•FO
RenderX

Table 7. Ordinal logistic regression results of the influencing factors of the remote education effect.

Variable	Regression coefficient	Standard error	Wald statistic	*P* value
Region (reference=center)				
East	–1.04	1.10	0.90	.34
West	–1.66	1.16	2.03	.15
Service mode				
B2B[a]mode				
Yes	5.70	4.68	1.48	.22
No[b]	N/A[c]	N/A	N/A	N/A
DTC[d]mode				
Yes	4.01	1.59	6.39	.01
No	N/A	N/A	N/A	N/A
B2B2C[e]mode				
Yes	–2.57	1.47	3.05	.08
No	N/A	N/A	N/A	N/A
Professional management department (reference=not established)				
Established	1.45	1.27	1.31	.25
Being established	3.67	1.78	4.28	.04
Investment amount (RMB)[f] (reference=less than 100,000 RMB)				
>5 million	0.60	2.26	0.07	.79
1-5 million	0.19	1.48	0.02	.90
500,000-1 million	–2.22	1.82	1.48	.22
100,000-500,000	0.39	1.12	0.12	.73
Funds				
Government financial support				
Yes	–2.22	1.22	3.31	.07
No	N/A	N/A	N/A	N/A
Hospital self-raising				
Yes	–2.18	1.53	2.03	.15
No	N/A	N/A	N/A	N/A
Research funding				
Yes	4.69	1.83	6.61	.01
No	N/A	N/A	N/A	N/A
Corporate sponsorship				
Yes	–1.77	1.65	1.15	.28
No	N/A	N/A	N/A	N/A
Frequency of remote education (reference=0-3 times a month)				
15 instances and above per month	5.95	2.11	7.94	.01
11-14 instances per month	–3.64	6.51	0.31	.58
7-10 instances per month	2.99	1.69	3.13	.08
4-6 instances per month	0.14	1.11	0.02	.90
Charge				
Yes	4.30	1.53	7.97	.01

XSL•FO
RenderX

Variable	Regression coefficient	Standard error	Wald statistic	*P* value
No	N/A	N/A	N/A	N/A

[a]B2B: business-to-business.

[b]Reference.

[c]N/A: not applicable.

[d]DTC: direct-to-consumer.

[e]B2B2C: business-to-business-to-customer.

[f]1 RMB=US $0.14.

Discussion

Principal Findings

Based on a national survey, this study analyzed the development of telemedicine in Chinese hospitals from multiple aspects, including implementation, application, and the key factors influencing telemedicine service effects. We found that telemedicine services were mainly carried out in the form of hardware videoconferences in B2B mode through VPN in Chinese tertiary hospitals, with various service types and a large service quantity, which were considered to have positive effects on the improvement of medical treatment in primary hospitals. Despite the rapid development of telemedicine in China, there are still problems such as the lack of uniform standards and laws, which reminds us that a lot of work is needed to improve the standardization of telemedicine services and establish legal protection for telemedicine services.

In terms of human resource and management, 31.7% (51/161) of the tertiary hospitals had no specific management departments, and some tertiary hospitals did not even have full-time staff. Further, the number of telemedicine staff with majors in computer science and communication and management was relatively small. As a burgeoning, cutting-edge, and multidisciplinary technology, telemedicine is in high need of a compound talent team that integrates the knowledge structure of science, engineering, and medicine. According to the technical guide of telemedicine in China, a tertiary hospital should set up an independent telemedicine department with medical, information technology, and management professionals and technical personnel [25]. However, many hospitals do not meet the requirements of the technical guide; therefore, more work is still needed to improve the organization of a telemedicine management department. Previous studies have suggested that sufficient capital investment and leadership attention are the primary factors affecting the development of remote education and remote pathology diagnosis [26]. In China, 35.4% (57/161) of the hospitals invest more than 500,000 RMB (approximately US $71,218) per year, and 28.0% (45/161) of the hospitals invest 100,000-500,000 RMB (approximately US $14,244-$71,218) per year, thereby providing financial support for telemedicine development. Government investments and hospital self-raising were the main sources of funds, which were consistent with the results of another telemedicine survey [27], thereby revealing the national attention on telemedicine.

Through the investigation, we found that B2B was the main telemedicine service mode in China. Historically, as the original intention of telemedicine is that experts in tertiary hospitals assist doctors in secondary or primary hospitals at different locations to provide solutions to the management of complicated diseases by videoconferencing, telemedicine services are still mainly limited to the B2B mode, and the DTC mode remains to be improved [28,29]. However, with the development of medical information, the DTC mode is considered to be a further developmental tendency for telemedicine [30,31], which will enable the patients to apply for telemedicine services with greater accessibility and convenience, such as choosing the doctors they prefer [32]. In 2018, a policy was issued in China to allow medical institutions to set up internet hospitals and to directly carry out telemedicine services for revisiting patients with common and chronic diseases. To date, about 158 internet hospitals in China have been built to provide web-based telemedicine consultations directly to the patients [33]. We believe that the telemedicine service mode will change from B2B to DTC in the near future.

According to the results of the network construction, VPN was the main type of network used for telemedicine. With the advantages of high security, flexible access, and low latency [34], VPN is the preferred network for telemedicine in the technical guide [25]. However, 44.7% (72/161) of hospitals still do not have access to VPN. A policy was introduced in China in 2018 to encourage telecom enterprises to provide high-quality dedicated internet access and VPN for medical institutions and to promote the construction of dedicated networks for telemedicine and guarantee the quality of medical-related data transmission services. It can be predicted that private networks will be the predominant type of network used for telemedicine services in the future. Telemedicine provides more chances for accessing patients' personal information. However, this is accompanied by a threat to the security of the medical data [35]. Therefore, security measures should be formulated in the implementation of telemedicine. In China, 91.3% (147/161) and 97.5% (157/161) of the tertiary hospitals have established data security measures and network security measures, respectively, such as firewall, network isolation, and internet behavior supervision, thereby reflecting their emphasis on information security.

The survey results showed that 62.1% (100/161) of the hospitals conducted hardware videoconferences instead of software videoconferences for teleconsultation. Despite the availability of low-cost videoconferencing software, due to the high requirements for information transmission and video quality, many hospitals have adopted high-performance hardware videoconferencing owing to the advantages of high definition, security, stability, and interoperability [36]. In 2017, an average

XSL•FO
RenderX

of 714 remote consultations were conducted in each tertiary hospital, indicating a huge increase compared to the number of teleconsultations in 2014 (no more than 7 times per quarter) [27]. For different types of remote diagnosis carried out in tertiary hospitals, the service quantity of tele-electrocardiography diagnosis ranked first, with an average number of 3342 cases in each hospital, which was related to the relatively low requirements on equipment and network speed [37]. Unlike the remote consultation and diagnosis, the application of telemedicine is relatively limited in intensive care and nursing and emergency care because of the immature technology, lack of legislation on charging standards and responsibility distinction, and the low acceptance by patients and medical staff [27,38,39].

Studies have revealed that the lack of uniform standards and laws was an essential factor hindering the development of telemedicine, which might lead to repeated and chaotic implementation of telemedicine [40]. Some scholars believe that since physical examination cannot be performed in telemedicine, medical safety cannot be guaranteed, and the responsibilities for the privacy and safety of patients are not clearly assigned to the doctors participating in telemedicine services [41-43]. Therefore, new legislation on telemedicine should be introduced to guide the implementation of telemedicine and assign the medical responsibilities clearly. Another factor hindering the development of telemedicine is the imperfect medical insurance system. In the United States of America, the proportion of medical insurance payment for telemedicine is 0%-67% [44], whereas in China, only some provinces such as Guizhou and Sichuan realize the medical insurance reimbursement. The government should formulate a medical insurance policy as well as develop a reasonable benefit distribution mechanism to promote the sustainable implementation of telemedicine in China.

This study shows that most hospitals believe that telemedicine is effective and could improve the medical service level of the hospitals, which was consistent with the results of many studies that positively evaluated telemedicine [45,46]. A study showed that as the frequency of consultation increased, primary care providers could significantly improve knowledge acquisition [47]. The DTC mode, support of scientific research funds, and service charges had a significant impact on the effects of both teleconsultation and remote education. Compared to the mainstream B2B mode, the obvious feature of the DTC mode is the participation of patients, which may help to obtain more accurate patient information during the consultation process and allow patients to interact with doctors. The supports for research funding and service charges can significantly increase the enthusiasm of doctors. With the encouragement of the government, many provinces in China have issued pricing and payment policies for telemedicine services. As a result, the charges for telemedicine services will be rule-based, more standardized, and reasonable in the future.

Comparison With Prior Work

Although previous studies have analyzed the utilization of telemedicine in a certain region, the overall development of telemedicine in China has not yet been studied, especially in terms of network construction, security measures, and hardware and software facilities. This study investigated the implementation and application of telemedicine from a national perspective, which will provide people with a comprehensive, multilevel, and multifaceted understanding of telemedicine development in China. Moreover, to our knowledge, this is the first study to use ordinal regression models to deeply analyze the factors influencing the effectiveness of telemedicine in multiple dimensions, including human resources, funding, management and service modes, networks, and charging, which will supply a reference for telemedicine planning.

Limitations

Although our findings provide a deep insight into the development of telemedicine in China, this study has several limitations. As the first nationwide survey on telemedicine in China, our sample is nationally representative and covers most areas of China. However, the sample size is still insufficient, as some areas with few tertiary hospitals were not included, such as Tibet; therefore, the scope of the research needs to be further expanded. Besides, we have only considered tertiary hospitals as the survey object; however, patients' attitude toward telemedicine is also important for the development of telemedicine, which will be the focus of the next study.

Conclusions

We conducted a quantitative analysis of the overall implementation and application of telemedicine in China with the data from 161 tertiary hospitals in 29 provinces, autonomous regions, and municipalities. Our findings revealed that telemedicine services were carried out in most parts of China, and most tertiary hospitals provided telemedicine services in B2B mode through the hardware videoconferencing. VPN was the most widely used type of telemedicine network, and audio-video terminals with large screens was the mainstream hardware. Teleconsultation, remote education, and telediagnosis were the main types of telemedicine service. Service modes, financial sources, network types, service charges, and medical experts are the main factors influencing the effect of teleconsultation and remote education. The management, uniform standards, and legislation still need to be improved for the sustainability of telemedicine in China. To plan the development of telemedicine further, our research provides a reference for policymakers to promote the implementation of DTC mode of telemedicine, expand the coverage of VPN, develop innovative service patterns such as remote nursing and remote care in intensive care units, and formulate complete telemedicine laws and regulations.

XSL•FO
RenderX

Acknowledgments

This research was supported by grants from the National Key Research and Development Program of China (Grant No. 2017YFC0909900), the National Natural Science Foundation of China (Grant No. 71673254), and the Innovation Research Team of Higher Education in Henan Province (Grant No. 20IRTSTHN028). Special thanks to the Telemedicine Information Professional Committee of China for providing a platform to conduct the telemedicine survey.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Web-based questionnaire survey (translated version).
[PDF File (Adobe PDF File), 358 KB - mhealth_v8i10e18426_app1.pdf]

Multimedia Appendix 2
Estimation results of ordinal regression analyzing the influencing factors of teleconsultation.
[PDF File (Adobe PDF File), 422 KB - mhealth_v8i10e18426_app2.pdf]

Multimedia Appendix 3
Estimation results of ordinal regression analyzing the influencing factors of remote education.
[PDF File (Adobe PDF File), 306 KB - mhealth_v8i10e18426_app3.pdf]

Multimedia Appendix 4
Data reporting guidelines and checklist for reporting results of internet e-surveys (CHERRIES).
[PDF File (Adobe PDF File), 23 KB - mhealth_v8i10e18426_app4.pdf]

References

1. World Health Organization. Telemedicine: Opportunities and developments in member states. In: Report on the second Global survey on eHealth (Global Observatory for eHealth Series, Volume 2). Geneva: World Health Organization; 2010.
2. Clark PA, Capuzzi K, Harrison J. Telemedicine: medical, legal and ethical perspectives. Med Sci Monit 2010 Dec;16(12):RA261-RA272. [Medline: 21119593]
3. Peltan ID, Poll JB, Guidry D, Brown SM, Beninati W. Acceptability and Perceived Utility of Telemedical Consultation during Cardiac Arrest Resuscitation. A Multicenter Survey. Annals ATS 2020 Mar 16;17(3):321-328. [doi: 10.1513/annalsats.201906-485oc]
4. Witkowski A, Łudzik J, Soyer HP. Telediagnosis with Confocal Microscopy: A Reality or a Dream? Dermatol Clin 2016 Oct;34(4):505-512. [doi: 10.1016/j.det.2016.05.013] [Medline: 27692456]
5. Hassan A, Ghafoor M, Tariq SA, Zia T, Ahmad W. High Efficiency Video Coding (HEVC)-Based Surgical Telementoring System Using Shallow Convolutional Neural Network. J Digit Imaging 2019 Dec;32(6):1027-1043. [doi: 10.1007/s10278-019-00206-2] [Medline: 30980262]
6. Tsirintani M. Strategic Procedures and Revisions for implementing Telemedicine and Telecare in Greece. Appl Clin Inform 2017 Dec 16;03(01):14-23. [doi: 10.4338/aci-2011-08-r-0048]
7. Uscher-Pines L, Mulcahy A, Cowling D, Hunter G, Burns R, Mehrotra A. Access and Quality of Care in Direct-to-Consumer Telemedicine. Telemed J E Health 2016 Apr;22(4):282-287 [FREE Full text] [doi: 10.1089/tmj.2015.0079] [Medline: 26488151]
8. Anand S, Fan VY, Zhang J, Zhang L, Ke Y, Dong Z, et al. China's human resources for health: quantity, quality, and distribution. Lancet 2008 Nov 15;372(9651):1774-1781. [doi: 10.1016/S0140-6736(08)61363-X] [Medline: 18930528]
9. National Health Commission of the People's Republic of China. Chinese Health Statistics Yearbook [In Chinese]. Beijing: Beijing Union Medical University Press; 2019.
10. Xie X, Liu P, Zheng Y, Zhou W, Zou J, Wang X, et al. Equity of health resource distribution in China during 2009–15: an analysis of cross-sectional nationwide data. The Lancet 2017 Dec;390:S6. [doi: 10.1016/s0140-6736(17)33144-6]
11. Sun J, Luo H. Evaluation on equality and efficiency of health resources allocation and health services utilization in China. Int J Equity Health 2017 Jul 14;16(1):127 [FREE Full text] [doi: 10.1186/s12939-017-0614-y] [Medline: 28709422]
12. Liu W, Liu Y, Twum P, Li S. National equity of health resource allocation in China: data from 2009 to 2013. Int J Equity Health 2016 Apr 19;15:68 [FREE Full text] [doi: 10.1186/s12939-016-0357-1] [Medline: 27093960]
13. Wang T, Li J, Zhu C, Hong Z, An D, Yang H, et al. Assessment of Utilization and Cost-Effectiveness of Telemedicine Program in Western Regions of China: A 12-Year Study of 249 Hospitals Across 112 Cities. Telemed J E Health 2016 Nov;22(11):909-920. [doi: 10.1089/tmj.2015.0213] [Medline: 27314300]

XSL•FO
RenderX

14. Menachemi N, Burke DE, Ayers DJ. Factors affecting the adoption of telemedicine--a multiple adopter perspective. J Med Syst 2004 Dec;28(6):617-632. [doi: 10.1023/b:joms.0000044964.49821.df] [Medline: 15615290]
15. Hanson RE, Truesdell M, Stebbins GT, Weathers AL, Goetz CG. Telemedicine vs Office Visits in a Movement Disorders Clinic: Comparative Satisfaction of Physicians and Patients. Mov Disord Clin Pract 2019 Jan;6(1):65-69 [FREE Full text] [doi: 10.1002/mdc3.12703] [Medline: 30746418]
16. Lee JY, Lee SWH. Telemedicine Cost-Effectiveness for Diabetes Management: A Systematic Review. Diabetes Technol Ther 2018 Jul;20(7):492-500. [doi: 10.1089/dia.2018.0098] [Medline: 29812965]
17. Zheng Y, Lin Y, Cui Y. Teledermatology in China: History, Current Status, and the Next Step. Journal of Investigative Dermatology Symposium Proceedings 2018 Dec;19(2):S71-S73 [FREE Full text] [doi: 10.1016/j.jisp.2018.09.003] [Medline: 30471754]
18. Zhai Y, Gao J, Chen B, Shi J, Wang L, He X, et al. Design and Application of a Telemedicine System Jointly Driven by Videoconferencing and Data Exchange: Practical Experience from Henan Province, China. Telemed J E Health 2020 Jan;26(1):89-100. [doi: 10.1089/tmj.2018.0240] [Medline: 31050599]
19. Cai H, Wang H, Guo T, Bao G. Application of Telemedicine in Gansu Province of China. PLoS One 2016;11(6):e0158026 [FREE Full text] [doi: 10.1371/journal.pone.0158026] [Medline: 27332894]
20. National Health Commission of the People's Republic of China. Transcript of the regular press conference of the National Health and Family Planning Commission on February 7, 2018. URL: http://www.nhc.gov.cn/wjw/xwdt/201802/edfe83b3fbd144cf8ef5592404210b65.shtml [accessed 2019-12-15]
21. Wang G, Zhang Z, Feng Y, Sun L, Xiao X, Wang G, et al. Telemedicine in the Management of Type 2 Diabetes Mellitus. Am J Med Sci 2017 Jan;353(1):1-5. [doi: 10.1016/j.amjms.2016.10.008] [Medline: 28104096]
22. Zhai Y, Zhu W, Cai Y, Sun D, Zhao J. Clinical- and cost-effectiveness of telemedicine in type 2 diabetes mellitus: a systematic review and meta-analysis. Medicine (Baltimore) 2014 Dec;93(28):e312 [FREE Full text] [doi: 10.1097/MD.0000000000000312] [Medline: 25526482]
23. Cheng YY, Wang H, Liu LL, Zhang LP, Shen QF, Meng MF. Application status and prospects of telemedicine in the field of burns. Chin J Burns 2019 Sep 20;35(9):697-700. [doi: 10.3760/cma.j.issn.1009-2587.2019.09.010] [Medline: 31594190]
24. He C, Zhou Q, Chen W, Tian J, Zhou L, Peng H, et al. Using an Internet-Based Hospital to Address Maldistribution of Health Care Resources in Rural Areas of Guangdong Province, China: Retrospective and Descriptive Study. JMIR Med Inform 2018 Dec 21;6(4):e51 [FREE Full text] [doi: 10.2196/medinform.9495] [Medline: 30578195]
25. National Health and Family Planning Commission of the People's Republic of China. Technical Guide for the Construction of Telemedicine Information System (2014 Edition). URL: http://www.nhc.gov.cn/mohwsbwstjxxzx/s8555/201501/120c912de1c4474c95b345d001271368/files/92f278e879d743de8e2ba5f38c1ed3ae.pdf [accessed 2019-12-08]
26. Alami H, Gagnon MP, Wootton R, Fortin JP, Zanaboni P. Exploring factors associated with the uneven utilization of telemedicine in Norway: a mixed methods study. BMC Med Inform Decis Mak 2017 Dec 28;17(1):180 [FREE Full text] [doi: 10.1186/s12911-017-0576-4] [Medline: 29282048]
27. Gong XY, Su T, Yang K, Zhang TS. Study on Telemedicine Development in China [In Chinese]. Chinese Journal of Health Informatics and Management. 2015 Dec. URL: http://med.wanfangdata.com.cn/Paper/Detail?id=PeriodicalPaper_zgwsxxglzz201502013&dbid=WF_QK [accessed 2020-03-04]
28. Becker CD, Dandy K, Gaujean M, Fusaro M, Scurlock C. Legal Perspectives on Telemedicine Part 1: Legal and Regulatory Issues. Perm J 2019;23 [FREE Full text] [doi: 10.7812/TPP/18-293] [Medline: 31314722]
29. Brous E. Legal Considerations in Telehealth and Telemedicine. AJN, American Journal of Nursing 2016;116(9):64-67. [doi: 10.1097/01.naj.0000494700.78616.d3]
30. Uscher-Pines L, Mulcahy A, Cowling D, Hunter G, Burns R, Mehrotra A. Access and Quality of Care in Direct-to-Consumer Telemedicine. Telemed J E Health 2016 Apr;22(4):282-287 [FREE Full text] [doi: 10.1089/tmj.2015.0079] [Medline: 26488151]
31. Ray KN, Shi Z, Poon SJ, Uscher-Pines L, Mehrotra A. Use of Commercial Direct-to-Consumer Telemedicine by Children. Acad Pediatr 2019 Aug;19(6):665-669 [FREE Full text] [doi: 10.1016/j.acap.2018.11.016] [Medline: 30639759]
32. Welch BM, Harvey J, O'Connell NS, McElligott JT. Patient preferences for direct-to-consumer telemedicine services: a nationwide survey. BMC Health Serv Res 2017 Nov 28;17(1):784 [FREE Full text] [doi: 10.1186/s12913-017-2744-8] [Medline: 29183372]
33. Xinhua News Agency. China builds 158 internet hospitals amid digital tide. 2019 May 8. URL: http://english.www.gov.cn/state_council/ministries/2019/05/08/content_281476650670642.htm [accessed 2019-12-15]
34. Hu X, Deng Y, Lu Z, Li S, Wang G, Lu S. Application of VPN technique in the construction of public health information system. J Huazhong Univ Sci Technolog Med Sci 2005;25(5):612-614. [doi: 10.1007/BF02896033] [Medline: 16463690]
35. Langarizadeh M, Moghbeli F, Aliabadi A. Application of Ethics for Providing Telemedicine Services and Information Technology. Med Arch 2017 Oct;71(5):351-355 [FREE Full text] [doi: 10.5455/medarh.2017.71.351-355] [Medline: 29284905]
36. Dinevski D, Kelc R, Dugonik B. Video communication in telemedicine. In: Graschew G, Roelofs TA, editors. Advances in Telemedicine: Technologies, Enabling Factors and Scenarios. Rijeka, Croatia: Intech Open; 2011.

XSL•FO
RenderX

37. Otto CM. Heartbeat: Telemedicine for echocardiography screening. Heart 2019 Feb;105(4):261-263. [doi: 10.1136/heartjnl-2019-314705] [Medline: 31693482]
38. Barbosa IDA, Silva KCDCDD, Silva VAD, Silva MJPD. The communication process in Telenursing: integrative review. Rev Bras Enferm 2016;69(4):765-772. [doi: 10.1590/0034-7167.2016690421i] [Medline: 27508484]
39. Moeckli J, Cram P, Cunningham C, Reisinger HS. Staff acceptance of a telemedicine intensive care unit program: a qualitative study. J Crit Care 2013 Dec;28(6):890-901. [doi: 10.1016/j.jcrc.2013.05.008] [Medline: 23906904]
40. Krupinski EA, Bernard J. Standards and Guidelines in Telemedicine and Telehealth. Healthcare (Basel) 2014 Feb 12;2(1):74-93 [FREE Full text] [doi: 10.3390/healthcare2010074] [Medline: 27429261]
41. Chaet D, Clearfield R, Sabin JE, Skimming K, Council on EthicalJudicial Affairs American Medical Association. Ethical practice in Telehealth and Telemedicine. J Gen Intern Med 2017 Oct;32(10):1136-1140 [FREE Full text] [doi: 10.1007/s11606-017-4082-2] [Medline: 28653233]
42. Hall JL, McGraw D. For telehealth to succeed, privacy and security risks must be identified and addressed. Health Aff (Millwood) 2014 Feb;33(2):216-221. [doi: 10.1377/hlthaff.2013.0997] [Medline: 24493763]
43. Cotet AM, Benjamin DK. Medical regulation and health outcomes: the effect of the physician examination requirement. Health Econ 2013 Apr;22(4):393-409. [doi: 10.1002/hec.2807] [Medline: 22450959]
44. Lin JC, Kavousi Y, Sullivan B, Stevens C. Analysis of Outpatient Telemedicine Reimbursement in an Integrated Healthcare System. Ann Vasc Surg 2020 May;65:100-106. [doi: 10.1016/j.avsg.2019.10.069] [Medline: 31678131]
45. Khairat S, Liu S, Zaman T, Edson B, Gianforcaro R. Factors Determining Patients' Choice Between Mobile Health and Telemedicine: Predictive Analytics Assessment. JMIR Mhealth Uhealth 2019 Jun 08;7(6):e13772 [FREE Full text] [doi: 10.2196/13772] [Medline: 31199332]
46. D'Haeseleer M, Eelen P, Sadeghi N, D'Hooghe MB, Van Schependom J, Nagels G. Feasibility of Real Time Internet-Based Teleconsultation in Patients With Multiple Sclerosis: Interventional Pilot Study. J Med Internet Res 2020 Aug 13;22(8):e18178 [FREE Full text] [doi: 10.2196/18178] [Medline: 32447274]
47. Zakaria A, Maurer T, Su G, Amerson E. Impact of teledermatology on the accessibility and efficiency of dermatology care in an urban safety-net hospital: A pre-post analysis. J Am Acad Dermatol 2019 Dec;81(6):1446-1452. [doi: 10.1016/j.jaad.2019.08.016] [Medline: 31415834]

Abbreviations

B2B: business-to-business
B2B2C: business-to-business-to-customer
DTC: direct-to-consumer
TIPC: Telemedicine Information Professional Committee of China
VPN: virtual private network

Edited by G Eysenbach; submitted 26.02.20; peer-reviewed by Y Liu, F Verbeke, F Mallor; comments to author 29.06.20; revised version received 29.08.20; accepted 02.10.20; published 23.10.20.

Please cite as:
Cui F, Ma Q, He X, Zhai Y, Zhao J, Chen B, Sun D, Shi J, Cao M, Wang Z
Implementation and Application of Telemedicine in China: Cross-Sectional Study
JMIR Mhealth Uhealth 2020;8(10):e18426
URL: http://mhealth.jmir.org/2020/10/e18426/
doi:10.2196/18426
PMID:33095175

©Fangfang Cui, Qianqian Ma, Xianying He, Yunkai Zhai, Jie Zhao, Baozhan Chen, Dongxu Sun, Jinming Shi, Mingbo Cao, Zhenbo Wang. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 23.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Home-Based Monitoring and Telemonitoring of Complicated Pregnancies: Nationwide Cross-Sectional Survey of Current Practice in the Netherlands

Josephus F M van den Heuvel[1], MD; Samira Ayubi[1], MD; Arie Franx[1,2], MD, PhD; Mireille N Bekker[1], MD, PhD

[1]Department of Obstetrics, University Medical Center, Utrecht University, Utrecht, Netherlands

[2]Department of Obstetrics and Gynaecology, Erasmus Medical Center, Erasmus University, Rotterdam, Netherlands

Corresponding Author:
Mireille N Bekker, MD, PhD
Department of Obstetrics
University Medical Center
Utrecht University
Lundlaan 6
Utrecht, 3584 EA
Netherlands
Phone: 31 088 75 55555
Email: m.n.bekker-3@umcutrecht.nl

Abstract

Background: Daily monitoring of fetal and maternal conditions in complicated pregnancies leads to recurrent outpatient visits or (prolonged) hospitalization. Alternatives for hospital admissions include home-based monitoring with home visits by professionals or telemonitoring with self-measurements performed by pregnant women and uploaded for in-clinic assessment. For both alternatives, cardiotocography and blood pressure measurement can be performed at home. It is unknown to what extent, for which reasons, and for which pregnancy complications these strategies are used.

Objective: This study aims to assess the current practice and attitudes concerning home-based monitoring (with daily home visits by professionals) and telemonitoring (using devices and the internet for daily self-recorded measurements) in high-risk pregnancies requiring maternal and fetal monitoring in the Netherlands.

Methods: This nationwide cross-sectional study involved sending a web-based survey to the obstetrics departments of all 73 hospitals in the Netherlands to be answered by 1 representative dedicated to pregnancy monitoring per hospital. The primary outcome was the provision of home-based monitoring or telemonitoring using cardiotocography between 1995 and 2018. The survey further addressed perspectives regarding the use of home-based monitoring and telemonitoring, including (contra)indications, advantages, and disadvantages for pregnant women and clinicians.

Results: The response rate for the provision of either home-based monitoring or telemonitoring was 100%. In 2018, 38% (28/73) of centers in the Netherlands offered either home-based monitoring or telemonitoring or both to pregnant women with complications. Home-based monitoring was offered in 26% (19/73) of the centers; telemonitoring, in 23% (17/73); and both in 11% (8/73). Telemonitoring was first offered in 2009, increasing from 4% (3/73) of hospitals in 2014 to 23% (17/73) in 2018. Responses were received from 78% (57/73) of the invited hospitals and analyzed. Of all 17 centers using telemonitoring, 59% (10/17) did not investigate perinatal outcomes, safety, and patient satisfaction prior to implementation. Other (6/17, 35%) telemonitoring centers are participating in an ongoing multicenter randomized clinical trial comparing patient safety, satisfaction, and costs of telemonitoring with standard hospital admission. Home-based monitoring and telemonitoring are provided for a wide range of complications, such as fetal growth restriction, pre-eclampsia, and preterm rupture of membranes. The respondents reported advantages of monitoring from home, such as reduced stress and increased rest for patients, and reduction of admission and possible reduction of costs. The stated barriers included lack of insurance reimbursement and possible technical issues.

Conclusions: Home-based monitoring is provided in 26% (19/73) and telemonitoring, in 23% (17/73) of hospitals in the Netherlands to women with pregnancy complications. Altogether, 38% (28/73) of hospitals offer either home-based monitoring or telemonitoring or both as an alternative to hospital admission. Future research is warranted to assess safety and reimbursement issues before more widespread implementation of this practice.

XSL•FO
RenderX

(JMIR Mhealth Uhealth 2020;8(10):e18966) doi:10.2196/18966

KEYWORDS

mobile health; telemonitoring; pregnancy complications; digital health; telemedicine

Introduction

Pregnancies with complications need close antenatal surveillance. Although 7-10 antenatal consultations are recommended in uncomplicated pregnancies, complications result in recurrent outpatient visits or hospital admission [1]. These complications include fetal growth restriction, pre-eclampsia, and preterm prelabor rupture of membranes (prevalence of 3%-7%, 1%-3%, and 1%-5%, respectively) [2-4]. Daily monitoring with cardiotocography (CTG), blood pressure measurements, and/or urine and blood analysis is recommended in international guidelines to assess maternal and fetal conditions in high-risk pregnancies [5-7]. Ultimately, hospitalization is indicated in up to 11% of all pregnancies, usually extending to delivery and the postpartum period [5-7]. Antenatal admissions pose psychological stress to pregnant women because of separation from family and home, lack of activity, and feelings of uncertainty [8,9]. In addition, hospital admissions increase health care costs and workload, especially in high-income countries that are already experiencing difficulties as a result of professional staff shortage [10].

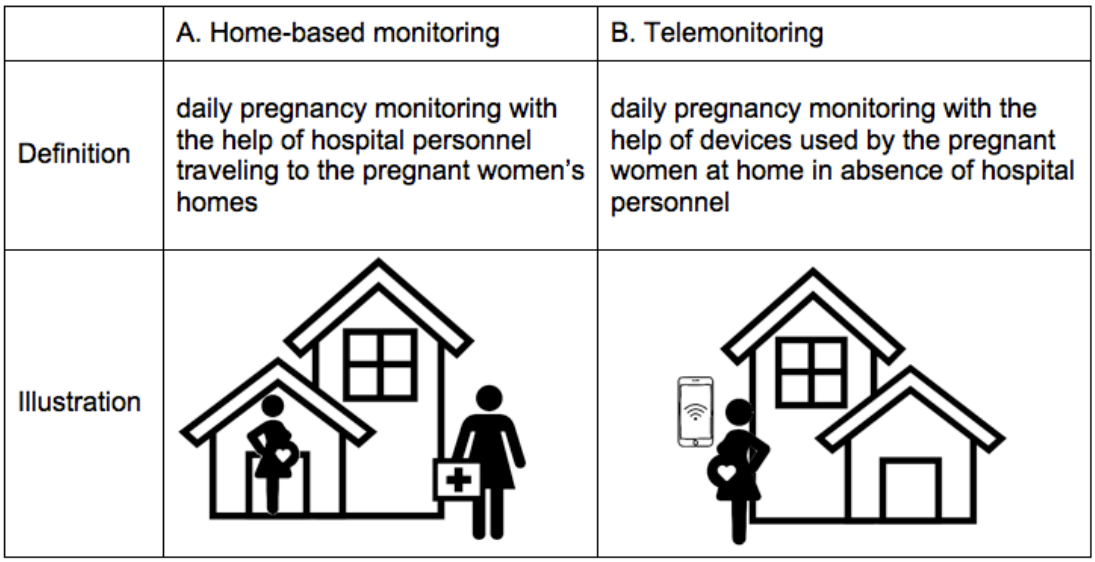
Since 1990, obstetrics departments in the Netherlands have been providing domiciliary care or "home-based monitoring" to women with high-risk pregnancies. As an alternative to clinical admission, home-based monitoring involves daily home visits from hospital-employed midwives or nurses for pregnant women with complications. Medical tests, including CTG, are performed at home, and the results are discussed with a supervising gynecologist (Figure 1A). Multiple randomized trials have proved that home-based monitoring with home visits is feasible and safe with regard to perinatal outcome [11-14]. Although these trials demonstrated satisfactory outcomes for both mother and child, the daily visits were also found to be time-consuming and, therefore, expensive.

The use of digital health for remote monitoring in pregnancy care is increasingly popular, as pregnant women are frequent users of smartphones, the internet, and health apps [15]. *Tele*monitoring is a relatively new approach in high-risk pregnancy and is recognized as an alternative to hospital admission or home-based monitoring with prenatal home visits. After training participants, daily measurements of blood pressure and CTG are self-recorded by the patient at home and sent via Bluetooth or WiFi to a secured digital platform. Using an internet connection, the data are integrated in the electronic patient file (Figure 1B). Patients are contacted by their clinician on a daily basis to discuss the presence of symptoms, tests results, and future management. Multiple telemonitoring platforms for remote CTG have been evaluated in prospective studies, and their feasibility, usability, accuracy of tracings, and acceptability by patients and clinicians are proven [15].

In general, digital health has the potential to improve access to care, disease monitoring, and patient satisfaction while reducing health care costs due to a reduction in visits and admissions. Currently, the clinical evidence for telemonitoring using CTG in complicated pregnancies is too scarce to support hypotheses regarding its effects on perinatal outcome, safety, patient preference, and costs.

A number of hospitals in the Netherlands currently provide either home-based monitoring or telemonitoring or both to women with high-risk pregnancies. It is unknown to what extent, for which reasons, and for which pregnancy complications these strategies are used. This information is relevant for clinicians planning to use a telemonitoring strategy in prenatal care. The aim of this nationwide survey study is to determine the number of hospitals in the Netherlands that provide home-based monitoring or telemonitoring or both, and to identify the current practice of out-of-hospital care in high-risk pregnancies.

Figure 1. Definition and illustration of (A) home-based monitoring and (B) telemonitoring in pregnancy.



XSL•FO
RenderX

Methods

We conducted a nationwide cross-sectional study using a web-based survey amongst obstetrics care professionals. All hospitals with pregnancy and childbirth care departments in the Netherlands (N=73) were invited to participate in our survey. They were asked to appoint one of their obstetrics care professionals dedicated to (remote) pregnancy monitoring as the department representative to answer the questions on behalf of the hospital. After receiving additional information about the purpose of the study, access was provided to our web-based survey. The survey link was sent via email in November 2018 and followed by a maximum of 3 email reminders. Nonrespondents were contacted once more by phone to answer the principal question "Does your center currently offer home-based monitoring or telemonitoring in pregnancy?"

The survey was self-developed and based on expert knowledge of home-based monitoring and telemonitoring in the Netherlands. A professor of obstetrics, a perinatologist, a hospital-based midwife, and 2 researchers, all with extensive experience in home-based monitoring of high-risk pregnancies, were involved in its development. It contained a maximum of 44 questions depending on whether home-based monitoring or telemonitoring was offered. The open and multiple-choice questions addressed 4 domains: (1) Basic demographics of the respondent, (2) home-based monitoring, (3) telemonitoring, and (4) advantages and disadvantages of home-based monitoring and telemonitoring, as perceived by the respondent (see Multimedia Appendix 1).

The survey sought information on the total number of births per year in order to compare the hospitals with reference to size. Regarding the provision of home-based monitoring or telemonitoring, the starting year and, if applicable, year of discontinuation were queried. We defined our study period from 1995 to 2018. Questions regarding indications, management protocols, and (dis)advantages of the strategies were asked bearing in mind the centers' practice for the year 2018. The introduction to our survey defined home-based monitoring as "daily pregnancy monitoring with help of hospital personnel traveling to the pregnant women's homes." Telemonitoring was defined as "daily pregnancy monitoring with the help of devices used by pregnant women at home in the absence of hospital personnel" (Figure 1). Simple descriptive statistics were used to describe the results. No ethical approval was required for this study because actual patients were not involved.

Results

Current Provision of Home-Based Monitoring and Telemonitoring in High-Risk Pregnancy

In 2018, 73 hospitals in the Netherlands provided pregnancy and childbirth care. The principle question, namely "Does your center currently offer home-based monitoring or telemonitoring in pregnancy?" was answered by all 73 invitees, resulting in a response rate of 100%.

In 2018, 26% (19/73) hospitals offered home-based monitoring with home visits by an obstetrics professional (nurse or midwife) for women with high-risk pregnancies. Nationwide, 23% (17/73) of hospitals offered telemonitoring in 2018 to women with high-risk pregnancies (Table 1). Moreover, 11% (8/73) of centers reported offering both home-based monitoring with home visits and telemonitoring to their patients.

In obstetrics departments with ≤1000 births per year, home-based monitoring and telemonitoring is limited to 0 and 1 centers, respectively. As for the different types of hospitals, 8 of 9 Dutch tertiary care hospitals with a neonatal intensive care unit facility currently offer home-based monitoring or telemonitoring or both for high-risk pregnancy management. The geographic distribution of hospitals with home-based monitoring and telemonitoring is displayed in Figure 2 for all 12 provinces of the Netherlands.

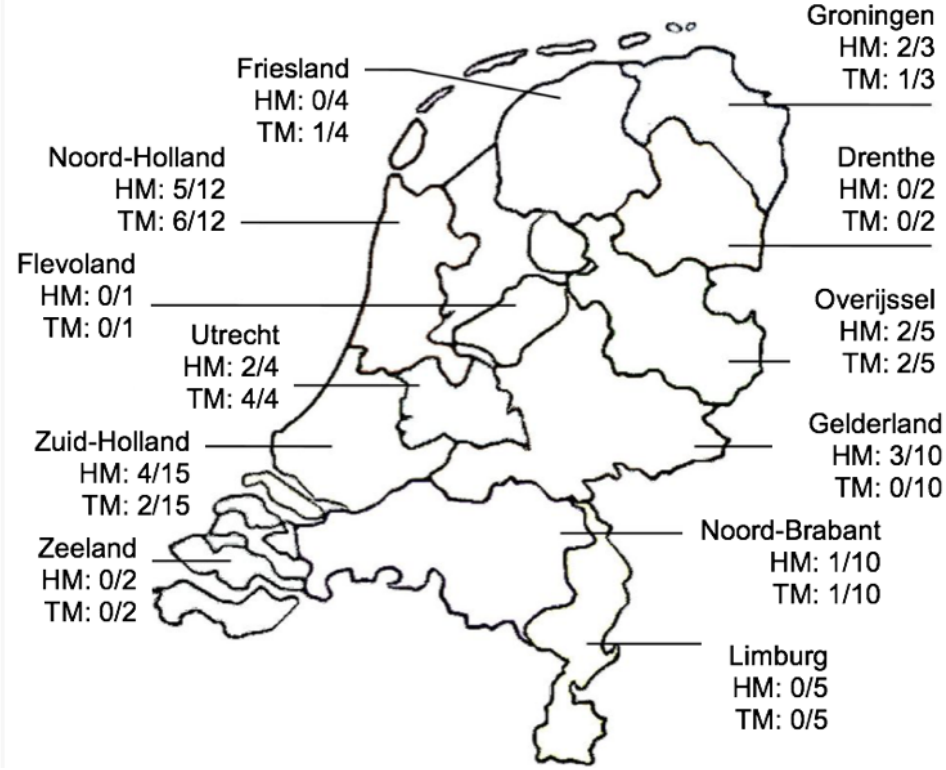
Table 1. Number of hospitals offering home-based monitoring and telemonitoring for high-risk pregnancies in 2018 in relation to the number of births per hospital per year.

Number births per hospital per year	Number of hospitals (N=73), n (%)[a]	Home-based monitoring (n=19), n (%)[a]	Telemonitoring (n=17), n (%)[a]
0-1000	15 (21)	0 (0)	1 (7)
1001-2000	35 (48)	9 (26)	8 (23)
2001-3000	21 (29)	9 (43)	6 (29)
>3000	2 (2)	1 (50)	2 (100)

[a]Percentages in these columns are based on the number of hospitals in each row.

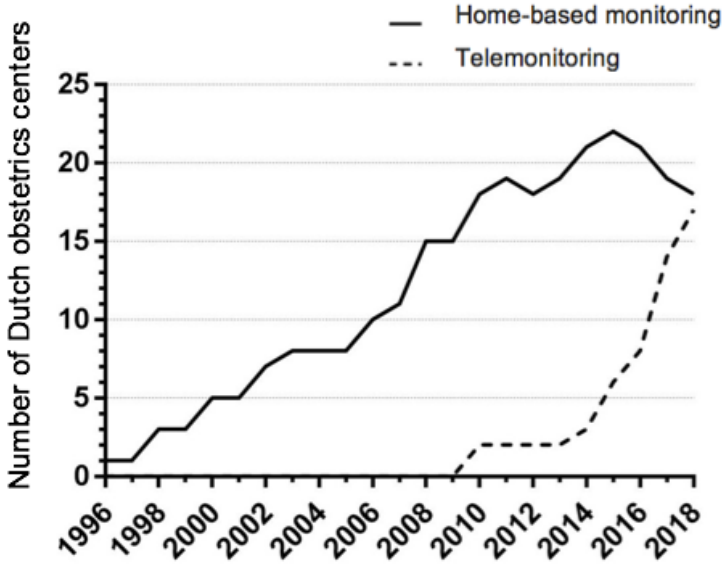
XSL•FO
RenderX

Figure 2. Geographic distribution of home-based monitoring and telemonitoring in 12 provinces of the Netherlands (N=73). HM: hospitals with home-based monitoring/all hospitals in this province; TM: hospitals with telemonitoring/all hospitals in this province.



For the studied period of 1995-2018, the trend line in Figure 3 shows that home-based monitoring in pregnancy has been offered since the mid-1990s. Most of these centers continued offering daily home visits over a longer period of time, reaching a peak in 2015. After the introduction of pregnancy telemonitoring in 2009, the trend line for telemonitoring use shows a steep increase from 2014 onwards, from 4% (3/73) of centers in 2014 to 23% (17/73) in 2018. This increase in number is accompanied by a slight drop in home-based monitoring provision.

Figure 3. Trend graph of obstetrics departments offering home-based monitoring and telemonitoring for high-risk pregnancies in the Netherlands.



Survey Results

Respondents' Characteristics

Of the total 73 invited hospitals, 57 participated in the web-based survey (response rate 78%). Of these respondents, 45% (26/57) worked in a teaching hospital, 39% (22/57) served in a nonteaching hospital, and 16% (9/57) worked in an tertiary care hospital with a neonatal intensive care unit. Moreover, 14% (8/57) reported 0-1000 births per year, 51% (29/57) had 1001-2000 births per year, 32% (18/57) recorded 2001-3000 births per year, and 3% (2/57) reported over 3000 births per year.

XSL•FO
RenderX

Declining Trend of Home-Based Monitoring

The results showed that 11% (6/57) of centers offered home-based monitoring in the years between 1995 and 2018 but stopped performing pregnancy monitoring with home visits. The median number of years of home-based monitoring provision was 7.5 years (range 2-18 years). Several reasons were provided for the discontinuation, such as very few possible candidates (3/6, 50%), problems with staff capacity (3/6, 50%), financial capacity issues (2/6, 33%), and switching over to telemonitoring without home visits (2/6, 33%).

Moreover, 42% (8/19) of hospitals with home-based monitoring considered switching to telemonitoring, stating that the latter provides higher patient satisfaction and does not require hospital staff to visit patients at home. Interestingly, 16% (3/19) of hospitals providing home-based monitoring did not change to telemonitoring because they are satisfied with their current home-based monitoring strategy. They stated that telemonitoring allows neither daily direct clinical assessment of the patient by a nurse/midwife nor the ability to monitor twin pregnancies.

Evaluation of Use

In 63% (12/19) of hospitals offering home-based monitoring, implementation of such monitoring was not preceded by a center-specific evaluation phase. However, home-based monitoring in these centers started mainly after the publication of the findings of 2 Dutch trials, which concluded positively about its patient safety and effects on satisfaction of care [7,8].

As for telemonitoring, 35% (6/17) of centers are participating in a multicenter randomized controlled trial comparing clinical hospital admission with telemonitoring in pregnancies requiring daily fetal monitoring. The aim of this trial is to compare patient safety, user satisfaction, and cost-effectiveness; its protocol can be found elsewhere [16]. The remaining 59% (10/17) of centers reported they neither participated in nor started evaluation of use of this novel strategy with daily self-measurements prior to implementation in complicated pregnancies in their centers.

Indications and Management in Home-Based Pregnancy Monitoring

Responding centers implementing either home-based monitoring or telemonitoring reported similar lists of pregnancy complications, which they considered eligible for daily monitoring outside their hospital (Table 2). Both fetal growth restriction and preterm rupture of membranes are considered eligible for home-based monitoring as well as telemonitoring in every center.

All (18/18, 100%) home-based monitoring centers reported that at home, midwives or nurses measure patients' blood pressure levels and perform CTGs during their visits. Fetal condition monitoring of both singleton and twin pregnancies using CTG is possible in 83% (15/18) of centers. Additionally, urine analysis (13/18, 72%), venous blood sampling (12/18, 67%), and medication administration (4/18, 22%) can be performed by professionals at home. This is in contrast to telemonitoring centers, where only blood pressure monitoring and CTG are performed by patients themselves at home.

Hospitals with either home-based monitoring or telemonitoring also reported on the following patient-specific contraindications for home-based monitoring: inability to follow instructions or difficulty in understanding the system (26/28, 93%), long home-to-hospital distance (25/28, 89%), existing antepartum hemorrhage (20/28, 71%), and vulnerable home situation or social issues experienced by the patient (14/28, 50%). Other mentioned general contraindications were gestational age<25 weeks and preterm premature rupture of membranes without engaged fetal head or breech. To ensure the safety of the patients by minimizing travel time if complications occur, the respondents clarified that patients must reside within a distance of 30-35 km from their hospital.

Table 2. Indications for home-based monitoring (n=19) and telemonitoring (n=17) in high-risk pregnancies.

Indications	Home-based monitoring centers, n (%)	Telemonitoring centers, n (%)
Fetal growth restriction	19 (100)	17 (100)
Preterm premature rupture of membranes	19 (100)	17 (100)
Prolonged prelabor rupture of membranes at term	5 (26)	2 (12)
Isolated oligohydramnios	10 (53)	4 (24)
Reduced fetal movement	15 (79)	15 (88)
Fetal anomalies requiring fetal monitoring	9 (47)	3 (18)
(Adverse) Obstetrics patient history[a]	16 (84)	15 (88)
Hypertensive pregnancy disorders	15 (79)	10 (59)
Cholestasis of pregnancy	14 (74)	5 (29)
Other maternal comorbidities[b]	11 (58)	4 (24)
Social or psychological distress	11 (58)	5 (29)

[a]For instance, intrauterine fetal death in a previous pregnancy.

[b]For instance, (gestational) diabetes mellitus, kidney disease, and cardiac disease requiring maternal monitoring.

XSL•FO
RenderX

Reported Advantages and Disadvantages of Home-Based Monitoring and Telemonitoring

The most frequently addressed advantages of home-based monitoring and telemonitoring for the patients, as perceived by the respondents, include more patient comfort and less emotional burden of hospitalization for the patient as they continue with daily (family) life and activities as much as possible. Other frequently mentioned advantages are summarized in Table 3.

Possible disadvantages of home-monitoring and telemonitoring for the patient include the possibility of a delay in providing help in case of an emergency or acute problem, because the patient is not physically present in the hospital. Technical and security issues regarding the devices are also mentioned (Table 3).

The respondents reported a number of perceived benefits of home-based monitoring and telemonitoring for the health care provider, the most important being the reduction in the number of admissions, which in turn may lower health care costs (45/57, 79%) and reduce the burden on hospital personnel (26/57, 46%). The most common disadvantages of home-based monitoring and telemonitoring for the clinicians are costs and reimbursement (38/57, 67%) and inability to conduct direct patient assessments (18/57, 31%). For home-based monitoring specifically, the most serious disadvantage was lack of sufficient obstetrics personnel to make home visits (22/57, 39%).

Table 3. Advantages and disadvantages of home-based monitoring and telemonitoring for patients according to the respondents (n=57).

Advantages or disadvantages	Values
Advantages, n (%)	
Improved patient comfort	40 (70)
Reduced (emotional) burden of admission	35 (61)
Reduced stress/more rest	25 (44)
Better patient autonomy	21 (37)
Higher patient satisfaction	8 (14)
Higher patient safety	7 (12)
Reduced over-medicalization during pregnancy	2 (4)
Disadvantages, n (%)	
Possible delay in providing help during emergencies or acute problems	38 (67)
No direct communication with the consulting gynecologist	18 (31)
Patients' inability to conduct CTG[a] at home	13 (23)
Technical issues	17 (30)
Inability to follow instructions	12 (21)

[a]CTG: cardiotocography.

Number of High-Risk Pregnant Women Managed From Home

All (19/19, 100%) of home-based monitoring centers reported monitoring a minimum of 745 to a maximum of 1140 patients with a singleton pregnancy via home visits in 2018.

All of the telemonitoring centers (17/17, 100%) reported monitoring a minimum of 400 to a maximum of 725 patients with a singleton pregnancy via remote monitoring devices in 2018.

Discussion

Main Findings

Our survey results show the current practice in the Netherlands regarding the use of home-based monitoring and telemonitoring in high-risk pregnancies. In 1995, pregnancy monitoring with daily home visits was available in only a few obstetrics hospitals; currently, it is used by 26% (19/73) of all hospitals in the Netherlands. The last 5 years have witnessed a steep increase in the provision of telemonitoring; as of 2018, it was used in 23% (17/73) of obstetrics departments. Furthermore, almost half (8/19, 42%) of the hospitals with home-based monitoring considered switching to telemonitoring using self-measurement of fetal and maternal parameters. For the telemonitoring centers, 59% (10/17) did not evaluate the use of this digital health strategy with daily self-measurements prior to implementation in their centers. Moreover, 35% (6/17) of centers are currently participating in an ongoing trial to compare traditional hospital admission and telemonitoring with regard to patient safety, satisfaction, and costs.

In 2018, 1145-1865 pregnant women were monitored from home with home visits or telemonitoring after diagnosis of 1 or more complications.

Strengths and Limitations

Our study involved a nationwide survey with a high response rate and included both secondary and tertiary referral centers, teaching and nonteaching centers, and a wide size range of units according to annual birth numbers. The responses to the survey depended on the voluntary participation of the invited hospitals, which could have led to selection bias. Furthermore, the

XSL•FO
RenderX

collected data were self-reported and hence subjective. Some of the results on the impact of remote monitoring were based on estimations by the respondents, which may limit the validity of the conclusions. Evaluations of the characteristics of pregnant women, relevant clinical outcomes (including safety), and user experiences are critical for future health care improvements using mobile health monitoring. However, this study was not devised to evaluate these outcomes, which might be considered a limitation.

Interpretation

The level of application of digital health in prenatal care is evident, with pregnancy telemonitoring being one of the most promising additions to new care models [9,15,17]. The respondents of our survey identified important (perceived) advantages of telemonitoring: improved patient-friendly care in response to their needs, increased patient satisfaction and autonomy, and reduced over-medicalization. These results are in line with those of previous research on patient experiences with digital health [15,18,19]. Furthermore, obstetrics care professionals underscore the importance of digital health in pregnancy care. For instance, a survey study conducted in Belgium concluded that 80% (28/35) of midwives and 67% (6/9) of obstetricians who used remote blood pressure monitoring in pregnancy perceived digital technologies to be an important component of prenatal monitoring [20]. Moreover, a survey amongst 89 German physicians concluded that nearly 70% considered apps for pregnancy monitoring reasonable [21]. Other reported advantages in favor of telemonitoring are the reductions in the number of admissions and the burden on hospital personnel [18,19]. Staff shortages are also driving a shift from hospital- to home-based care.

In the Netherlands, approximately 170,000 children are delivered per year from both uncomplicated and complicated pregnancies. We estimated earlier that 11% of pregnant women need antenatal hospital admission because of complications, which equals to 18,700 women yearly [2-4]. Using our respondents' results, we calculated that 1145-1865 pregnant women were monitored from home in 2018. This number indicates that roughly 6%-10% of antenatal hospital admissions were replaced by home-based monitoring or telemonitoring in 2018. Although exact numbers on the length of hospitalization during high-risk pregnancies are lacking, we can use these values to estimate the possible impact of home-based and telemonitoring on admissions during pregnancy. If home-based monitoring or telemonitoring services in pregnancy were to replace 5 days or nights of hospitalization per pregnant woman, the number of hospital admission days would reduce by 5,725-9,325 annually.

Studies on telemonitoring implementation using patient-recorded daily CTG are limited. Despite the inadequate knowledge of the effects of pregnancy telemonitoring on perinatal outcomes, patient experiences, and cost-effectiveness, this study shows that telemonitoring is becoming increasingly popular in the Netherlands. Although not mentioned by our respondents, legal concerns such as third party control and use of data limit the widespread use of digital health interventions. In the Netherlands, external companies providing devices, software, and storage of patient data for telemonitoring must provide certain data security assurances. Evidence from clinical trials and health technology assessments will help to better estimate the exact budgetary impact from several different (ie, societal, insurance, and hospital) perspectives. The costs involved in the development, use, and maintenance of the devices, as well as the manner in which they are imbedded in current practice, should also be calculated to assess the added value of pregnancy telemonitoring. Our survey respondents reported challenges with reimbursement, since no insurance coverage for pregnancy telemonitoring exists in the Netherlands. Financial issues were another primary reason mentioned by our respondents (especially the smaller obstetrics units) to not offer either home-based monitoring or telemonitoring. It is well known that insurance companies only cover well-researched digital health interventions in accordance with their economic evaluations [22,23]. To compare daily telemonitoring at home with traditional hospital care for complicated pregnancies, a multicenter randomized controlled trial, HOTEL (HOspital admission versus TELemonitoring in high risk pregnancy), is currently recruiting participants in 6 Dutch hospitals [16]. This trial aims to compare both strategies with regard to perinatal outcomes, patient satisfaction, and cost-effectiveness.

Recommendations for Research and Practice

Our survey study provides information about the current practice and trends in the Netherlands regarding home-based monitoring and telemonitoring in perinatal care. More detailed information on the barriers and facilitators, from both the patients' and health care providers' viewpoints, may help develop other innovative strategies in perinatal care. However, evidence on the medical outcomes and patient safety with telemonitoring is still lacking, and more information is required before implementing such innovations in the target population. We must expand our knowledge of these forms of care in order to continue moving forward with digital health innovations. Consensus on the implementation and research agenda can pave the road to the widespread use of digital health services. Additional trials and stakeholders' views of digital health care are needed to develop insurance reimbursement systems for such remote monitoring innovations in pregnancy.

Conclusion

In 2018, 26% (19/73) of hospitals in the Netherlands offered home-based monitoring and 23% (17/73) offered telemonitoring to their patients with pregnancy complications. These increasingly popular forms of home-based care allow an increasing number of pregnant women in need of daily monitoring to stay at home and avoid hospital admission. Additionally, the survey respondents shed light on multiple possible advantages and disadvantages of home-based monitoring and telemonitoring in pregnancy. These results can contribute to future evaluations of digital innovations in pregnancy care, as further research on their safety, experience, and cost-effectiveness is warranted before more widespread implementation.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of the survey used in this study.
[DOCX File , 18 KB - mhealth_v8i10e18966_app1.docx]

References

1. National Institute of Clinical Excellence (NICE). Antenatal care for uncomplicated pregnancies - Clinical guideline [CG62]. 2008 Mar 26. URL: https://www.nice.org.uk/guidance/cg62 [accessed 2018-11-15]
2. Romo A, Carceller R, Tobajas J. Intrauterine growth retardation (IUGR): epidemiology and etiology. Pediatr Endocrinol Rev 2009 Feb;6 Suppl 3:332-336. [Medline: 19404231]
3. Lamain-de Ruiter M, Kwee A, Naaktgeboren CA, Louhanepessy RD, De Groot I, Evers IM, et al. External validation of prognostic models for preeclampsia in a Dutch multicenter prospective cohort. Hypertens Pregnancy 2019 May;38(2):78-88. [doi: 10.1080/10641955.2019.1584210] [Medline: 30892981]
4. van der Ham DP, Vijgen S, Nijhuis J, van Beek JJ, Opmeer BC, Mulder ALM, PPROMEXIL trial group. Induction of labor versus expectant management in women with preterm prelabor rupture of membranes between 34 and 37 weeks: a randomized controlled trial. PLoS Med 2012;9(4):e1001208 [FREE Full text] [doi: 10.1371/journal.pmed.1001208] [Medline: 22545024]
5. National National Institute of Clinical Excellence (NICE). Hypertension in pregnancy: diagnosis and management - Clinical guideline [CG107]. 2010 Aug 25. URL: https://www.nice.org.uk/guidance/cg107 [accessed 2018-11-15]
6. Royal College of Obstetricians and Gynaecologists. Green-top Guideline No 31. Small-for-Gestational-Age Fetus, Investigation and Management (Green-top Guideline No. 31). 2013 Mar 23. URL: https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg31/ [accessed 2018-11-15]
7. National Institute of Clinical Excellence (NICE). Preterm labour and birth - NICE guideline [NG25]. 2015 Oct 20. URL: https://www.nice.org.uk/guidance/ng25 [accessed 2018-11-15]
8. Gourounti C, Karpathiotaki N, Vaslamatzis G. Psychosocial stress in high risk pregnancy. Int Arch Med 2015;8(95). [doi: 10.3823/1694]
9. van den Heuvel JFM, Teunis CJ, Franx A, Crombag NMTH, Bekker MN. Home-based telemonitoring versus hospital admission in high risk pregnancies: a qualitative study on women's experiences. BMC Pregnancy Childbirth 2020 Feb 04;20(1):77 [FREE Full text] [doi: 10.1186/s12884-020-2779-4] [Medline: 32019499]
10. Huynh L, McCoy M, Law A, Tran KN, Knuth S, Lefebvre P, et al. Systematic literature review of the costs of pregnancy in the US. Pharmacoeconomics 2013 Nov;31(11):1005-1030. [doi: 10.1007/s40273-013-0096-8] [Medline: 24158771]
11. Iedema-Kuiper H. Geintegreerde thuiszorg bij risicozwangeren, een gerandomiseerd onderzoek ; (with a summary in English). Utrecht, the Netherlands: Universiteit Utrecht, Faculteit Geneeskunde; 1996.
12. Monincx WM, Zondervan HA, Birnie E, Ris M, Bossuyt PM. High risk pregnancy monitored antenatally at home. Eur J Obstet Gynecol Reprod Biol 1997 Dec;75(2):147-153. [doi: 10.1016/s0301-2115(97)00122-x] [Medline: 9447367]
13. Brooten D, Youngblut JM, Brown L, Finkler SA, Neff DF, Madigan E. A randomized trial of nurse specialist home care for women with high-risk pregnancies: outcomes and costs. Am J Manag Care 2001 Aug;7(8):793-803 [FREE Full text] [Medline: 11519238]
14. Harrison MJ, Kushner KE, Benzies K, Kimak C, Jacobs P, Mitchell BF. In-home nursing care for women with high-risk pregnancies: outcomes and cost. Obstet Gynecol 2001 Jun;97(6):982-987. [doi: 10.1016/s0029-7844(01)01360-6] [Medline: 11384707]
15. van den Heuvel JF, Groenhof TK, Veerbeek JH, van Solinge WW, Lely AT, Franx A, et al. eHealth as the Next-Generation Perinatal Care: An Overview of the Literature. J Med Internet Res 2018 Jun 05;20(6):e202 [FREE Full text] [doi: 10.2196/jmir.9262] [Medline: 29871855]
16. van den Heuvel JFM, Ganzevoort W, De Haan-Jebbink JM, van der Ham DP, Deurloo KL, Seeber L, et al. HOspital care versus TELemonitoring in high-risk pregnancy (HOTEL): study protocol for a multicentre non-inferiority randomised controlled trial. BMJ Open 2019 Oct 28;9(10):e031700. [doi: 10.1136/bmjopen-2019-031700] [Medline: 31662396]
17. Lanssens D, Vandenberk T, Thijs IM, Grieten L, Gyselaers W. Effectiveness of Telemonitoring in Obstetrics: Scoping Review. J Med Internet Res 2017 Sep 27;19(9):e327 [FREE Full text] [doi: 10.2196/jmir.7266] [Medline: 28954715]
18. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 03;7(8):e016242. [doi: 10.1136/bmjopen-2017-016242]
19. O'Brien E, Rauf Z, Alfirevic Z, Lavender T. Women's experiences of outpatient induction of labour with remote continuous monitoring. Midwifery 2013 Apr;29(4):325-331. [doi: 10.1016/j.midw.2012.01.014] [Medline: 23159160]
20. Lanssens D, Vandenberk T, Lodewijckx J, Peeters T, Storms V, Thijs IM, et al. Midwives', Obstetricians', and Recently Delivered Mothers' Perceptions of Remote Monitoring for Prenatal Care: Retrospective Survey. J Med Internet Res 2019 Apr 15;21(4):e10887 [FREE Full text] [doi: 10.2196/10887] [Medline: 30985286]

XSL•FO
RenderX

21. Grassl N, Nees J, Schramm K, Spratte J, Sohn C, Schott TC, et al. A Web-Based Survey Assessing the Attitudes of Health Care Professionals in Germany Toward the Use of Telemedicine in Pregnancy Monitoring: Cross-Sectional Study. JMIR Mhealth Uhealth 2018 Aug 08;6(8):e10063 [FREE Full text] [doi: 10.2196/10063] [Medline: 30089606]
22. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
23. Dorsey ER, Topol EJ. State of Telehealth. N Engl J Med 2016 Jul 14;375(2):154-161. [doi: 10.1056/NEJMra1601705] [Medline: 27410924]

Abbreviations

CTG: cardiotocography
HOTEL: HOspital admission versus TELemonitoring in high risk pregnancy

Edited by G Eysenbach; submitted 30.03.20; peer-reviewed by D Lanssens, E van der Velde, JB Park; comments to author 04.05.20; revised version received 29.05.20; accepted 03.09.20; published 28.10.20.

Please cite as:
van den Heuvel JFM, Ayubi S, Franx A, Bekker MN
Home-Based Monitoring and Telemonitoring of Complicated Pregnancies: Nationwide Cross-Sectional Survey of Current Practice in the Netherlands
JMIR Mhealth Uhealth 2020;8(10):e18966
URL: http://mhealth.jmir.org/2020/10/e18966/
doi:10.2196/18966
PMID:33112250

©Josephus F M van den Heuvel, Samira Ayubi, Arie Franx, Mireille N Bekker. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 28.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Attitudes and Expectations of Health Care Professionals Toward App-Based Therapy in Patients with Osteoarthritis of the Hip or Knee: Questionnaire Study

Johanna Theresia Biebl[1], MD; Stephan Huber[2], MD; Marzena Rykala[1], BSc; Eduard Kraft[1], MD; Andreas Lorenz[1], MD

[1]Department of Orthopaedics, Physical Medicine and Rehabilitation, University Hospital, LMU Munich, Munich, Germany

[2]Kaia Health GmbH, Munich, Germany

Corresponding Author:
Johanna Theresia Biebl, MD
Department of Orthopaedics, Physical Medicine and Rehabilitation
University Hospital, LMU Munich
Marchioninistr. 15
Munich, 81377
Germany
Phone: 49 89 4400 74070
Email: johanna.biebl@med.uni-muenchen.de

Abstract

Background: The use of mobile health (mHealth) apps is becoming increasingly widespread. However, little is known about the attitudes, expectations, and basic acceptance of health care professionals toward such treatment options. As physical activity and behavior modification are crucial in osteoarthritis management, app-based therapy could be particularly useful for the self-management of this condition.

Objective: The objective of the study was to determine the expectations and attitudes of medical professionals toward app-based therapy for osteoarthritis of the hip or knee.

Methods: Health care professionals attending a rehabilitation congress and employees of a university hospital were asked to fill out a questionnaire consisting of 16 items. A total of 240 questionnaires were distributed.

Results: A total of 127 participants completed the questionnaire. At 95.3% (121/127), the approval rate for app-based therapy for patients with osteoarthritis of the hip or knee was very high. Regarding possible concerns, aspects related to data protection and privacy were primarily mentioned (41/127, 32.3%). Regarding potential content, educational units, physiotherapeutic exercise modules, and practices based on motivation psychology were all met with broad approval.

Conclusions: The study showed a high acceptance of app-based therapy for osteoarthritis, indicating a huge potential of this form of treatment to be applied, prescribed, and recommended by medical professionals. It was widely accepted that the content should reflect a multimodal therapy approach.

(JMIR Mhealth Uhealth 2020;8(10):e21704) doi:10.2196/21704

KEYWORDS

mobile health; digital health; self-management; osteoarthritis; smartphone; patient education; exercise therapy

Introduction

According to the World Health Organization, mobile health (mHealth) refers to medical procedures in private and public health care that are applied by means of mobile devices using various technologies [1]. Mobile devices such as smartphones are widely used in the population and have a large number of sensors that can measure vital signs and other health-related data and display patients' progress [2]. Combining sensor data with actively provided information by users and interaction with health care professionals (HCPs) opens up new possibilities for diagnosis and intervention [3].

Osteoarthritis (OA) is the most common joint disease and can lead to severe pain, impaired physical activity, and severely restricted health-related quality of life [4]. The incidence and prevalence of OA will continue to rise in the future due to an

XSL•FO
RenderX

aging society [5]. The importance of patient education and physical exercise as part of therapy is undisputed [6]. In their systematic review, Hagen et al [7] showed that in the field of community-based care, only 38.8% of patients received the recommendation to perform physical exercise or a corresponding prescription. Patient education and advice on self-management strategies were offered to only 35.4% of the patients. Thus, there is a discrepancy between accepted recommendations and the reality of care for patients with OA.

Educational content and suitable exercises could also be taught using an app. It has been shown that health apps are well suited for implementing sustainable behavior changes in the daily life of chronically ill patients [8]. App-based therapeutic options are already available for various chronic conditions, including insomnia, diabetes, chronic knee pain, and low back pain [9-13]. Many of these apps contain communication possibilities and offer sources of information and options for documentation, such as diaries. In addition, planning tools such as appointment reminders and medication schedules are intended to improve adherence to therapy. Aspects that also play a role in the management of OA of the hip and especially of the knee joint, such as weight reduction, have already been successfully addressed in other contexts using app-based approaches [14,15]. Regarding telemedical care for OA, a Swedish research group has translated a conventional OA self-management program—the Management of Patients with Osteoarthritis (BOA) program—into a digital form called Joint Academy. The BOA program was developed on the basis of existing evidence, national and international treatment guidelines, and patient interviews [16]. In the digital form, there is a platform for patients that offers exercises, physiotherapeutic counselling, support from other affected people, and educational content [17]. Participants who used the Joint Academy platform approximately 5 days a week showed an improvement in pain and physical functioning [18]. Another existing digital therapy is the so-called Hinge Health program designed for patients with chronic knee pain, including patients with OA. Among other components, this program offers active exercises in which patients wear portable bands with motion sensors, allowing feedback on their exercise performance [11,19]. In their randomized controlled trial, Mecklenburg et al [11] detected that patients with chronic knee pain who were treated with the Hinge Health program for 12 weeks had significantly better results in terms of pain, physical functioning, surgery, and understanding of the disease than a control group.

Because app-based therapy is a novel technique, interest in the acceptance and expectations of medical professionals regarding app-based therapy is growing. For instance, an Australian study [20] found that approximately two-thirds of the participating general practitioners used apps themselves within their professional activities, and approximately one-half of the respondents recommended the use of apps to their patients. In another study [21], interviews with primary care physicians identified barriers and facilitators for the implementation of apps for the self-management of diabetes. Moreover, there has been research regarding the attitudes of HCPs toward app-based therapy for depression. While only 21.1% had used app-based therapy with their patients before, 66.0% believed that outcomes would improve if apps were integrated into the treatment of depression [22]. Kessel et al [23] conducted an online survey specifically to evaluate the attitudes and expectations of HCPs regarding the use of telemedicine and apps in the field of oncology, and they detected a broad overall support for these forms of care: 88.9% of respondents considered telemedicine to be useful and 84.3% were in favor of an oncological app in addition to standard care. However, to our knowledge, expectations of HCPs regarding app-based therapy for patients with OA have not yet been recorded in a structured way. In the field of musculoskeletal diseases, a recent review by Najm et al [24] showed that the involvement of physicians and other medical professionals in the development and design of apps has been low to date and that their increased participation would be preferable.

Under these circumstances, this study aimed to determine the expectations and attitudes of medical professionals toward app-based therapy for OA of the hip or knee joint. Based on the results, an app is to be developed that meets the expectations of potential mediators of the app and takes into account their clinical experience.

Methods

The study was approved by the Ethics Committee of Ludwig Maximilian University of Munich (LMU Munich), Munich, Germany (reference number 19-627). Written informed consent was obtained from all participants before the survey began.

Study Design

A questionnaire with 16 main items was developed based on recommendations by Langbecker et al [25]. After literature research, we conducted interviews with health experts from different professions and specialties. The collected information was categorized and structured. Based on these data, national guidelines [26,27], and the care standards of our university hospital (LMU Munich), the questionnaire's content was defined by an interdisciplinary team of physicians from various disciplines, psychologists, physiotherapists, and persons knowledgeable in the development of medical mobile apps. The questionnaire was then pretested on a collective of HCPs with regard to comprehensibility and clarity.

Of the 16 questions, 11 were closed and 5 were semiopen. Multiple answers were possible. Two of the main items consisted of 12 and 5 subitems, respectively, each of which was to be assessed on a 5-point Likert scale (1=not useful, 2=rather not useful, 3=partially useful, 4=rather useful, 5=useful). One of the items was used to collect personal data and consisted of four subitems. For some questions, it was possible to specify "no comment" as the answer.

The questions related to possible advantages and disadvantages of the app-based therapy, possible educational content, meaningful exercises, and possible problem areas, as well as to the idea of embedding the app in existing technical systems and possible connections (eg, to so-called wearables—devices that can be worn on the body and use computer technologies).

XSL•FO
RenderX

The participants were asked to indicate their gender, length of professional experience, occupation, and field of activity. Professional experience was categorized in 5-year steps (less than 5 years, 5-10 years, 10-15 years, 15-20 years, 20-25 years, 25-30 years, and more than 30 years).

The questionnaire survey was conducted in an anonymous form. The questionnaires (see Multimedia Appendix 1) were handed out to employees at the University Hospital, LMU Munich, and to attendees at a rehabilitation congress at the same institution. Inclusion criteria were having a self-reported degree in a regulated medical profession and age over 18 years. A total of 240 questionnaires were handed out. No incentives were offered for participation.

In order to avoid a selection bias toward individuals with higher technological affinity, the survey was deliberately distributed as a paper questionnaire. Nevertheless, we have based the presentation of the results, as far as possible, on the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [28] to comply as closely as possible with standards in this field of research.

Statistical Analysis

The data were analyzed using descriptive statistics. If no information was provided for a question, it was taken into account when calculating the percentages and is indicated accordingly. The number of participants who had answered the respective question is shown in parentheses.

In order to detect possible correlations between years of professional experience and approval of app-based therapy, several contingency tables were analyzed. First, the levels of work experience were dichotomized with 6 different cutoffs of years of professional experience, and 6 2x2 contingency tables were built. Second, the level of work experience was dichotomized as individual groups of professional experience versus all other experience levels, and 6 more 2x2 tables were created. All contingency tables were analyzed using the Fisher exact test. Statistical analysis was conducted using SPSS software (version 21.0; IBM Corp).

Results

A total of 127 HCPs submitted completed questionnaires (response rate 52.9%). The characteristics of the sample are listed in Table 1.

In response to the question of whether they would generally recommend an app-based therapy to their patients for the treatment of OA of the knee or hip, 89.0% (105/118) replied yes. Figure 1 provides an illustration of the answer to this question divided according to patients' occupational groups. There was no statistically significant difference in participants' attitude toward recommendation of an app-based therapy based on their years of professional experience (Table 2).

XSL•FO
RenderX

Table 1. Sample characteristics (N=127).

Characteristics	n (%)
Gender	
Female	78 (61.4)
Male	49 (38.6)
Professional experience (years)	
Less than 5	36 (28.3)
5-10	21 (16.5)
10-15	17 (13.4)
15-20	12 (9.4)
20-25	13 (10.2)
25-30	12 (9.4)
More than 30	16 (12.6)
Occupation/training	
Physicians	43 (33.9)
Hospital sector	24 (18.9)
Outpatient sector	13 (10.2)
Medical activity in other areas	6 (4.7)
Nursing staff	3 (2.4)
Therapeutic occupations	44 (34.6)
Occupational therapists	6 (4.7)
Massage therapists	3 (2.4)
Physiotherapists	33 (26.0)
Speech therapists	2 (1.6)
Health care management assistants	5 (3.9)
Psychologists/psychotherapists	6 (4.7)
Medical students	9 (7.1)
Other health care professions	17 (13.3)
Field of activity	
Surgical medicine	11 (8.7)
Conservative medicine	84 (66.2)
Both conservative and surgical medicine	14 (11.0)
No specification provided	18 (14.2)

XSL•FO
RenderX

Figure 1. General approval of app-based therapy for osteoarthritis categorized according to profession.

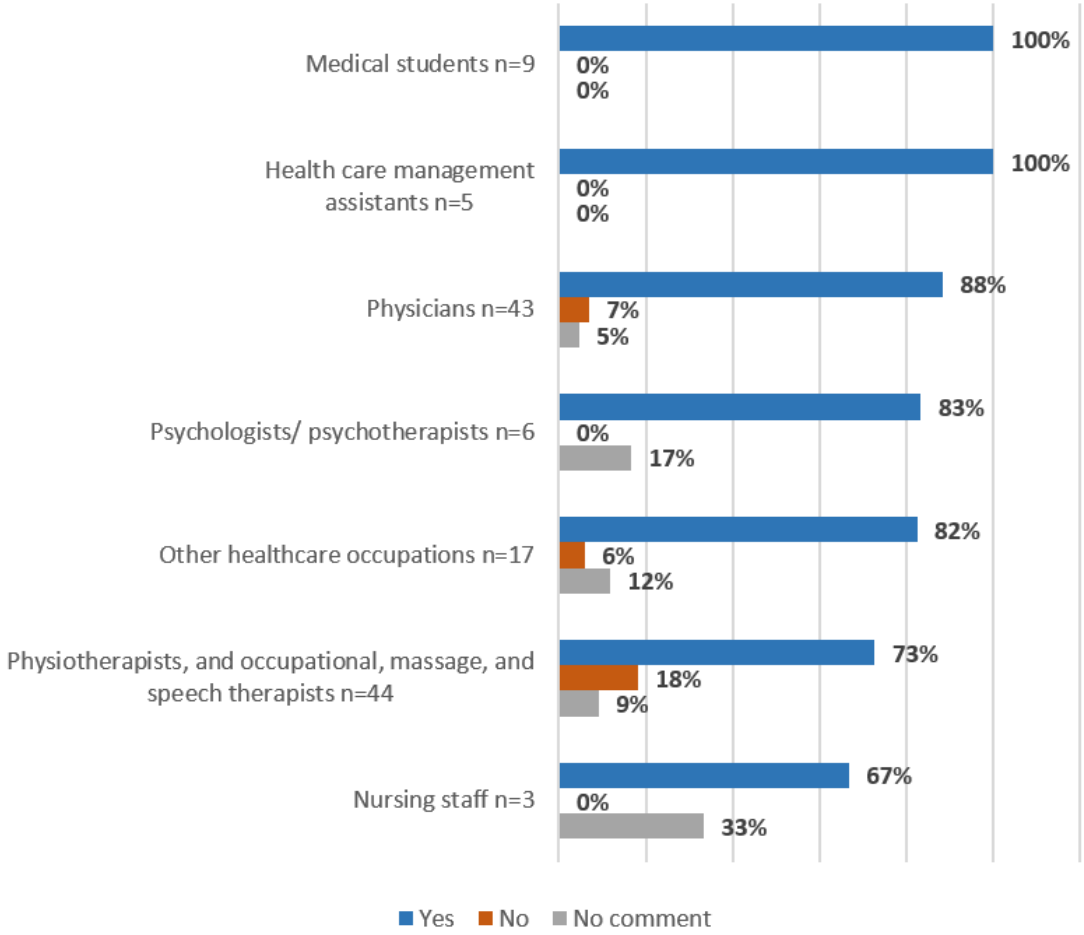


Table 2. Differences in approval of app-based therapy according to level of professional experience. *P* values are provided for comparison of each group to all other groups of professional experience.

Levels of professional experience (years)	*P* value[a]
<5	.51
5-10	>.99
10-15	>.99
15-20	>.99
20-25	.13
25-30	>.99
>30	.38
<10	.39
<15	.24
<20	.39
<25	.49

[a]Calculated using two-sided Fisher exact test.

The specification that an app should be used in addition to conventional therapy was accepted by 95.3% (121/127).

In regard to the arguments against the use of an app to support the treatment of OA, 32.3% (41/127) of the participants cited data protection and data privacy problems as their main concerns. Concerns about the safety of patients were cited by 24.4% (31/127) of participants, and lack of evidence was cited by 20.5% (26/127). While 14.2% (18/127) of participants feared that the use of an app might impair the doctor-patient relationship, 30.7% (39/127) of participants stated that they had no reservations.

Concerning perceived advantages, 67.7% (86/127) of participants saw the flexible access to the information source as an advantage of app-based OA therapy. The flexible use of

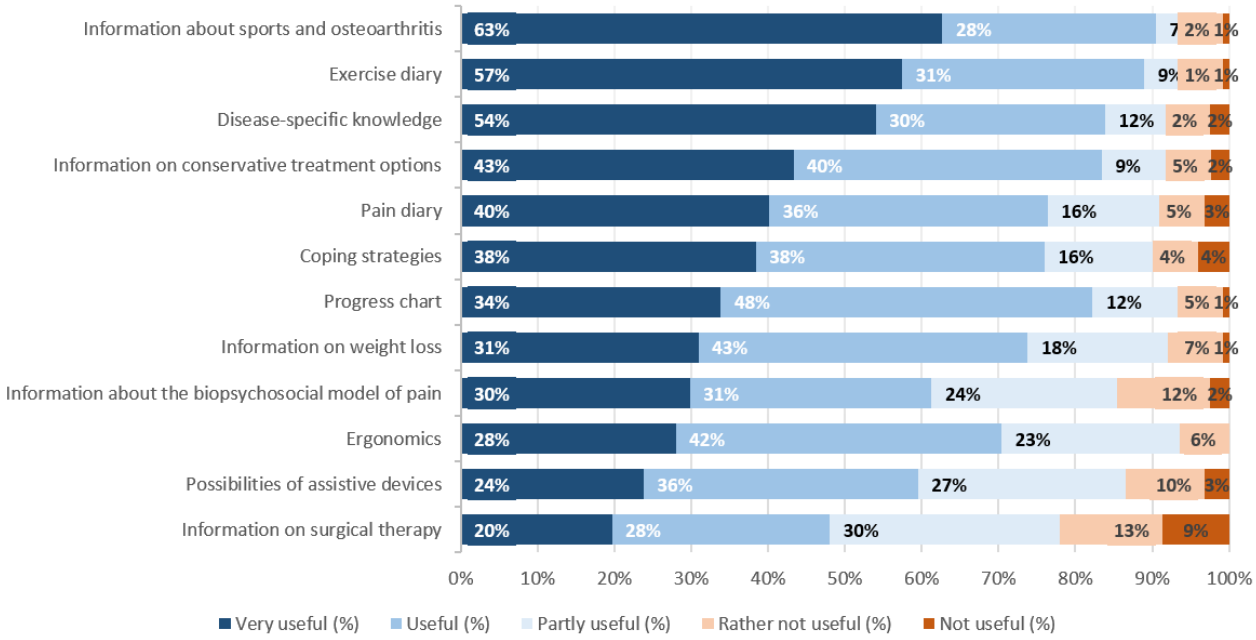
XSL•FO
RenderX

exercises was viewed as an advantage by 78.7% (100/127), while 63.0% (80/127) perceived the strengthening of competence in disease management to be a positive aspect of app-based therapy. The independence of appointments with health care providers was seen as an opportunity by 44.9% (57/127) of participants. It was stated by 25.2% (32/127) of participants that they considered a reduction in the number of prescriptions that might result from providing patients with an app to be valuable. The health education content that was emphasized by health care professionals as being potentially important content for an OA app is shown in Figure 2.

Participants' attitudes toward possible exercise modules are shown in Figure 3. Approximately 77.2% (98/115) of the participants said they would welcome the integration of coaching procedures into the app, but 14.8% (17/115) said they were against it. It was believed by 76.0% (77/115) of participants that patients should receive feedback (eg, via SMS text messaging), while 33.0% (38/115) did not recommend this function.

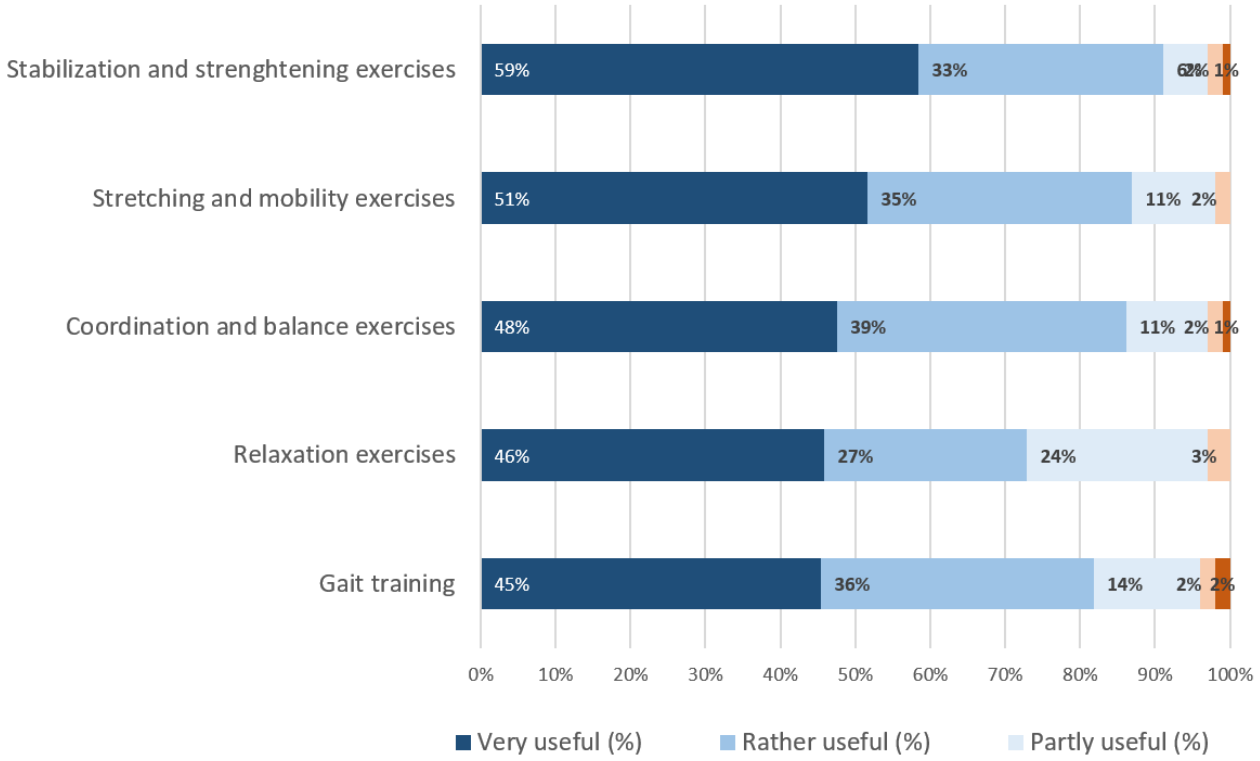
Furthermore, participants were asked about their opinion on the extent to which an OA app should be connected to other telemedical systems. The results of this question are shown in Figure 4.

Figure 2. Health care professionals´ opinions regarding possible educational content for an osteoarthritis app.



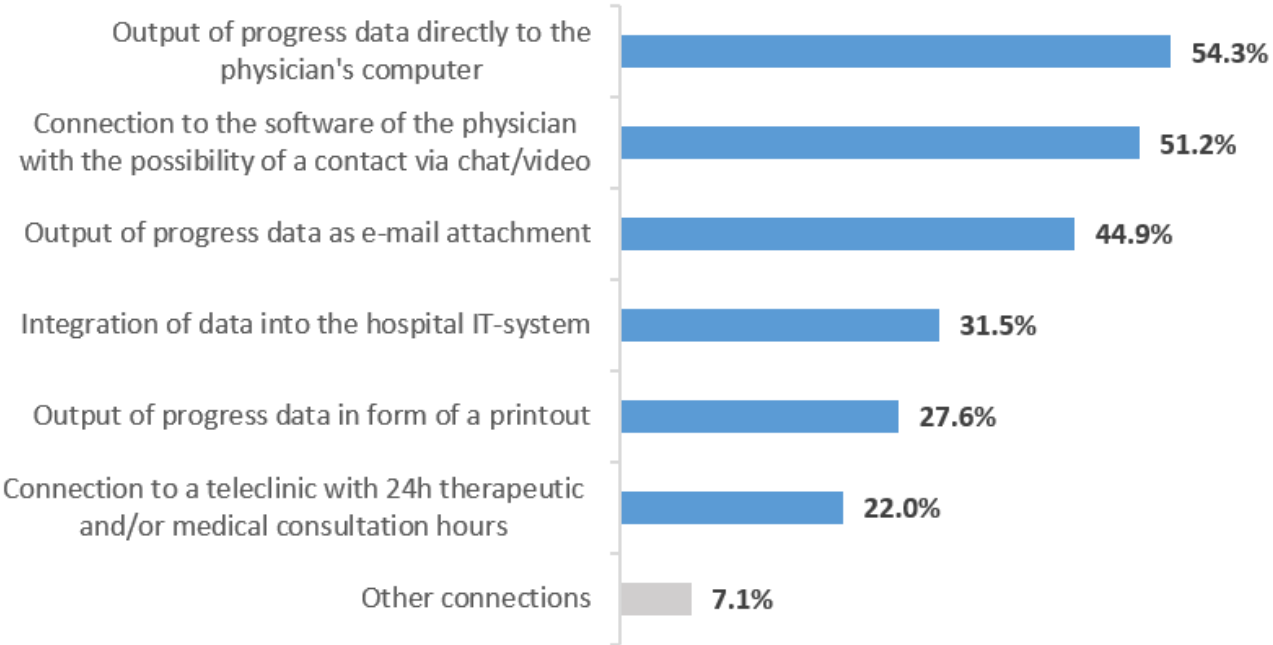
XSL•FO
RenderX

Figure 3. Health care professionals' opinions regarding different types of exercises.



In addition, the participants were asked for which patient groups an OA app could be particularly useful. Here, 31.5% (40/127) stated that a corresponding app could be of preventive value for all persons with a BMI above 25 kg/m^2 and an age over 65 years. Free access for all individuals who want to remain active in old age was advocated by 54.3% (69/127). Approximately 50.4% (64/127) of participants were in favor of the use of an app for patients with OA after prior consultation with a general practitioner, orthopedist, physician, or pain therapist. Among the participants, 22.8% (29/127) were in favor of using an app for all OA patients who had previously received at least 18 therapy sessions with an outpatient physiotherapist, while 48.8% (62/127) believed that patients should be treated with an OA app after 18 physiotherapy units with instruction of the exercises included in the app. The approval rate for patients with OA who have already received multimodal therapy to use an app for continuous follow-up and therapy was 44.1% (56/127).

Figure 4. Consent to telemedical connections. IT Information Technology; 24h: 24-hour.



XSL•FO
RenderX

Of the devices that could potentially improve the app's offering, 70.9% (90/127) of participants named fitness wristbands. The connection to a digital scale (eg, to check weight reduction) was considered useful by 37.0% (47/127) of participants. A connection to a blood pressure monitor for measurement at home was supported by 37.8% (48/127) of participants, while 25.2% (32/127) found a connection to a blood glucose meter useful. Approximately 15.0% (19/127) of participants stated that they thought that connecting external devices would not enrich the app.

Discussion

Principal Findings

Our survey showed that the vast majority of the HCPs surveyed, regardless of their prior professional experience or their discipline, were in favor of using a medical mobile app for patients with OA of the hip or knee. This suggests that HCPs would be likely to integrate well-developed medical apps into therapeutic regimens for patients with OA. The findings also indicate the large, untapped potential of HCPs to raise awareness of mHealth apps and to guide such novel treatment approaches.

We had expected that practitioners with a longer history of treating patients "conventionally" would be more cautious about this novel form of therapy. This assumption was also based on previous research that showed a rather reserved attitude of older HCPs regarding the use of mobile apps in their everyday work [29]. In our study, however, the different groups of professional experience did not differ significantly in their approval of app-based therapy.

Peeters et al [30] stated that patients who use technology to cope with their complaints have more disease-specific knowledge and a better understanding of their condition. However, it is precisely those patient groups who rarely use these resources who would benefit most from app-based information services and interventions [31]. In a survey by Rasche et al [32], users of health apps and general apps who were over 60 years old indicated that they obtain information about apps from family and friends, the internet, digital distribution platforms (eg, the App Store), magazines, television, and experts, with experts being the least used source of information. Therefore, medical professionals could play an important mediating role by reducing the inhibitions of chronically ill patients to use apps to manage their condition with a medical app.

Regarding the arguments against the use of digital care for patients with OA, data privacy issues were most frequently cited. These concerns should be taken seriously, and respective concepts for data privacy protection should be applied and presented in a transparent and comprehensible manner for both experts and users. The General Data Protection Regulation, which is now valid in the European Union, was developed to ensure transparency and reliability in the use of personal data. Even if there are still some uncertainties regarding the concrete practical implementation, there is a clear set of rules for the use of data with regard to app-based therapy [33]. Considerations for patient safety were also mentioned. Accordingly, there should be comprehensive concepts for patient protection. These could include 24-hour customer support, an integrated evaluation of red flags, and clear instructions on when to seek further medical help. Furthermore, vigilance systems can be implemented by manufacturers to collect available information on the safety profile of their apps in real-world use.

In addition to the interaction between physician and patient, Miyamoto et al [34] identified the integration of apps into existing health care services as a key element in initiating behavior changes in patients. Patients participating in the study wanted their collected data to be put into the context of their existing medical records using health apps so that they could receive individual and optimal medical advice based on the synopsis of their findings [34]. However, with regard to embedding an app in existing systems, a relatively large number of participants in our survey showed a certain reluctance. This could possibly be due to fears that app-based interactions with patients would be incalculable, difficult to plan, and involve additional work. In order to achieve wide acceptance, certain concepts might be advantageous to ensure that the individual practitioner is not confronted with unexpected, urgent requests with a direct need for action, even outside of office hours. Here, for example, a central primary contact who could process user requests could be established by the provider of the app.

The extent to which medical professionals advocate for the use of coaching strategies and individualized feedback (eg, via SMS text messaging) was another item in our questionnaire. More than two-thirds of the participants supported this. Behavior therapy strategies could help patients to learn how to initiate sustainable behavior changes on the one hand and how to deal with their illness on the other. For this purpose, an app design would be conceivable in which positive feedback, rewards for reaching previously defined goals, assistance with motivation problems, and the possibility of contacting experts or other interested parties could be integrated [35]. Frequently, therapeutic strategies developed for personal interaction between practitioner and patient are integrated into apps. To what extent modifications are useful and necessary here should be the subject of further research. In addition, the existing high drop-out rate in app-based therapy, as described by Krebs and Duncan [36], could possibly be overcome using approaches based on motivation psychology.

The active participation of medical professionals in the development of digital health apps would be helpful and desirable to ensure that HCPs' expectations are met. Noergaard et al [37] underlined the great value of participatory development of health care services. In addition to the involvement of experts, the participation of patients in the development of digital services is essential [37]. Therefore, as a next step, surveys should be conducted with people affected by OA in order to gain precise knowledge of their needs and expectations. Findings from these future studies might serve to increase the acceptance and adherence among patients, which might potentially increase recommendations by HCPs when implemented.

Based on the findings of our survey, the following content should be considered when developing an app: (1) knowledge units, (2) exercise modules that cover a wide physiotherapeutic

XSL•FO
RenderX

spectrum, and (3) psychological content in the form of motivation-promoting strategies and relaxation techniques. A combination of the various components of medical treatment, pain psychological strategies, physical activity, and patient information would represent a multimodal therapy approach and reflect current evidence and national guidelines [26,38-40]. Conceptual considerations of the biopsychosocial model of pain should also be included [41]. To our knowledge, there is currently no app for the indication of OA that completely covers all of these aspects and meets the declared expectations of HCPs.

For OA of the hip and knee joint, the 6-minute walking test is an important assessment tool. Stienen et al [42] evaluated an app-based 6-minute walking test for patients with degenerative diseases of the lumbar spine, which proved to be highly reliable with the results of the usual, nondigital execution of the test. In addition to other assessments, this test is also an important tool in the evaluation and follow-up of OA of the hip or knee joint [43,44]. The embedding of such assessment tools into an OA app could make a decisive contribution to the high quality of the app. Data show that for patients with OA of the knee, walking can have a beneficial effect on symptoms and functioning [45,46]. As a strategy to promote regular walking, tools just as pedometers could be integrated into an OA app.

Limitations

One limitation of our study is the rather low participation of physicians who perform surgery in the survey population. This could explain why knowledge units about surgical options met with the least approval. Furthermore, the survey was conducted at a congress for rehabilitation medicine. This resulted in a certain preselection of participants in terms of the specialties of participants and a bias toward more academically oriented professionals choosing to attend an academic conference. Due to the survey having partially closed questions, a certain loss of information and a selection bias cannot be excluded. Other approaches, most notably qualitative methods, might lead to different outcomes when assessing the expectations of HCPs toward apps for OA. Furthermore, the high approval rate for app-based therapy may have been influenced by the fact that, as is known, people who are interested in a certain topic are more likely to participate in a corresponding survey [47].

Conclusions

In our survey, there is a very positive attitude of HCPs toward app-based therapy for patients with OA of the hip or knee, indicating untapped potential in the development of an appropriate app. Because HCPs, in principle, see great opportunities in app-based therapy, well-thought-out, secure apps should stand a great chance of being recommended and used in practice. It turned out that an app structure with various modules consisting of knowledge transfer, physical exercises, and practices based on motivation psychology was widely supported. Future studies in the field should address patients' expectations regarding mHealth treatments for OA to ensure these expectations are known and met.

Acknowledgments

The study was part of the project TELE-CORRECT funded by the Bavarian Ministry of economic affairs, regional development, and energy.

Conflicts of Interest

SH is an employee of Kaia Health, a manufacturer of medical mobile apps, and receives cash and options.

Multimedia Appendix 1
Questionnaire (English version).
[DOC File , 81 KB - mhealth_v8i10e21704_app1.doc]

References

1. World Health Organization. mHealth: New Horizons for Health through Mobile Technologies: Based on the Findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3). Geneva: WHO Press; 2011.
2. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: Experiences, Possibilities, and Perspectives. JMIR mHealth uHealth 2014 May 16;2(2):e24. [doi: 10.2196/mhealth.3328]
3. Trifan A, Oliveira M, Oliveira JL. Passive Sensing of Health Outcomes Through Smartphones: Systematic Review of Current Solutions and Possible Limitations. JMIR Mhealth Uhealth 2019 Aug 23;7(8):e12649. [doi: 10.2196/12649]
4. Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clin Exp Res 2005 Aug 25;17(4):255-263. [doi: 10.1007/BF03324607] [Medline: 16285189]
5. Elders MJ. The increasing impact of arthritis on public health. J Rheumatol Suppl 2000 Oct;60:6-8. [Medline: 11032095]
6. Skou ST, Roos EM. Physical therapy for patients with knee and hip osteoarthritis: supervised, active treatment is current best practice. Clin Exp Rheumatol 2019;37 Suppl 120(5):112-117. [Medline: 31621559]
7. Hagen KB, Smedslund G, Østerås N, Jamtvedt G. Quality of Community-Based Osteoarthritis Care: A Systematic Review and Meta-Analysis. Arthritis Care Res (Hoboken) 2016;68(10):1443-1452. [doi: 10.1002/acr.22891]
8. Holman H, Lorig K. Patient Self-Management: A Key to Effectiveness and Efficiency in Care of Chronic Disease. Public Health Rep 2004;119(3):239-243. [doi: 10.1016/j.phr.2004.04.002]

XSL•FO
RenderX

9. Luik AI, Kyle SD, Espie CA. Digital Cognitive Behavioral Therapy (dCBT) for Insomnia: a State-of-the-Science Review. Curr Sleep Med Rep 2017 May 8;3(2):48-56. [doi: 10.1007/s40675-017-0065-4]
10. Zhang L, He X, Shen Y, Yu H, Pan J, Zhu W, et al. Effectiveness of Smartphone App–Based Interactive Management on Glycemic Control in Chinese Patients With Poorly Controlled Diabetes: Randomized Controlled Trial. J Med Internet Res 2019 Dec 9;21(12):e15401. [doi: 10.2196/15401]
11. Mecklenburg G, Smittenaar P, Erhart-Hledik JC, Perez DA, Hunter S. Effects of a 12-Week Digital Care Program for Chronic Knee Pain on Pain, Mobility, and Surgery Risk: Randomized Controlled Trial. J Med Internet Res 2018 Apr 25;20(4):e156. [doi: 10.2196/jmir.9667]
12. Shebib R, Bailey JF, Smittenaar P, Perez DA, Mecklenburg G, Hunter S. Randomized controlled trial of a 12-week digital care program in improving low back pain. NPJ Digit Med 2019 Jan 7;2(1). [doi: 10.1038/s41746-018-0076-7]
13. Toelle TR, Utpadel-Fischler DA, Haas K, Priebe JA. App-based multidisciplinary back pain treatment versus combined physiotherapy plus online education: a randomized controlled trial. NPJ Digit Med 2019 May 3;2(1). [doi: 10.1038/s41746-019-0109-x]
14. Alnasser A, Kyle J, Aloumi N, Al-Khalifa A, Marais D. The Twazon Arabic Weight Loss App: App-Based Intervention for Saudi Women With Obesity. JMIR Mhealth Uhealth 2019 May 28;7(5):e10923. [doi: 10.2196/10923]
15. Bennett GG, Steinberg D, Askew S, Levine E, Foley P, Batch BC, et al. Effectiveness of an App and Provider Counseling for Obesity Treatment in Primary Care. American Journal of Preventive Medicine 2018 Dec;55(6):777-786. [doi: 10.1016/j.amepre.2018.07.005]
16. Thorstensson CA, Garellick G, Rystedt H, Dahlberg LE. Better Management of Patients with Osteoarthritis: Development and Nationwide Implementation of an Evidence-Based Supported Osteoarthritis Self-Management Programme. Musculoskelet. Care 2014 Oct 24;13(2):67-75. [doi: 10.1002/msc.1085]
17. Dahlberg LE, Grahn D, Dahlberg JE, Thorstensson CA. A Web-Based Platform for Patients With Osteoarthritis of the Hip and Knee: A Pilot Study. JMIR Res Protoc 2016 Jun 03;5(2):e115. [doi: 10.2196/resprot.5665]
18. Dahlberg LE, Dell'Isola A, Lohmander LS, Nero H. Improving osteoarthritis care by digital means - Effects of a digital self-management program after 24- or 48-weeks of treatment. PLoS ONE 2020 Mar 4;15(3):e0229783. [doi: 10.1371/journal.pone.0229783]
19. Smittenaar P, Erhart-Hledik JC, Kinsella R, Hunter S, Mecklenburg G, Perez D. Translating Comprehensive Conservative Care for Chronic Knee Pain Into a Digital Care Pathway: 12-Week and 6-Month Outcomes for the Hinge Health Program. JMIR Rehabil Assist Technol 2017 Apr 05;4(1):e4. [doi: 10.2196/rehab.7258]
20. Byambasuren O, Beller E, Glasziou P. Current Knowledge and Adoption of Mobile Health Apps Among Australian General Practitioners: Survey Study. JMIR Mhealth Uhealth 2019 Jun 3;7(6):e13199. [doi: 10.2196/13199]
21. Ayre J, Bonner C, Bramwell S, McClelland S, Jayaballa R, Maberly G, et al. Factors for Supporting Primary Care Physician Engagement With Patient Apps for Type 2 Diabetes Self-Management That Link to Primary Care: Interview Study. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11885. [doi: 10.2196/11885]
22. Kerst A, Zielasek J, Gaebel W. Smartphone applications for depression: a systematic literature review and a survey of health care professionals' attitudes towards their use in clinical practice. Eur Arch Psychiatry Clin Neurosci 2019 Jan 3;270(2):139-152. [doi: 10.1007/s00406-018-0974-3]
23. Kessel KA, Vogel MM, Schmidt-Graf F, Combs SE. Mobile Apps in Oncology: A Survey on Health Care Professionals' Attitude Toward Telemedicine, mHealth, and Oncological Apps. J Med Internet Res 2016 Nov 24;18(11):e312. [doi: 10.2196/jmir.6399]
24. Najm A, Gossec L, Weill C, Benoist D, Berenbaum F, Nikiphorou E. Mobile Health Apps for Self-Management of Rheumatic and Musculoskeletal Diseases: Systematic Literature Review. JMIR Mhealth Uhealth 2019 Nov 26;7(11):e14730. [doi: 10.2196/14730]
25. Langbecker D, Caffery LJ, Gillespie N, Smith AC. Using survey methods in telehealth research: A practical guide. J Telemed Telecare 2017 Jul 20;23(9):770-779. [doi: 10.1177/1357633x17721814]
26. S2k-Leitlinie Gonarthrose. AWMF Online. URL: https://www.awmf.org/uploads/tx_szleitlinien/033-004l_S2k_Gonarthrose_2018-01_1-verlaengert.pdf [accessed 2020-10-08]
27. Bannuru R, Osani M, Vaysbrot E, Arden N, Bennell K, Bierma-Zeinstra S, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthritis and Cartilage 2019 Nov;27(11):1578-1589. [doi: 10.1016/j.joca.2019.06.011]
28. Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34]
29. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T 2014 May;39(5):356-364 [FREE Full text] [Medline: 24883008]
30. Peeters J, Wiegers T, Friele R. How technology in care at home affects patient self-care and self-management: a scoping review. Int J Environ Res Public Health 2013 Oct 29;10(11):5541-5564 [FREE Full text] [doi: 10.3390/ijerph10115541] [Medline: 24173139]
31. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health App Use Among US Mobile Phone Users: Analysis of Trends by Chronic Disease Status. JMIR Mhealth Uhealth 2017 Dec 19;5(12):e197. [doi: 10.2196/mhealth.7832]

XSL•FO
RenderX

32. Rasche P, Wille M, Bröhl C, Theis S, Schäfer K, Knobe M, et al. Prevalence of Health App Use Among Older Adults in Germany: National Survey. JMIR Mhealth Uhealth 2018 Jan 23;6(1):e26. [doi: 10.2196/mhealth.8619]
33. McCall B. What does the GDPR mean for the medical community? The Lancet 2018 Mar;391(10127):1249-1250. [doi: 10.1016/s0140-6736(18)30739-6]
34. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking Health Data Is Not Enough: A Qualitative Exploration of the Role of Healthcare Partnerships and mHealth Technology to Promote Physical Activity and to Sustain Behavior Change. JMIR mHealth uHealth 2016 Jan 20;4(1):e5. [doi: 10.2196/mhealth.4814]
35. Abraham, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
36. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR mHealth uHealth 2015 Nov 04;3(4):e101. [doi: 10.2196/mhealth.4924]
37. Noergaard B, Sandvei M, Rottmann N, Johannessen H, Wiil U, Schmidt T, et al. Development of a Web-Based Health Care Intervention for Patients With Heart Disease: Lessons Learned From a Participatory Design Study. JMIR Res Protoc 2017 May 17;6(5):e75. [doi: 10.2196/resprot.7084]
38. Flor H, Fydrich T, Turk DC. Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. Pain 1992;49(2):221-230. [doi: 10.1016/0304-3959(92)90145-2]
39. Hurley MV, Walsh NE, Mitchell H, Nicholas J, Patel A. Long-term outcomes and costs of an integrated rehabilitation program for chronic knee pain: A pragmatic, cluster randomized, controlled trial. Arthritis Care Res 2012 Jan 25;64(2):238-247. [doi: 10.1002/acr.20642]
40. Krauß I, Steinhilber B, Haupt G, Miller R, Martus P, Janßen P. Exercise therapy in hip osteoarthritis--a randomized controlled trial. Dtsch Arztebl Int 2014 Sep 01;111(35-36):592-599 [FREE Full text] [doi: 10.3238/arztebl.2014.0592] [Medline: 25249361]
41. Engel G. The need for a new medical model: a challenge for biomedicine. Science 1977 Apr 08;196(4286):129-136. [doi: 10.1126/science.847460]
42. Stienen M, Gautschi OP, Staartjes VE, Maldaner N, Sosnova M, Ho AL, et al. Reliability of the 6-minute walking test smartphone application. J Neurosurg Spine 2019 Sep 13:1-8. [doi: 10.3171/2019.6.SPINE19559] [Medline: 31518975]
43. Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. BMC Musculoskelet Disord 2005 Jan 28;6(1). [doi: 10.1186/1471-2474-6-3]
44. Lord SR, Menz HB. Physiologic, psychologic, and health predictors of 6-minute walk performance in older people. Archives of Physical Medicine and Rehabilitation 2002 Jul;83(7):907-911. [doi: 10.1053/apmr.2002.33227]
45. Evcik D, Sonel B. Effectiveness of a home-based exercise therapy and walking program on osteoarthritis of the knee. Rheumatology International 2002 Jul 1;22(3):103-106. [doi: 10.1007/s00296-002-0198-7]
46. White DK, Tudor-Locke C, Zhang Y, Fielding R, LaValley M, Felson DT, et al. Daily Walking and the Risk of Incident Functional Limitation in Knee Osteoarthritis: An Observational Study. 2014 Aug 26;66(9):1328-1336. [doi: 10.1002/acr.22362]
47. Robert M. Groves. Nonresponse Rates and Nonresponse Bias in Houshold Surveys. The Public Opinion Quarterly 2006;70:646-675. [doi: 10.1093/poq/nfl033]

Abbreviations

BOA: Management of Patients with Osteoarthritis program
HCP: health care professional
LMU Munich: Ludwig Maximilian University of Munich
mHealth: mobile health
OA: osteoarthritis

Edited by G Eysenbach; submitted 22.06.20; peer-reviewed by A Trifan, L Sandal; comments to author 11.07.20; revised version received 19.08.20; accepted 03.09.20; published 28.10.20.

Please cite as:
Biebl JT, Huber S, Rykala M, Kraft E, Lorenz A
Attitudes and Expectations of Health Care Professionals Toward App-Based Therapy in Patients with Osteoarthritis of the Hip or Knee: Questionnaire Study
JMIR Mhealth Uhealth 2020;8(10):e21704
URL: http://mhealth.jmir.org/2020/10/e21704/
doi:10.2196/21704
PMID:33112255

XSL•FO
RenderX

©Johanna Theresia Biebl, Stephan Huber, Marzena Rykala, Eduard Kraft, Andreas Lorenz. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 28.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessing the Food and Drug Administration's Risk-Based Framework for Software Precertification With Top Health Apps in the United States: Quality Improvement Study

Noy Alon[1]; Ariel Dora Stern[2], PhD; John Torous[1], MD

[1]Beth Israel Deaconess Medical Center, Boston, MA, United States

[2]Harvard Business School, Boston, MA, United States

Corresponding Author:
John Torous, MD
Beth Israel Deaconess Medical Center
330 Brookline Avenue
Boston, MA, 02215
United States
Phone: 1 7143359858
Email: jtorous@bidmc.harvard.edu

Abstract

Background: As the development of mobile health apps continues to accelerate, the need to implement a framework that can standardize the categorization of these apps to allow for efficient yet robust regulation is growing. However, regulators and researchers are faced with numerous challenges, as apps have a wide variety of features, constant updates, and fluid use cases for consumers. As past regulatory efforts have failed to match the rapid innovation of these apps, the US Food and Drug Administration (FDA) has proposed that the Software Precertification (Pre-Cert) Program and a new risk-based framework could be the solution.

Objective: This study aims to determine whether the risk-based framework proposed by the FDA's Pre-Cert Program could standardize categorization of top health apps in the United States.

Methods: In this quality improvement study during summer 2019, the top 10 apps for 6 disease conditions (addiction, anxiety, depression, diabetes, high blood pressure, and schizophrenia) in Apple iTunes and Android Google Play Store in the United States were classified using the FDA's risk-based framework. Data on the presence of well-defined app features, user engagement methods, popularity metrics, medical claims, and scientific backing were collected.

Results: The FDA's risk-based framework classifies an app's risk by the disease condition it targets and what information that app provides. Of the 120 apps tested, 95 apps were categorized as targeting a nonserious health condition, whereas only 7 were categorized as targeting a serious condition and 18 were categorized as targeting a critical condition. As the majority of apps targeted a nonserious condition, their risk categorization was largely determined by the information they provided. The apps that were assessed as not requiring FDA review were more likely to be associated with the integration of external devices than those assessed as requiring FDA review (15/58, 26% vs 5/62, 8%; P=.03) and health information collection (24/58, 41% vs 9/62, 15%; P=.008). Apps exempt from the review were less likely to offer health information (25/58, 43% vs 45/62, 72%; P<.001), to connect users with professional care (7/58, 12% vs 14/62, 23%; P=.04), and to include an intervention (8/58, 14% vs 35/62, 55%; P<.001).

Conclusions: The FDA's risk-based framework has the potential to improve the efficiency of the regulatory review process for health apps. However, we were unable to identify a standard measure that differentiated apps requiring regulatory review from those that would not. Apps exempt from the review also carried concerns regarding privacy and data security. Before the framework is used to assess the need for a formal review of digital health tools, further research and regulatory guidance are needed to ensure that the Pre-Cert Program operates in the greatest interest of public health.

(JMIR Mhealth Uhealth 2020;8(10):e20482) doi:10.2196/20482

KEYWORDS

mobile health; smartphone; Food and Drug Administration; software; mobile phone

XSL•FO
RenderX

Introduction

Background

The development of mobile health apps has been increasing in recent years; recent estimates found that approximately 325,000 mobile health apps are available in the marketplace [1]. However, a consequence of rapid technological development is that many health apps remain to be unevaluated by researchers [2]. Thus, clinicians and patients are largely uninformed about the efficacy of these apps and lack data on their potential to benefit health and/or cause harm.

Despite a lack of evidence and in the absence of direct regulation, smartphone ownership and interest in health apps remain to be high among patients [3-5] and those who might not have a diagnosis but are seeking to improve their well-being [6-10]. However, the majority of the population has still not downloaded health apps [3,9-11], and clinicians are hesitant to recommend apps [12] because of concerns over privacy, data security [6-9,11], and app effectiveness [7,12]. As such, the need for evidence, guidance, and thoughtful regulation in digital health is clear [13]. More concrete government regulations have the potential to set a quality baseline and reduce the number of unsubstantiated claims made by health apps. These measures could increase clinicians' and patients' trust in digital health tools [14].

In the past, the US Food and Drug Administration (FDA) focused its regulatory efforts on a small subset of mobile medical apps: those that provided treatment or diagnosis to users and those that were an extension of or transformed into regulated medical devices [15]. These mobile medical apps would be subjected to a formal FDA review and the same regulatory requirements as other medical devices [15]. This review process requires app developers to register their organization and product and provide information regarding the design procedure, facilities, and how their app will be described. Depending on the classification of their product, developers must also submit a premarket notification or approval with supporting clinical evidence [16]. However, the FDA has acknowledged that this framework is not well suited for rapid development and changes made to many health apps [17].

The Software Precertification Pilot Program

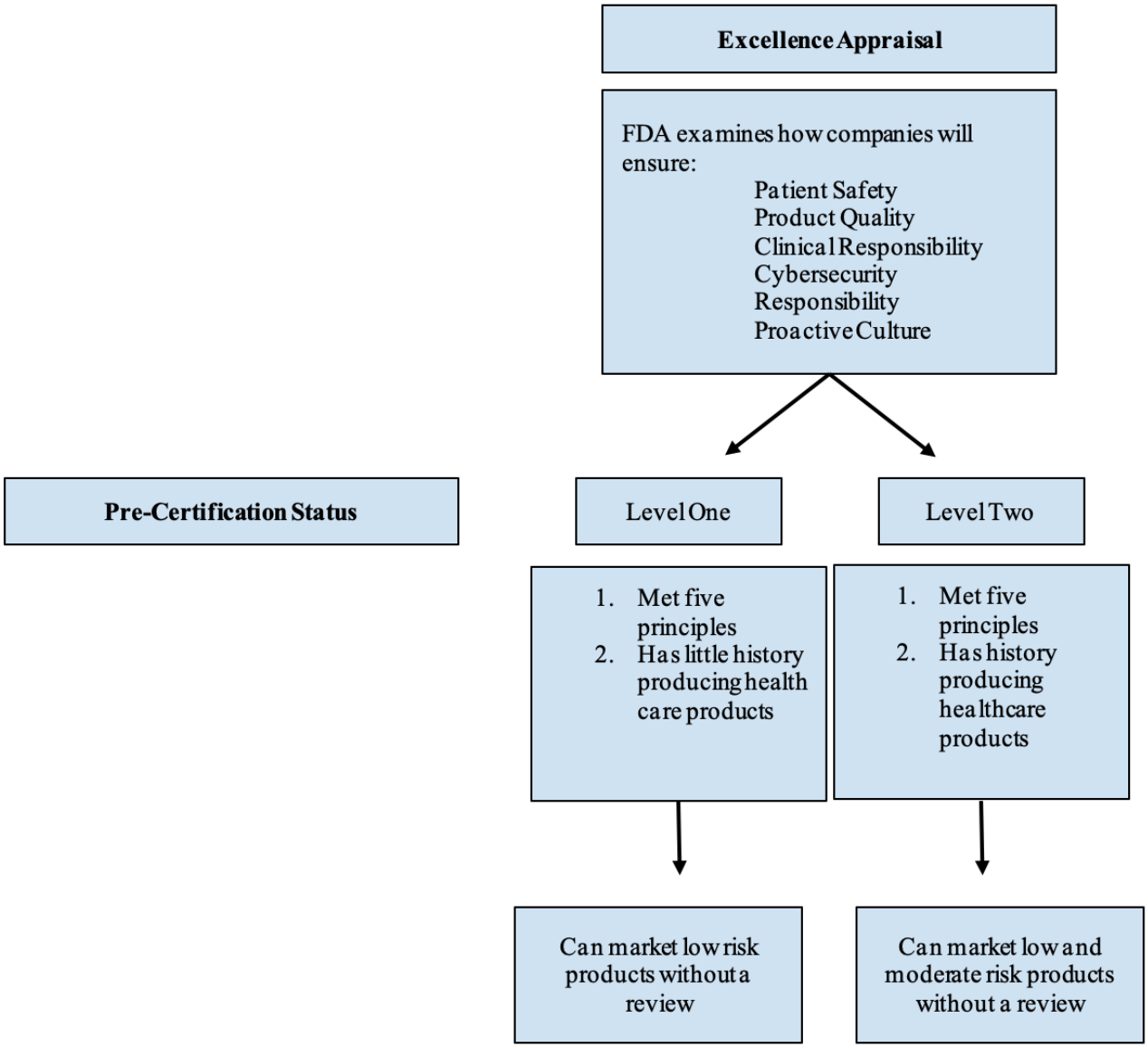
As a result, in June 2018, the FDA published a working model for its Software Precertification (Pre-Cert) Pilot Program and released a *Test Plan* for the program in early 2019 [18,19]. This program hopes to provide a more efficient review process for software-only products that would reduce the regulatory burden of entering certain software product markets and would encourage software developers to advance the capabilities of their products [20]. The Pre-Cert Program is designed to address many of the tensions between software development and traditional regulated medical technologies [17,18], such as the tradition of regular product updating (by software developers) versus testing and quality assurance before infrequent and discrete product updates (by medical device developers) [21,22]. The FDA's work on piloting the Pre-Cert Program has continued in recent months, with ongoing evaluation and mid-2019 reporting on how mock reassessments of already-approved software products would have fared under a streamlined regulatory review process [23].

Under the Pre-Cert Program, FDA regulators plan to first evaluate digital health app developers and not the apps themselves [17]. In its current form, the program will only apply to developers marketing software as a medical device (SaMD), which the FDA defines as software that is intended to be used for medical purposes without a hardware extension [24]. Medical purposes applicable to SaMD are defined by the FDA as including but not limited to the diagnosis, prevention, monitoring, treatment, and alleviation of disease or injury [25].

Under Pre-Cert, FDA regulators would first examine companies through an *excellence appraisal*, during which the FDA would review app developers' policies and practices to determine if and how a developer's policies enable the organization to excel in 5 proposed excellence principles: (1) patient safety, (2) product quality, (3) clinical responsibility, (4) cybersecurity responsibility, and (5) a proactive culture [18]. If a developer is deemed to have met all 5 principles, the FDA will grant it 1 of 2 precertification statuses: A *Level One or Level Two Pre-Cert.* A *Level One Pre-Cert* would enable organizations to market lower-risk software products without any regulatory review, but moderate- and high-risk products would receive the benefit of undergoing a *streamlined* (abbreviated) review process. This status will be given to developer companies that have met the 5 principles but have less experience producing health care products. A *Level Two Pre-Cert* would enable developer companies to market low- and moderate-risk products without any regulatory review (but would require review for high-risk products) and would only be rewarded to developers that both excelled in the 5 principles *and* have a history of producing safe and effective health care products [18]. This process is shown in Figure 1.

XSL•FO
RenderX

Figure 1. Precertification status determination process. This figure is an overview of how the FDA will determine the precertification status of different organizations. FDA: Food and Drug Administration.



After a developer's Pre-Cert status is granted, the type of review (if any review is necessary) that each new software product will undergo would be determined by its risk profile. In addition to the developer's Pre-Cert status, each new software must complete a risk analysis, and together, these designations will determine if a review is necessary. Using a risk-based framework developed by the International Medical Device Regulators Forum (IMDRF) SaMD Working Group, software developers will perform this risk analysis and determine an SaMD's risk by considering the severity of the medical condition it targets and the type of information the app offers [26]. The IMDRF framework categorized medical conditions as nonserious, serious, or critical, and the FDA has further specified the characteristics of each categorization for the Pre-Cert Program. Similarly, the IMDRF broke down the significance of app-provided information into informing clinical care, driving clinical management, or treating and diagnosing, and the FDA uses these categories in the Pre-Cert Program [18]. The combination of these two-dimensional categorizations, coupled with an organization's Pre-Cert status, will then jointly determine whether the FDA would perform a regulatory review for a given SaMD product. If necessary, a review would then be completed before an SaMD product can be marketed. Importantly, the FDA plans to continue regulating products that come to market through the Pre-Cert process by continuously examining their *real-world performance* in the postmarket setting [18,27].

The Pre-Cert Program hopes to streamline the FDA's review process by incorporating FDA oversight during the development of precertified organizations' apps and not just when the app is finalized. The FDA also hopes to minimize the burden on developers to prove their product's efficacy and safety, but the list of reduced requirements has yet to be finalized [17].

Current Frameworks

The FDA's effort to modernize its regulatory framework is not unique, as multiple guidelines attempting to clarify and streamline government regulations of digital health tools have been implemented in both the United States and Europe. In conjunction with the FDA and other departments, the Federal

XSL•FO
RenderX

Trade Commission has developed a web-based survey helping app developers identify what federal regulations pertain to their app [28]. In Europe, the European Commission eHealth Action Plan was first adopted in 2004 and has worked to clarify policies and present the possibilities of using digital health tools to populations throughout the European Union [29]. In 2017, the World Health Organization published guidelines for classifying digital health interventions [30]. In March 2019, the UK National Institute for Health and Care Excellence published a framework establishing standards of evidence for these technologies [31].

In addition, the clinical community has already begun the process of evaluating apps (including highlighting concerns around both patient privacy and app efficacy) [32-35] and thus has gained insight into how patients use apps and the current quality of available apps. As the clinical community has gained experience with digital health and patients and clinicians will be the ones using and recommending these tools, we sought to determine if their findings and concerns are reflected in this new model. To achieve this goal, we simulated the Pre-Cert Program's risk categorization of SaMD products and considered whether such a risk-based framework would be able to differentiate top health apps based on their features. We aimed to determine if there is a correlation between apps' features, attributes, and functionality and their Pre-Cert risk categorization. This correlation could indicate that the framework is reliable and accurate and thus would provide a standardized way to characterize these technologies.

Methods

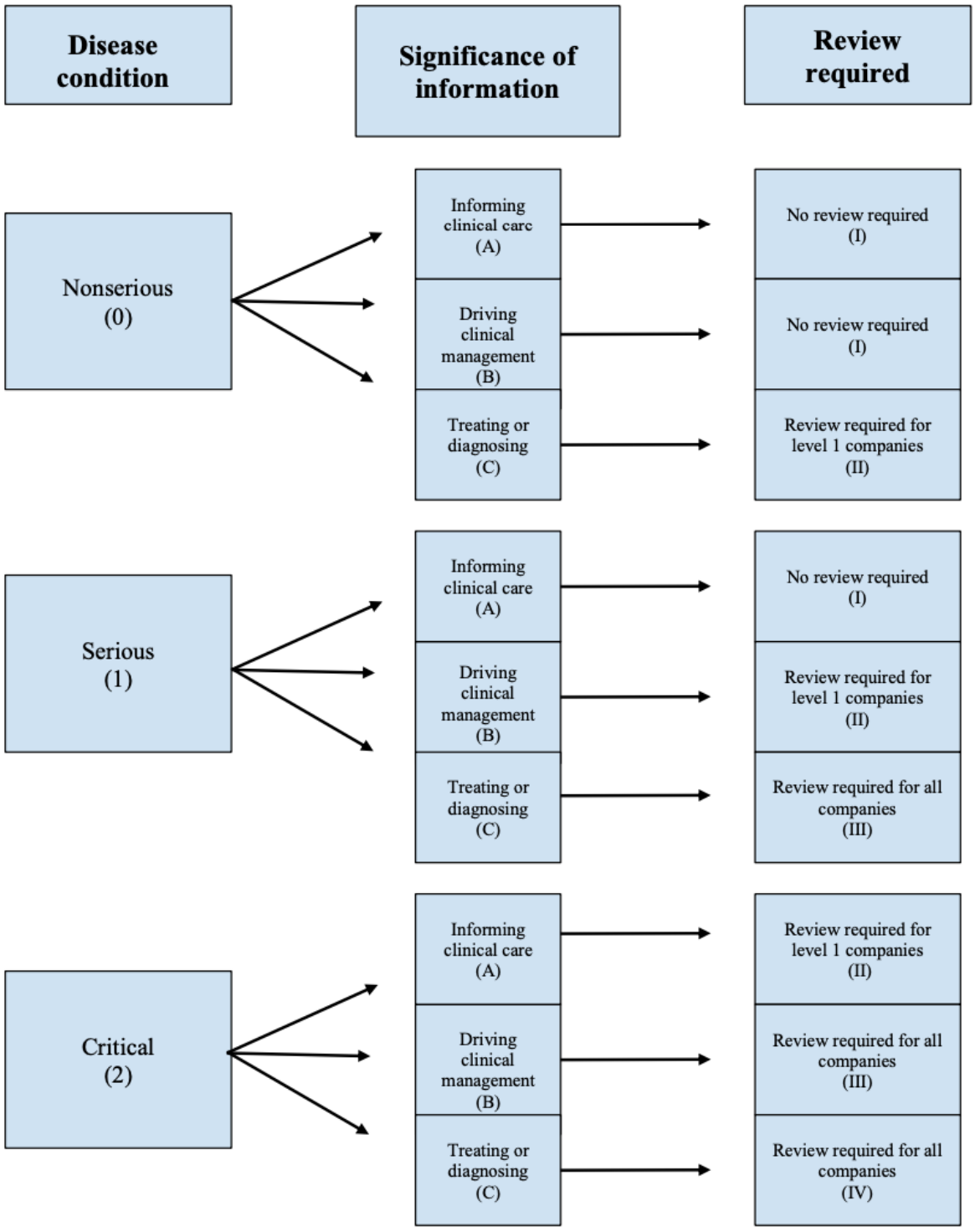
Data Collection

We based our analysis on the methods outlined in a study by Wisniewski et al [32] following published and peer-reviewed methods for app selection and classification that attempt to standardize app categorization. Patient feedback was involved in the design of the codebook used in the study, determining what app features should be included. On June 20, 2019, the top 10 apps in both Apple iTunes and Android Google Play Store in the United States were selected out of a total of 120 apps for 6 common disease conditions: addiction, anxiety, depression, diabetes, high blood pressure, and schizophrenia. As 10 apps for each disease condition were chosen across 2 platforms, 20 apps in total were assessed for each condition. The top 10 apps represent the sample that consumers would likely be exposed to first and thus, most likely to use. As a result, although this sample is a convenience sample, it has clinical and real-world relevance. As the methodology was based on Wisniewski et al [32], the same disease conditions were studied to maintain consistency. In total, 2 independent coders downloaded, used, and evaluated the apps. Dummy profiles were created, and each app was tested with data provided by the researchers. Any disagreements between coders were resolved via discussion until a consensus was reached. All the raw data and significance testing performed were reviewed by a clinician who offered guidance on the clinical relevance of the data. Each coder completed a spreadsheet that prompted information identical to that of Wisniewski et al [32], including the presence of app attributes (such as a written privacy policy), information on how the app gathered and returned data, stated patient engagement methods, visible popularity metrics, stated medical claims, and the presence (or absence) of scientific evidence (ie, evidence-based claims). Wisniewski et al [32] is useful for further reference. Only the presence of observable features was coded for each app. Features such as *ease of use* are dependent on the user and were not included in the codebook in an attempt to achieve reliability between coders in an app's categorization. A risk-based framework, following the characteristics outlined in the FDA's current draft of the Pre-Cert Program, was added to the coding procedure to simulate applying the model in potential regulatory use [18]. A correlation between the presence of apps' features, attributes, and functionality and their Pre-Cert risk categorization could point to the framework's reliability and accuracy when categorizing apps. For example, if an app's ability to access the phone's camera is found to differentiate between apps that require a review and those that are exempt, then this result suggests that the Pre-Cert's risk categorization is not based on subjective measures and has the potential to weed out the apps that pose a greater risk to consumers. Thus, both metrics were coded for, and comparisons were made between them.

We translated both the disease condition and significance of information categories to a numerical scale, allowing for easier data analysis: apps that were deemed to target a nonserious condition were rated as 0, whereas serious and critical conditions were given a 1 and 2, respectively. If an app targeted several diagnoses, it was categorized by the most severe disease condition described. Regarding the information provided, informing clinical care was rated an "A," whereas apps that drove clinical management or treated and diagnosed users were given a "B" and "C," respectively. Using the FDA's current guidelines, we coded apps as informing clinical care if they simply provided information. Any personal data entry that was used to monitor symptoms was coded as driving clinical management, whereas apps providing treatment and diagnosis were differentiated from other functionalities. An app's review required status, the classification that decides whether an FDA review would be required under the Pre-Cert Program, was determined by the combination of both criteria and given the numerical value that the IMDRF working group had previously attributed to each category, ranging from 1 to 4 (I, II, III, and IV). For example, a meditation app claiming to alleviate anxiety and stress would have been coded as targeting a nonserious condition (0) and providing treatment (C). Under the Pre-Cert Program, this app would be given a review level of II, which requires *Level One Pre-Cert* organizations to undergo a streamlined review, whereas *Level Two* organizations are exempt from any FDA review. The rating system used is shown in Figure 2. Following the previous literature, we coded only for the presence of and not the quality of app features to maintain the reliability of the data.

XSL•FO
RenderX

Figure 2. Risk categorization rating system. This figure shows how the Pre-Cert Program uses the disease condition an app targets (0-2) and what information that app provides (A-C) to determine what review that app must undergo (I-IV).



Statistical Analysis

Following coding and data reconciliation, apps were dichotomized into exemption from a review or requiring a review. Apps given an IMDRF categorization of "I" would be exempt from any regulatory review, whereas type II, III, and IV apps would undergo some form of review depending on the precertification status of the organization. As types II, III, and IV apps would undergo some form of review, we grouped these apps together. The data were further stratified by categorical measures, such as which disease condition they targeted. Two-sided *t* tests of differences between categories under the assumption of equal variances were performed using Microsoft Excel 2019 (version 16.0.6742.2048) to determine statistical significance.

XSL•FO
RenderX

Results

Principal Results

Of the 120 total Apple and Android apps examined in the simulation, 95 (79.2%) were categorized as targeting a nonserious health condition, whereas only 7 (5.8%) apps targeted a serious condition, including one app that targeted addiction, 5 that targeted depression, and one that targeted anxiety. The remaining 18 (15.0%) apps targeted a critical condition; however, all apps in this group targeted schizophrenia.

Review required status—that is, the classification that determines if an FDA review would be required under the Pre-Cert Program and represented by code I, II, III, or IV (Figure 2)—was largely determined by the information that the app's developer provided. Of 120 apps, 30 (25.0%) were found to have informed clinical care, whereas driving clinical management and treating or diagnosing each had 42 (35.0%) apps fitting their respective descriptions. The significance of information, which is coded A, B, or C (Figure 2), for the remaining 6 (5.0%) apps was unclassifiable, as these apps were not intended for or did not claim to provide any health-related advice or treatment. As these apps did not offer any health information, they were deemed to be exempt from an FDA regulatory review. These apps were included in the original sample as a general search in both app stores was performed in an attempt to mimic the experience of consumers if they searched for health apps. As a result, not all apps were necessarily marketed under the store's medical category.

When comparing the reliability between Apple and Android apps, no statistically significant differences were found between whether or not review was required for each disease condition between the platforms (addiction: 1.25 vs 1.22, P=.72; anxiety: 1.8 vs 2, P=.34; depression: 1.8 vs 2, P=.55; high blood pressure: 1.22 vs 1.3, P=.41; and schizophrenia: 2.22 vs 2.56; P=.57). This suggests that the Pre-Cert's categorization is reliable across platforms. As both platforms are available for consumers to choose from, it is worthwhile to note the lack of statistical significance between them.

Stratification by Review Required Status

The number of apps and their features were stratified by whether a review was required and are shown in Figure 3. The features coded for are grouped thematically, as features associated with gathering data are labeled (in) and those involved with user engagement or presenting information are labeled (out; Figure 3). App attributes associated with privacy, medical claims, presence of scientific evidence, connection to professional care, and use of rewards or inventions are listed below.

Two-sided t tests comparing all app features described between apps that the Pre-Cert Program exempted from review and those that would require a regulatory review were performed (Figure 3). Apps with a review level of 2 or greater were combined because of the small number of apps under the third and fourth review levels (III and IV) and because all these apps would require a review. Apps that did not require any review (review level I) were more likely than those that definitely or may have required review to integrate data from external devices such as smartwatches (15/58, 26% of review level I apps vs 5/62, 8% of reviews II, III, *and IV level apps*; P=.03) and collect health information such as step counts (24/58, 41% vs 9/62, 15%; P=.008). Apps expected to be exempt from FDA review (I) were less likely to offer information or reference facts (25/58, 43% vs 45/62, 72%; P<.001), were less likely to connect to professional care (7/58, 12% vs 14/62, 23%; P=.04), and were less likely to include an intervention (8/58, 14% vs 35/62, 55%; P<.001) than those requiring a review. User-given star ratings also significantly differed between apps that did not require review versus those that definitely or may have required review (4.48 vs 4.13; P=.003), suggesting that streamlined apps were rated more highly than those definitely or potentially requiring a review. Notably, there were no statistically significant differences between the 2 potential review profiles in the provision of information about the ability for data deletion (37/58, 64% vs 34/62, 55%; P=.20) or the average days since the app's last update (189 vs 264; P=.60).

The mean values and SDs of proxies for app popularity (star ratings, number of reviews, and days since the last update) and data on days since each app's last update are summarized in Table 1.

XSL•FO
RenderX

Figure 3. App features by review required. The orange bars represent apps that would undergo a regulatory review in the Pre-Cert Program (review levels II, III, and IV), and the blue bars represent apps exempt from review (review level I).

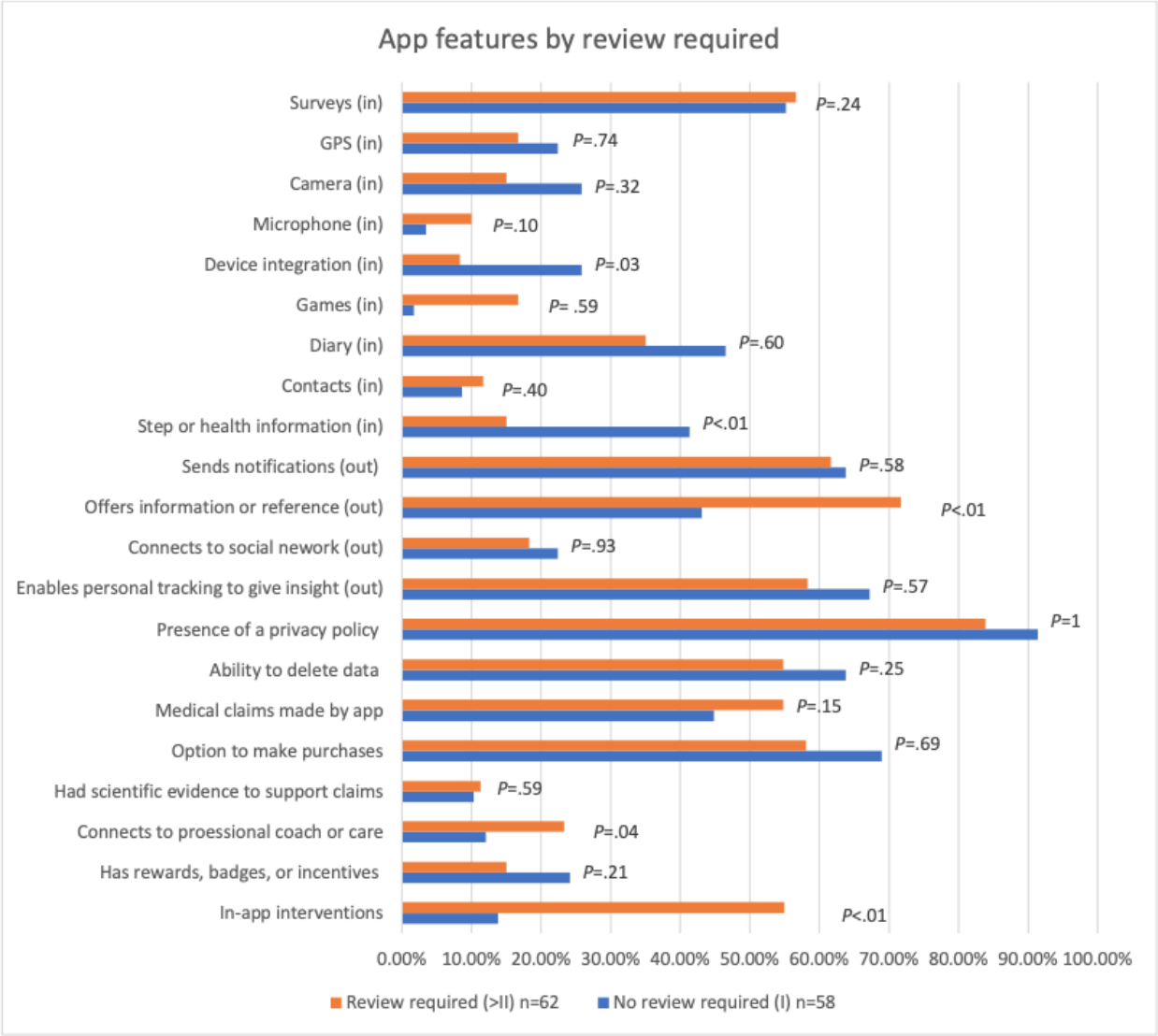


Table 1. Popularity metrics and update history by review required.

Review required	No review required (n=58)	Review required (n=62)
User star ratings, mean (SD)	4.48 (0.6)	4.13 (1)
Number of ratings, mean (SD)	14,554 (42,409)	14,018 (70,454)
Days since last update, mean (SD)	189 (335)	264 (338)

Stratification by Targeted Disease

When the data were stratified by targeted disease, the number of apps requiring review within each condition varied dramatically. In total, 16 apps targeting addiction required no review (I), while only 4 apps would undergo a review (II, III, and IV). A similar trend is seen in apps targeting high blood pressure as 15 of these apps were determined to be exempt from a review (I), leaving only 5 apps to undergo a review (II, III, and IV). Notably, all apps targeting diabetes were deemed to be exempt from any review (I), whereas anxiety, depression, and schizophrenia apps comprised the majority of those definitely or possibly requiring review (II, III, and IV). This result was driven by the finding that most anxiety and depression apps offered treatment and schizophrenia was classified as a critical disease condition, which resulted in a higher likelihood that review would be required.

When stratified by disease, the samples become underpowered because of the small number of apps in each subgroup. However, it is worth noting that some of the differences noted earlier remain. For example, only 6% (1/16) of addiction apps exempt from review offered an intervention, whereas 100% (4/4) of addiction apps requiring a review did. The same trend holds true for apps targeting anxiety (0/3, 0% vs 16/17, 94%) and

XSL•FO
RenderX

depression (1/6, 17% vs 11/14, 79%). In addition, none of the apps targeting depression that were exempt from review offered external information, whereas 79% (11/14) of those requiring reviews did. Notable results are summarized in Tables 2 and 3, whereas a full table comparing app features and whether review would be required, stratified by targeted disease, is provided in Multimedia Appendix 1.

Table 2. Apps' features for addiction, anxiety, and depression by review required, stratified by targeted disease.

App features	Addiction apps exempt from review (n=16)	Addiction apps requiring reviews (n=4)	Anxiety apps exempt from review (n=3)	Anxiety apps requiring reviews (n=17)	Depression apps exempt from review (n=6)	Depression apps requiring reviews (n=14)
Device integration, n (%)	0 (0)	0 (0)	0 (0)	2 (12)	1 (17)	0 (0)
Steps or health information, n (%)	0 (0)	0 (0)	1 (33)	3 (18)	1 (17)	3 (21)
Offer information, n (%)	2 (13)	4 (100)	2 (67)	12 (71)	0 (0)	11 (79)
Connect to professional care, n (%)	5 (31)	1 (25)	1 (33)	6 (35)	0 (0)	6 (43)
In-app interventions, n (%)	1 (6)	4 (100)	0 (0)	16 (94)	1 (17)	11 (79)
User star ratings, mean (SD)	4.7 (0.18)	4.68 (0.17)	4.5 (0.52)	4.65 (0.19)	4.55 (0.23)	4.2 (0.56)

Table 3. Apps' features for diabetes, high blood pressure, and schizophrenia.

App features	Diabetes apps exempt from review (n=20)	High blood pressure apps exempt from review (n=15)	High blood pressure apps requiring review (n=5)	Schizophrenia apps exempt from review (n=4)	Schizophrenia apps requiring review (n=16)
Device integration, n (%)	11 (55)	3 (20)	0 (0)	0 (0)	3 (19)
Steps or health information, n (%)	8 (40)	1 (7)	0 (0)	0 (0)	3 (19)
Offer information, n (%)	14 (70)	5 (33)	1 (20)	2 (50)	15 (94)
Connect to professional care, n (%)	1 (5)	0 (0)	0 (0)	0 (0)	1 (6)
In-app interventions, n (%)	6 (30)	0 (0)	1 (20)	0 (0)	1 (6)
User star ratings, mean (SD)	4.48 (0.47)	4.13 (0.94)	3.28 (1.19)	4.18 (0.57)	2.5 (1.9)

Discussion

Principal Findings

After coding for the presence of observable features of top health apps, we found attributes that differentiated the apps that would likely undergo an FDA regulatory review under the Pre-Cert Program versus those that would not. Apps offering interventions were most likely to require a review (II, III, and IV), whereas monitoring apps were more likely to be streamlined. In addition, apps requiring FDA review were more likely to offer references and connect users to professional care than streamlined apps. This distinction between formal medical advice and user-led data embedded in the FDA's risk categorization demonstrates a promising foundation for the framework. Apps gearing themselves toward providing more formal care, such as interventions or references, have the potential to elicit greater harm than monitoring apps if these features are erroneous. Consumers are using these apps for treatment or diagnosis and are being exposed to the information provided by these health apps. Monitoring apps largely rely on data provided by the user rather than on the supply of novel information. The Pre-Cert's risk categorization's ability to differentiate between apps relying on formal medical advice versus user-led data and require a review from the former indicates its potential to catch apps that pose a greater risk.

When the data were stratified by targeted disease, the sample became underpowered, and we were unable to perform significance testing. However, although the small size of each subgroup is a limitation in this study, the observed trends offer valuable and novel insight into how the Pre-Cert Program's categorization should be refined before full implementation. For example, in the subgroup analysis, the percentage of apps offering an intervention differed dramatically between those exempt from review and those requiring a review for apps

XSL•FO
RenderX

targeting addiction, anxiety, and depression. This finding hints that the presence of an intervention is one of the strongest associations for apps requiring a review. However, this metric is challenging to reliably differentiate from apps that simply monitor symptoms. In particular, in the mental health setting, it has been established that individuals who monitor their symptoms feel better [36,37], and the line between ecological momentary assessments versus intervention is blurred with apps. Determining whether an app drives clinical management or provides treatment is pivotal for that app's risk categorization under Pre-Cert. Therefore, the FDA should set clear guidelines and give examples of what apps they consider provide an intervention versus monitor symptoms. Although we coded 24 features of apps (Figure 3) and (Table 1), only a few were found to differentiate between apps that would require FDA review and those that would have been streamlined. In addition, as described earlier, the strongest association between apps requiring a review that are offering an intervention remains to be difficult to clearly define. These results suggest that if the FDA wants to implement an appropriately detailed risk-based framework that can address the features of already-existing apps, they will need to publish more explicit guidelines and likely require extensive training for coders in order to obtain high interrater reliability and avoid possible misclassifications.

A specific criterion that the FDA should set more explicit guidelines around is the disclosure of apps' data policies. At present, the framework does not reflect if or how an app discloses how users can delete their information. For example, only 64% (37/58) of apps evaluated as likely to be exempt from review provided information about data deletion, although *cybersecurity responsibility* remains one of the Pre-Cert Program's 5 excellence principles. Apps requiring review levels of 2 or higher had a similar percentage of 55% (34/62). This result indicates that apps containing features and functions that pose legitimate risks on user privacy and data security are exempt from review under the Pre-Cert Program. Patients and clinicians who today use medical apps note *privacy and security* to be one of their top concerns with mobile health, meaning there is an opportunity for the FDA to be more explicit when assessing app developers' data policies.

The Pre-Cert Program and its risk categorization are still in their early developmental stages, as the FDA continues to test myriad aspects of the program. In an update summarizing testing performed through May 2019, the FDA described their refinement of this review determination process and admitted that further insight from patients and the digital health community is needed [23]. Our results can help inform the program in these pivotal early stages. To make the framework more useful, more data concerning its strengths and weaknesses are necessary. A novel challenge for evaluating apps is that their use case is not static and will vary based on the patients' clinical needs and treatment goals. For example, a mindfulness app may be a well-being tool, exempt from regulation in some contexts, but could also be recommended by clinicians in the treatment of major depressive disorders. This fluidity of purposes is different from that of a traditional medical device, for example, a pacemaker, which has a narrow and well-defined use among patients. In addition, further clarification as to what the FDA defines as a nonserious, serious, and critical condition and what regulators consider to be informing clinical care, driving management, or treating and diagnosing should be published. Finally, as the FDA plans to surveil an app's *real-world performance* in a postmarket setting, it will be important for them to publish guidelines of how they will ensure that not only the app's quality remains consistent but also that the developer continues to excel in the 5 criteria that granted them Pre-Cert status initially. Creating optimal systems is complex and will require the right combination of diverse stakeholder involvement. Thus, clinicians, patients, and leaders in the digital health community should be fully incorporated into this process and will likely welcome the opportunity to provide feedback.

Limitations

Our results must be interpreted in light of several limitations. First, we examined only 120 apps out of thousands that are currently marketed. As we took a convenience sample, there is the possibility that this sample does not reflect the top 10 apps presented to every consumer upon their search. In addition, at present, it is unclear which apps will need to be regulated with Pre-Cert and which will voluntarily partake—although as many health-related apps at present make clinical claims, many likely would fall under the scope of regulation. Second, our ratings were obtained by 2 reviewers and checked by a third reviewer. No third-party standards currently exist for determining how to score apps and to maintain validity, although we used published evaluation standards from previous research. Our research team only coded those features that could be verified, meaning that more subjective aspects of software products, such as app usability, were not coded. Third, we recognize that apps targeting certain disease conditions will have some inherent features and classifications. For example, apps targeting diabetes are more likely to integrate external devices than apps targeting other disease conditions because of their connection to blood glucose monitors. We attempted to minimize this effect by having a large sample of apps across a diverse range of conditions. In addition, the Pre-Cert's risk categorization will be used to classify these mobile health apps and will also need to account for these inherent features. Finally, we acknowledge that this risk-based framework is still in its testing phase and that revisions and additions will be made that are likely to increase the clarity of the criteria. Indeed, we expect this and applaud ongoing FDA efforts to pilot its framework and invite feedback from user communities.

Comparison With Prior Work

In the current state of digital health, the need to provide these more explicit guidelines is clear. There remains a lack of standards-based and reliable regulatory frameworks and evaluations that assess app quality, which, in turn, diminish consumers' confidence in digital health [38]. Researchers have attempted to bridge this gap by testing the reliability and accuracy of current rating systems [13,33,35] and deriving their own frameworks [34,38]. Our study fits into this landscape as we simulated using a proposed framework, the Pre-Cert Program's risk categorization, to classify top health apps. We provide novel information, as the Pre-Cert's framework has not been used on a sample of currently available apps. Thus, our

XSL•FO
RenderX

study identifies the potential of the program and areas for its improvement, which can inform other app classification initiatives. In addition, our studies and others similar to ours could potentially be used to surveil apps after they have reached the market. We used a peer-reviewed and published method that standardizes app selection and classifies apps based on 24 features. The iterations of this study over time can track if an app updates its features or changes its policies. Therefore, the community could play a role in the postmarket surveillance that FDA plans to implement in the Pre-Cert Program. We were limited by the fact that this risk categorization has not been previously studied as other rating systems. The lack of guidance and clarifications when discrepancies arose is an area of concern as digital health remains to be an evolving and sometimes ambiguous landscape.

Conclusions

The Pre-Cert Program's risk-based framework for assessing digital health apps and other SaMD products offers a promising foundation for enforcing appropriate digital health regulation while facilitating innovation and the use of technological advancements. However, the limited differences in our sample between apps likely requiring regulatory review and those that likely do not suggest that more detailed criteria are needed. We believe that additional exercises such as those done in this study, which can shed light on how the framework is likely to play out in the context of real-world digital health products, will be of high value. On the basis of such research, regulatory guidelines could be clarified and specified before the framework is deployed in the complex and dynamic landscape of digital health.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' Contributions

JT conceived the study. NA and student volunteers collected the data, and NA conducted the analysis, which was reviewed by JT. NA drafted the manuscript, and all authors contributed significantly to editing and the final version.

Conflicts of Interest

JT receives unrelated research support from Otsuka. Other authors declare no conflicts of interest.

Multimedia Appendix 1
Apps reviewed.
[DOCX File , 17 KB - mhealth_v8i10e20482_app1.docx]

References

1. Pohl M. 325,000 Mobile Health Apps Available in 2017: Android Now the Leading Mhealth Platform. Research 2 Guidance. 2017. URL: https://research2guidance.com/325000-mobile-health-apps-available-in-2017/ [accessed 2019-08-06]
2. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 7;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
3. Beard C, Silverman AL, Forgeard M, Wilmer MT, Torous J, Björgvinsson T. Smartphone, social media, and mental health app use in an acute transdiagnostic psychiatric sample. J Med Internet Res Mhealth Uhealth 2019 Jun 7;7(6):e13364 [FREE Full text] [doi: 10.2196/13364] [Medline: 31199338]
4. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of 'Mhealth' among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2016 Mar;42(2):448-455 [FREE Full text] [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
5. Robertson MC, Tsai E, Lyons EJ, Srinivasan S, Swartz MC, Baum ML, et al. Mobile health physical activity intervention preferences in cancer survivors: a qualitative study. J Med Internet Res Mhealth Uhealth 2017 Jan 24;5(1):e3 [FREE Full text] [doi: 10.2196/mhealth.6970] [Medline: 28119278]
6. Stawarz K, Preist C, Coyle D. Use of smartphone apps, social media, and web-based resources to support mental health and well-being: online survey. J Med Internet Res Ment Health 2019 Jul 12;6(7):e12546 [FREE Full text] [doi: 10.2196/12546] [Medline: 31301126]
7. Schueller SM, Neary M, O'Loughlin K, Adkins EC. Discovery of and interest in health apps among those with mental health needs: survey and focus group study. J Med Internet Res 2018 Jun 11;20(6):e10141 [FREE Full text] [doi: 10.2196/10141] [Medline: 29891468]
8. Thornton LK, Kay-Lambkin FJ. Specific features of current and emerging mobile health apps: user views among people with and without mental health problems. Mhealth 2018;4:56 [FREE Full text] [doi: 10.21037/mhealth.2018.11.04] [Medline: 30701174]
9. Vangeepuram N, Mayer V, Fei K, Hanlen-Rosado E, Andrade C, Wright S, et al. Smartphone ownership and perspectives on health apps among a vulnerable population in East Harlem, New York. Mhealth 2018;4:31 [FREE Full text] [doi: 10.21037/mhealth.2018.07.02] [Medline: 30221166]

XSL•FO
RenderX

10. Montagni I, Cariou T, Feuillet T, Langlois E, Tzourio C. Exploring digital health use and opinions of university students: field survey study. J Med Internet Res Mhealth Uhealth 2018 Mar 15;6(3):e65 [FREE Full text] [doi: 10.2196/mhealth.9131] [Medline: 29549071]
11. Torous J, Wisniewski H, Liu G, Keshavan M. Mental health mobile phone app usage, concerns, and benefits among psychiatric outpatients: comparative survey study. J Med Internet Res Ment Health 2018 Nov 16;5(4):e11715 [FREE Full text] [doi: 10.2196/11715] [Medline: 30446484]
12. Byambasuren O, Beller E, Glasziou P. Current knowledge and adoption of mobile health apps among australian general practitioners: survey study. J Med Internet Res Mhealth Uhealth 2019 Jun 3;7(6):e13199 [FREE Full text] [doi: 10.2196/13199] [Medline: 31199343]
13. Torous J, Firth J, Huckvale K, Larsen ME, Cosco TD, Carney R, et al. The emerging imperative for a consensus approach toward the rating and clinical recommendation of mental health apps. J Nerv Ment Dis 2018 Aug;206(8):662-666. [doi: 10.1097/NMD.0000000000000864] [Medline: 30020203]
14. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. NPJ Digit Med 2020;3:14 [FREE Full text] [doi: 10.1038/s41746-019-0212-z] [Medline: 32047860]
15. Policy for Device Software Functions and Mobile Medical Applications. Food and Drug Adminstration. 2019. URL: https://www.fda.gov/media/80958/download [accessed 2019-08-07]
16. Overview of Device Regulation. US Food and Drug Administration. 2020. URL: https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/overview-device-regulation [accessed 2020-03-01]
17. Digital Health Innovation Action Plan. US Food and Drug Administration. URL: https://www.fda.gov/media/106331/download [accessed 2019-08-07]
18. Developing Software Pre-Certification Program: A Working Model. US Food and Drug Administration. 2018 Jun. URL: https://www.fda.gov/media/113802/download [accessed 2019-08-07]
19. Software Precertification Program: 2019 Test Plan. US Food and Drug Administration. 2019 Jan. URL: https://www.fda.gov/media/119723/download [accessed 2019-09-25]
20. Niewood T. New Class of Therapeutics Calls for New Regulations: Thoughts on the Fda's New Process for Regulating Digiceuticals. Recon Strategy - Strategic Healthcare Partners. 2018 Nov 15. URL: http://reconstrategy.com/2018/11/new-class-of-therapeutics-calls-for-new-regulations-thoughts-on-the-fdas-new-process-for-regulating-digiceuticals/ [accessed 2019-08-08]
21. Peiris D, Miranda JJ, Mohr DC. Going beyond killer apps: building a better mHealth evidence base. Br Med J Glob Health 2018;3(1):e000676. [doi: 10.1136/bmjgh-2017-000676] [Medline: 29527353]
22. Gordon WJ, Stern AD. Challenges and opportunities in software-driven medical devices. Nat Biomed Eng 2019 Jul;3(7):493-497. [doi: 10.1038/s41551-019-0426-z] [Medline: 31278389]
23. Software Pre-Certification Program 2019 Mid-Year Update. US Food and Drug Administration. 2019. URL: https://www.fda.gov/media/129047/download [accessed 2019-08-08]
24. Software as a Medical Device. US Food and Drug Administration. 2018 Apr 12. URL: https://www.fda.gov/medical-devices/digital-health/software-medical-device-samd [accessed 2019-08-12]
25. Spanou D. Software as a Medical Device (SaMD): Key Definitions. International Medical Device Regulators Forum (IMDRF). 2103 Dec 9. URL: http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf [accessed 2020-03-01]
26. Global Approach to Software as a Medical Device Software as a Medical Device. US Food and Drug Administration. 2107 Jun 12. URL: https://www.fda.gov/medical-devices/software-medical-device-samd/global-approach-software-medical-device-software-medical-device [accessed 2019-08-07]
27. Digital Health Software Precertification (Pre-Cert) Program. US Food and Drug Administration. 2020 Sep 14. URL: https://www.fda.gov/medical-devices/digital-health/digital-health-software-precertification-pre-cert-program [accessed 2019-08-07]
28. Federal TC. FTC Releases New Guidance For Developers of Mobile Health Apps. Federal Trade Commission | Protecting America's Consumers. 2106 Apr 5. URL: https://www.ftc.gov/news-events/press-releases/2016/04/ftc-releases-new-guidance-developers-mobile-health-apps [accessed 2020-03-01]
29. ehealth Action Plan 2012-2020: Innovative Healthcare for the 21st Century. European Commission. 2012 Dec 7. URL: https://ec.europa.eu/digital-single-market/en/news/ehealth-action-plan-2012-2020-innovative-healthcare-21st-century
30. Classification of Digital Health Interventions v 1.0. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/260480/WHO-RHR-18.06-eng.pdf;jsessionid=EF5C7B012765D3D18F62E3B3C71E67DC?sequence=1 [accessed 2020-03-01]
31. Evidence Standards Framework for Digital Health Technologies. National Institute for Health and Care Excellence. 2019 Mar. URL: https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf [accessed 2020-03-01]
32. Wisniewski H, Liu G, Henson P, Vaidyam A, Hajratalli NK, Onnela J, et al. Understanding the quality, effectiveness and attributes of top-rated smartphone health apps. Evid Based Ment Health 2019 Feb;22(1):4-9 [FREE Full text] [doi: 10.1136/ebmental-2018-300069] [Medline: 30635262]

XSL•FO
RenderX

33. Carlo AD, Ghomi RH, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. NPJ Digit Med 2019;2:54 [FREE Full text] [doi: 10.1038/s41746-019-0129-6] [Medline: 31304400]
34. Henson P, David G, Albright K. Deriving a Practical Framework for the Evaluation of Health Apps. Lancet Digit Health. 2019. URL: https://doi.org/10.1016/S2589-7500(19)30013-5 [accessed 2020-09-23]
35. LeBeau K, Huey LG, Hart M. Assessing the quality of mobile apps used by occupational therapists: evaluation using the user version of the mobile application rating scale. J Med Internet Res Mhealth Uhealth 2019 May 1;7(5):e13019 [FREE Full text] [doi: 10.2196/13019] [Medline: 31066712]
36. Osborn J, Ajakaiye A, Cooksley T, Subbe CP. Do mHealth applications improve clinical outcomes of patients with cancer: a critical appraisal of the peer-reviewed literature. Support Care Cancer 2020 Mar;28(3):1469-1479 [FREE Full text] [doi: 10.1007/s00520-019-04945-4] [Medline: 31273501]
37. Walsh S, Golden E, Priebe S. Systematic review of patients' participation in and experiences of technology-based monitoring of mental health symptoms in the community. Br Med J Open 2016 Jun 21;6(6):e008362. [doi: 10.1136/bmjopen-2015-008362] [Medline: 27329437]
38. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. NPJ Digit Med 2019;2:38. [doi: 10.1038/s41746-019-0111-3] [Medline: 31304384]

Abbreviations

FDA: Food and Drug Administration
IMDRF: International Medical Device Regulators Forum
Pre-Cert: Software Precertification
SaMD: Software as a Medical Device

Edited by G Eysenbach; submitted 19.05.20; peer-reviewed by G Adrian, J Taylor, B Rudd; comments to author 29.06.20; revised version received 07.07.20; accepted 22.07.20; published 26.10.20.

Please cite as:
Alon N, Stern AD, Torous J
Assessing the Food and Drug Administration's Risk-Based Framework for Software Precertification With Top Health Apps in the United States: Quality Improvement Study
JMIR Mhealth Uhealth 2020;8(10):e20482
URL: http://mhealth.jmir.org/2020/10/e20482/
doi:10.2196/20482
PMID:32927429

©Noy Alon, Ariel Dora Stern, John Torous. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 26.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementation of a Home-Based mHealth App Intervention Program With Human Mediation for Swallowing Tongue Pressure Strengthening Exercises in Older Adults: Longitudinal Observational Study

HyangHee Kim[1,2], PhD, CCC-SLP; Nam-Bin Cho[1], BA; Jinwon Kim[3], MSc; Kyung Min Kim[1], BA; Minji Kang[1], MA; Younggeun Choi[3], MSc; Minjae Kim[3], BSc; Heecheon You[3], PhD; Seok In Nam[4], PhD; Soyeon Shin[1], BA

[1]Graduate Program in Speech-Language Pathology, Yonsei University College of Medicine, Seoul, Republic of Korea

[2]Department and Research Institute of Rehabilitation Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea

[3]Department of Industrial and Management Engineering, Pohang University of Science and Technology, Pohang, Republic of Korea

[4]Graduate School of Social Welfare, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
HyangHee Kim, PhD, CCC-SLP
Graduate Program in Speech-Language Pathology
Yonsei University College of Medicine
11-12 Yondaedongmun-gil
Seodaemun-gu
Seoul, 03721
Republic of Korea
Phone: 82 2 2228 3900
Email: h.kim@yonsei.ac.kr

Abstract

Background: Tongue pressure is an effective index of swallowing function, and it decreases with aging and disease progression. Previous research has shown beneficial effects of swallowing exercises combined with myofunctional tongue-strengthening therapy on tongue function. Tongue exercises delivered through mobile health (mHealth) technologies have the potential to advance health care in the digital age to be more efficient for people with limited resources, especially older adults.

Objective: The purpose of this study is to explore the immediate and long-term maintenance effects of an 8-week home-based mHealth app intervention with biweekly (ie, every 2 weeks) human mediation aimed at improving the swallowing tongue pressure in older adults.

Methods: We developed an mHealth app intervention that was used for 8 weeks (3 times/day, 5 days/week, for a total of 120 sessions) by 11 community-dwelling older adults (10 women; mean age 75.7 years) who complained of swallowing difficulties. The app included a swallowing monitoring and intervention protocol with 3 therapy maneuvers: effortful prolonged swallowing, effortful pitch glide, and effortful tongue rotation. The 8-week intervention was mediated by biweekly face-to-face meetings to monitor each participant's progress and ability to implement the training sessions according to the given protocol. Preintervention and postintervention isometric and swallowing tongue pressures were measured using the Iowa Oral Performance Instrument. We also investigated the maintenance effects of the intervention on swallowing tongue pressure at 12 weeks postintervention.

Results: Of the 11 participants, 8 adhered to the home-based 8-week app therapy program with the optimal intervention dosage. At the main trial end point (ie, 8 weeks) of the intervention program, the participants demonstrated a significant increase in swallowing tongue pressure (median 17.5 kPa before the intervention and 26.5 kPa after the intervention; P=.046). However, long-term maintenance effects of the training program on swallowing tongue pressure at 12 weeks postintervention were not observed.

Conclusions: Swallowing tongue pressure is known to be closely related to dysphagia symptoms. This is the first study to demonstrate the effectiveness of the combined methods of effortful prolonged swallowing, effortful pitch glide, and effortful tongue rotation using mobile app training accompanied by biweekly human mediation in improving swallowing tongue pressure in older adults. The mHealth app is a promising platform that can be used to deliver effective and convenient therapeutic service

XSL•FO
RenderX

to vulnerable older adults. To investigate the therapeutic efficacy with a larger sample size and observe the long-term effects of the intervention program, further studies are warranted.

International Registered Report Identifier (IRRID): RR2-10.2196/19585

(JMIR Mhealth Uhealth 2020;8(10):e22080) doi:10.2196/22080

KEYWORDS

mHealth; older adults; swallowing tongue pressure; Iowa Oral Performance Instrument; app; swallowing disorders; swallowing maneuver; effortful prolonged swallowing; effortful pitch glide; effortful tongue rotation

Introduction

The tongue plays an important role in both the oral and pharyngeal stages of swallowing, especially in proper bolus formation and propulsion [1]. Thus, its impairment is heavily linked to swallowing difficulties or dysphagia by inhibiting adequate bolus formation and transportation within the oral and pharyngeal cavities [2]. During liquid consumption, the bolus is initially held in the anterior part of the mouth or in the chamber between the tongue surface and the hard palate.

In older adults, sarcopenia, which is defined as loss of muscle mass due to aging, contributes to lingual dysfunction, resulting in swallowing impairments [3]. Previous studies have shown that sarcopenia and tongue muscle degeneration are highly related to aging [3-5] and negatively influence tongue pressure generation and swallowing in older adults, escalating the risk of aspiration and dysphagia [3,4,6].

The ability to generate tongue pressure is critical for liquid bolus propulsion in swallowing and is also an effective index for evaluating swallowing function [7,8]. Tongue pressure commonly decreases with aging [2,4], partly due to sarcopenia of the tongue [1,4]. Specifically, reduction in tongue pressure in older adults might cause an increase in oropharyngeal residue and consequently increase the risk of aspiration and pneumonia [7,9] due to reduction in proper bolus control and transfer [1]. Approximately 13% to 54% of older individuals have reported swallowing difficulties, with the rate varying with age, underlying diseases, and care level [10-12]. Major swallowing symptoms, such as coughing, choking [11,13], and malnourishment in older adults [6], are reportedly influenced by tongue dysfunction and tongue pressure measures [14-16].

Swallowing tongue pressure is generated when an individual swallows different types of food or liquid or during dry swallowing. The decrease in swallowing tongue pressure in older adults is prevalent [7,17] depending on the types and the viscosity of the intake [1,7,17,18] and during dry swallowing [7]. As the decline in tongue pressure due to aging may be indicative of risk factors related to dysphagia [17], it is crucial to pay attention to swallowing tongue pressure measures in order to incorporate swallowing rehabilitation.

Past studies have shown beneficial effects of swallowing exercises combined with myofunctional tongue-strengthening exercises in older adults, as well as in adults with various types of disease. Among diverse rehabilitative techniques using swallowing interventions, the Mendelsohn maneuver [19] and effortful swallowing are swallowing-focused methods. The implementation of both treatment methods has shown a positive effect on tongue pressure [20], as well as on pharyngeal [21] and hyolaryngeal functions [22]. Moreover, the combination of the two methods has led to increased anterior and posterior tongue pressure [23] and decreased aspiration rates in dysphagic patients [24].

Meanwhile, indirect methods of swallowing treatment include effortful pitch glide and effortful tongue rotation. Effortful pitch glide does not invoke the principle of task specificity for swallowing [25,26]. However, when effortful pitch glide was combined with the Mendelsohn maneuver, its positive effect on the activation of suprahyoid muscles, the functioning of which is regarded as essential in hyolaryngeal excursion (and thus in swallowing), was observed [27]. Such results may also account for the observed coordination of suprahyoid muscle activity, hyoid movement, and tongue pressure production during tongue squeezing [28]. A significant increase in maximum tongue pressure was also observed after effortful tongue rotation [29] and directional exercises of the tongue (ie, elevation, protrusion, and lateralization) [30].

Repetition and intensity of training are key components needed to facilitate experience-dependent neural plasticity in motor behaviors [25] and swallowing [26]. Training with evident goals, however, requires participants to regularly visit swallowing clinics. Several studies have provided information on factors associated with the difficulty of ensuring regular visits to clinics among older adults, including expenses, distance, transportation, and the time required [31,32]. Therefore, mobile health (mHealth) technologies may have great potential in advancing health care in the digital age to be more efficient for people [33] with limited resources, especially for older adults.

Thus, this study explores the immediate and long-term maintenance effects of an 8-week home-based mHealth app intervention program with human mediation aimed at improving the swallowing tongue pressure in older adults. We further highlight several clinical issues raised during the intervention program that needed to be addressed. This study was part of an ongoing larger 7-year national grant project entitled “Development and Implementation of Swallowing Evaluation and Intervention Program for Older Adults.”

Methods

Participants

The participants in this study were identical to those who took part in the usability study of the mHealth app [34]. The data for the usability evaluation [34] and the current study (ie, swallowing training results) were collected from these

XSL•FO
RenderX

participants in the same time frame. A total of 11 older adults (10 women; mean age 75.7 years, SD 4.96; mean education 10.4 years, SD 3.83; mean Korean Mini-Mental State Examination score 28.1, SD 1.72) from 2 district-run senior welfare centers in Seoul participated in the program. The inclusion criteria were (1) older than 65 years; (2) complaints of swallowing difficulties (ie, increased aspiration rate and foreign body sensation in throat); (3) a Korean Mini-Mental State Examination score within normal limits; (4) no problem with vision, hearing, and motor functions required for using a tablet; and (5) no hearing problems that may prevent the participant from following directions. The exclusion criteria were (1) any history of neurological disorders (ie, stroke and Parkinson disease) and (2) nonoral feeding (ie, nasogastric tube and percutaneous endoscopic gastrostomy). Self-reported swallowing rating was indicated as responses to the survey item "I have a swallowing problem" (with a scale of 0=never, 1=seldom, 2=sometimes, 3=often, and 4=always). Table 1 presents the characteristics of the participants. All participants gave informed, written consent as per the Declaration of Helsinki before conducting the study (institutional review board No. PIRB-2019-E024).

Table 1. Participant characteristics.

Participant No.	Gender	Age (years)	Education (years)	K-MMSE[a] score	Self-reported swallowing rating[b]
1	Female	80	6	30	3
2	Female	71	16	29	3
3	Female	67	12	30	2
4	Female	83	9	24	4
5	Female	77	12	27	3
6	Female	71	9	29	2
7	Female	82	6	29	3
8	Female	73	12	28	3
9	Female	76	5	27	2
10	Male	75	16	28	2
11	Female	78	6	29	2
Total, mean (SD)	N/A[c]	75.73 (4.96)	9.45 (4.80)	28.18 (1.72)	2.64 (0.67)

[a]K-MMSE: Korean Mini-Mental State Examination. A perfect score is 30.

[b]Response to the survey item "I have a swallowing problem" (with a scale of 0=never, 1=seldom, 2=sometimes, 3=often, and 4=always).

[c]N/A: not applicable.

App Therapy Contents

The app contents included a swallowing monitoring and intervention protocol with 3 therapeutic methods or maneuvers: (1) effortful prolonged swallowing, which incorporated effortful swallowing [35] with the Mendelsohn maneuver [19,36], (2) the effortful pitch glide exercise [37], and (3) effortful tongue rotation [29]. The 3 effortful maneuvers were named A Successful Swallow with Effortful Trainings (ASSET). Methods for each training maneuver used in this study are shown in Table 2.

Participants were required to perform a total of 120 sessions (3 times/day, 5 days/week for 8 weeks). A 1-day session involved 20 repetitions of each of the 3 exercises (effortful prolonged swallowing, effortful pitch glide, effortful tongue rotation) in the morning, afternoon, and evening, with each time-specific session taking approximately 30 minutes. If the individuals failed to conduct the exercises on time, they were allowed to make up for the missed sessions on the same day only.

XSL•FO
RenderX

Table 2. Procedures for each training maneuver.

Exercise	Procedure
Effortful prolonged swallowing	(Phase 1: water swallow) Step 1. Hold 5 mL of water in your mouth. Step 2. Push all muscles around your neck and swallow as hard as you can (at this time, maintain your muscular strength for 2 seconds). (Phase 2: dry swallow) Step 1. Collect saliva in your mouth. Step 2. Push all muscles around your neck and swallow as hard as you can (at this time, maintain your muscular strength for 2 seconds).
Effortful pitch glide	Step 1. Make an elongated “ee” sound at a comfortable pitch. Step 2. Gradually raise your pitch to the highest pitch possible. Step 3. Keep the pitch at the highest range as long as possible.
Effortful tongue rotation	Step 1. Stretch your tongue and move it to one side of your cheek. Step 2. Rotate your tongue fully around the space between your teeth and back of the lips once while maintaining your strength for 5 seconds.

App Features

We developed a tablet-based mHealth app (Android operating system) called 365 Healthy Swallowing Coach, which enables older adults to execute the given 3 types of swallowing training maneuvers without a clinician physically present. It contains an educational program, a feedback system, an adherence monitoring system, and a tailored training setting. The mobile app was designed to be downloaded and used on a tablet, the Samsung Galaxy Tab A (model No. SM-P580; Samsung), and to be user-friendly for older adults. Examples of user interface designs for older adults with decreased cognitive and psychomotor functions include (1) tabs arranged by use sequence for ease of navigation, (2) buttons containing both icons and text for ease of understanding, and (3) apposite button activation ranges, button sizes, and interbutton spaces empirically identified for the trade-off between ease of activation and prevention of inadvertent activation [34]. Furthermore, the mHealth app supports a variety of features that could enable older adults to properly and effectively train themselves. Examples of training compliance enablers include videos, animations, multimodal content, a biofeedback system for ease of understanding [38], and an automatic training data-logging system and training scheduler for ease of adherence monitoring [34]. A usability and feasibility test of the 365 Healthy Swallowing Coach app reported significantly higher usability scores from older adults with higher levels of education and smart device usage. Moreover, an increase in favorable responses regarding app acceptability, training program utilization, emotional responses, and learning experience with the intervention program was observed as the participants continued to use the app [34].

Educational Program

The educational content of the 365 Healthy Swallowing Coach app includes introductory videos on the basics of swallowing mechanisms and swallowing disorders, as well as training methods for each exercise (ie, effortful prolonged swallowing, effortful pitch glide, and effortful tongue rotation). Information on the basics of swallowing is introduced through a self-produced animation in the current study. The animation educates users on the physiological changes that occur during swallowing, such as epiglottic closure for airway protection, and promotes awareness on the danger of aspiration and pneumonia. The educational video for each exercise method consists of a demonstration carried out by a researcher and a practice trial. Users can practice each method after observing the demonstrations, which include a live-action video for effortful prolonged swallowing and effortful tongue rotation or a real-time pitch glide graph for effortful pitch glide. These videos are always shown prior to the beginning of the actual training sessions to ensure that users perform the exercises in an appropriate way. Screenshots of the educational content are shown in Figures 1 and 2.

XSL•FO
RenderX

Figure 1. A screenshot of an example introductory video on the basics of swallowing mechanisms (describing hyolaryngeal excursion and epiglottic closure) and dysphagia. The caption in the video states, “At this point, the epiglottis closes as the hyoid bone and the larynx rise, protecting the airway.”

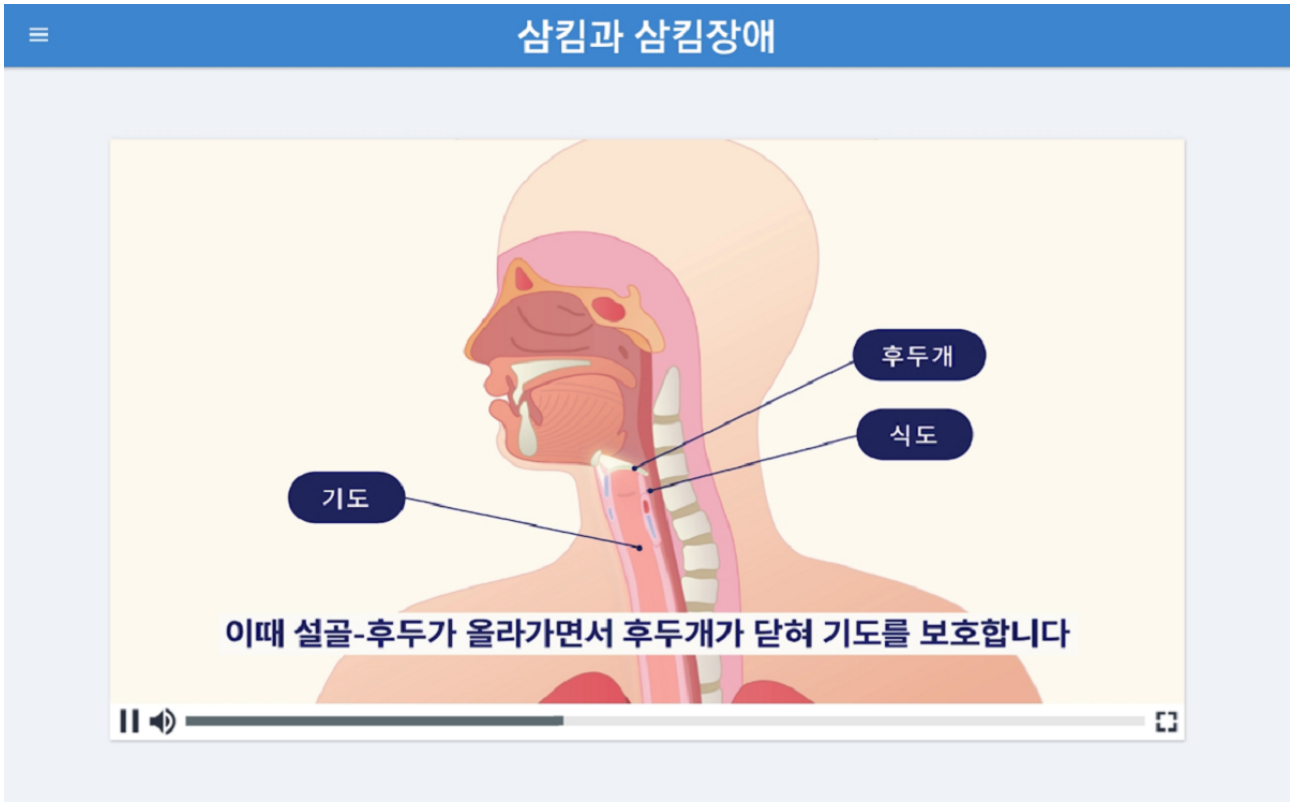


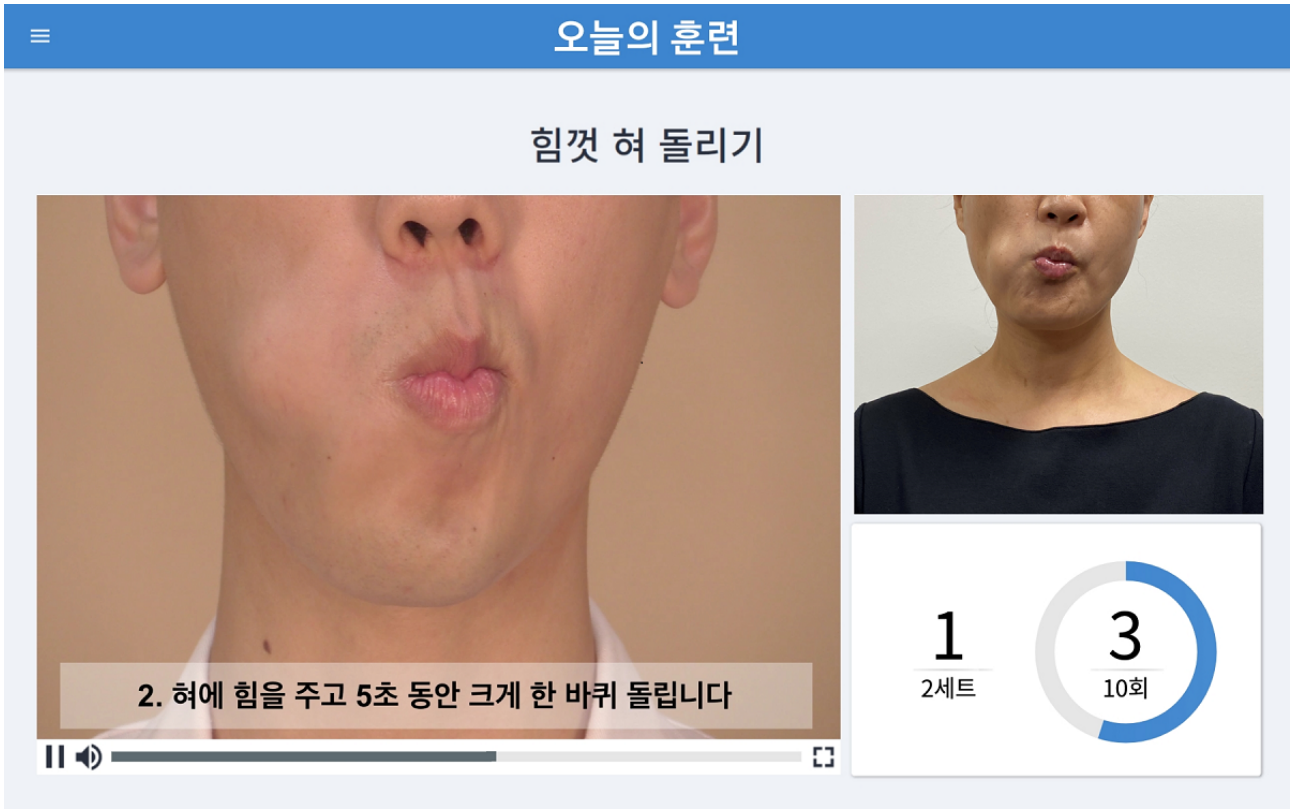
Figure 2. A screenshot of the educational demo video of the effortful prolonged swallowing exercise maneuver in the 365 Healthy Swallowing Coach app, with training procedures (top right) and preparation materials (bottom right). The caption in the video says, “First, hold 5 mL of water in your mouth.”

XSL•FO
RenderX

Biofeedback System

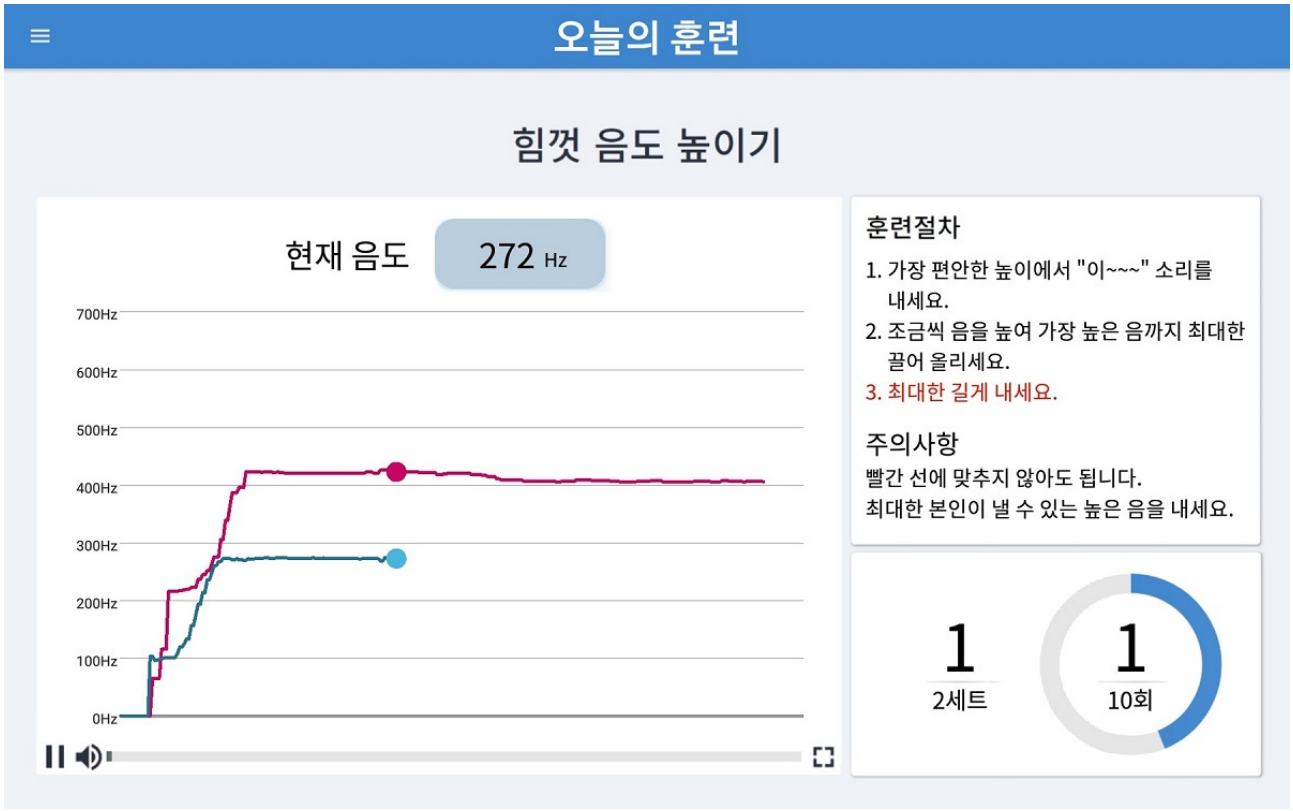
While the user exercises, the app provides instant visual biofeedback. For the effortful prolonged swallowing and effortful tongue rotation exercises, the screen displays a demo video and a mirror function on the same screen so that users can follow the demonstration while watching themselves train, as seen in Figure 3. The screenshot of the effortful pitch glide exercise displays a real-time pitch glide graph with 2 colors (Figure 4). The red line represents the pitch of the demonstration audio, and the blue line represents the user's pitch. Users glide and adjust their pitch according to the shape of the red line. The biofeedback system allows users to appropriately perform the exercise methods each time by providing instant feedback regarding their performance.

Figure 3. A screenshot of the visual biofeedback in the "Today's Training" section, with the effortful tongue rotation demo video displayed on the left, the mirror function on the right, and the number of repetitions on the bottom right. The caption on the screen says, "2. Rotate your tongue fully around the space between your teeth and the back of the lips once while maintaining your strength for 5 seconds."



XSL•FO
RenderX

Figure 4. A screenshot of the visual biofeedback in the "Today's Training" section during the effortful pitch glide maneuver, with a real-time pitch graph on the left, the training procedures on the right, and the number of repetitions indicated on the bottom right.

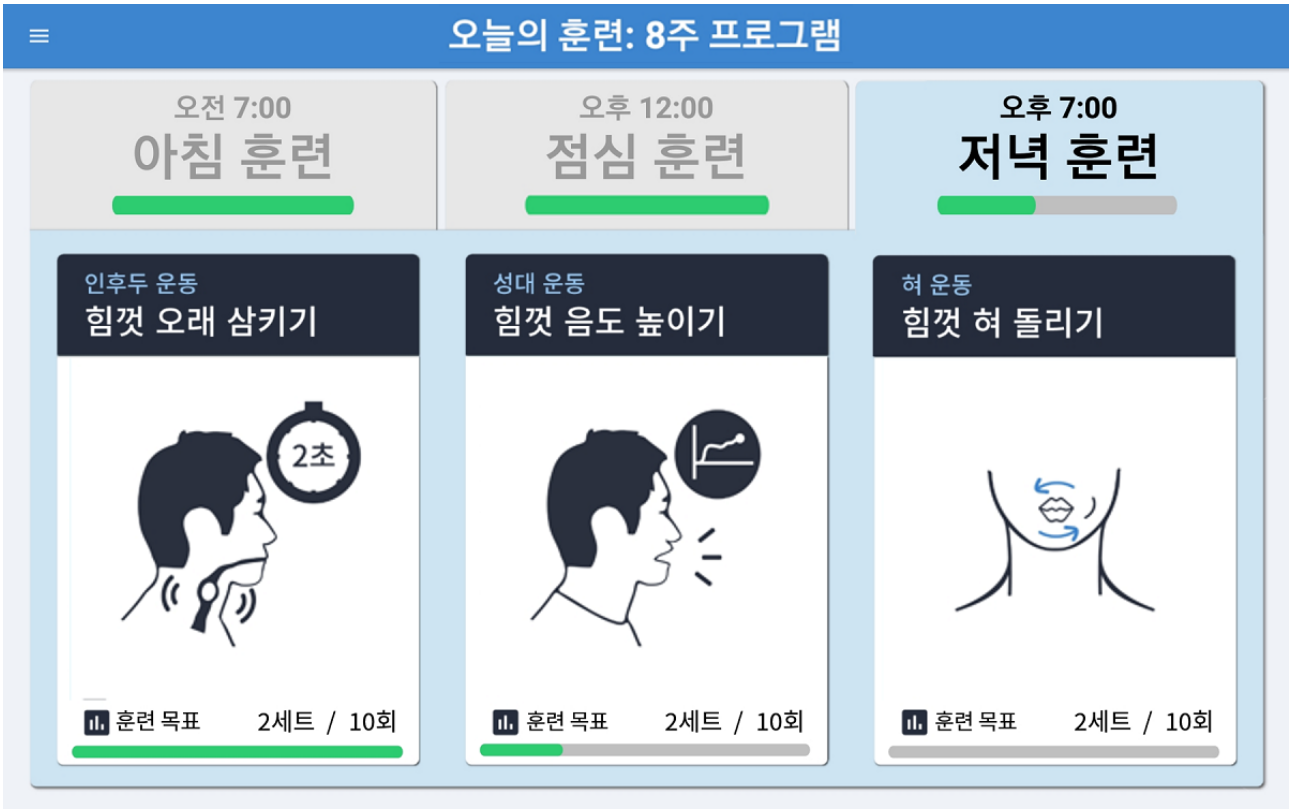


Adherence Monitoring System

Users can monitor their progress through an automatic data-logging system that guides them through the performance of rigorous daily sessions. Figure 5 shows a screenshot of the exercise selection page for the swallowing function training. The exercise selection screen and the exercise record screen of the app show the time-specific and daily progress, respectively. The exercise selection screen initially shows empty progress bars on the selection tabs of each time-specific training session (ie, morning, afternoon, and evening training) on the top portion of the screen and on each exercise (ie, effortful prolonged swallowing, effortful pitch glide, and effortful tongue rotation) on the bottom portion of the screen (Figure 5). Meanwhile, the overall progress can also be checked on the exercise record screen, where a monthly calendar is provided, with the daily performance rate marked with a green (performance rate of 100%), orange (50% to 99%), or red (less than 50%) circle around a specific date. By using the adherence monitoring system of the mHealth app, users can constantly check their progress, which can help minimize the number of missed exercises.

XSL•FO
RenderX

Figure 5. A screenshot of the exercise selection section (for the swallowing function training) of the 365 Healthy Swallowing Coach app, with green bars indicating exercise- and time-specific progress.



Procedures

A flowchart of the 8-week home-based intervention study with biweekly (ie, every 2 weeks) human mediation is presented in Figure 6. Two weeks before the beginning of the intervention program, participants' demographic information and medical histories were collected, and their isometric and swallowing tongue pressures were measured by a clinician. Subsequently, they were educated on how to operate the designated tablet and the swallowing function training contents of the 365 Healthy Swallowing Coach app before starting their home training using the mobile device (orientation stage). Upon commencing the intervention program, the researchers held biweekly face-to-face meetings with each of the participants for the purpose of checking the participant's ability to carry out the training sessions without assistance and perform the exercises in accordance with the given protocol. In this way, the researchers acted as evaluators who checked and monitored the training status of the participants. However, since the primary purpose of the face-to-face meetings was strictly confined to monitoring the participants' progress, the researchers minimized giving advice that might directly influence their self-administered performance. Each face-to-face meeting was held for approximately 30 minutes. During the meetings, the participants were asked to navigate through the app (ie, from the main screen of the tablet to the training selection screen) and perform 1 training session as an example while the researchers monitored their presentation. If the participants failed to operate the app properly or performed training inaccurately, they were advised by the researchers to try to correct themselves. After each participant finished a training session and closed the app, the researchers then checked the participant's progress by reviewing the sessions tracked in their tablet.

At postintervention (ie, immediately after the 8-week intervention), the participants' isometric and swallowing tongue pressures were evaluated, as previously performed during the preintervention evaluation stage. No additional intervention or intervention-related mediation took place after the postintervention evaluation stage. Then, at 12 weeks postintervention (ie, 20th week from the beginning of the intervention program), the participants' isometric and swallowing tongue pressures were reevaluated in an identical manner as in the preintervention and postintervention evaluation stages in order to evaluate the maintenance effect of the intervention program.

Swallowing and isometric tongue pressures were measured using the Iowa Oral Performance Instrument (IOPI) (model 2.1; IOPI Medical LLC). The IOPI is a handheld device with a connecting tube and a measurement bulb (Figure 7) used to measure tongue pressure against the hard palate in an air-filled bulb.

XSL•FO
RenderX

Figure 6. Flowchart of the 8-week home-based intervention and maintenance study.

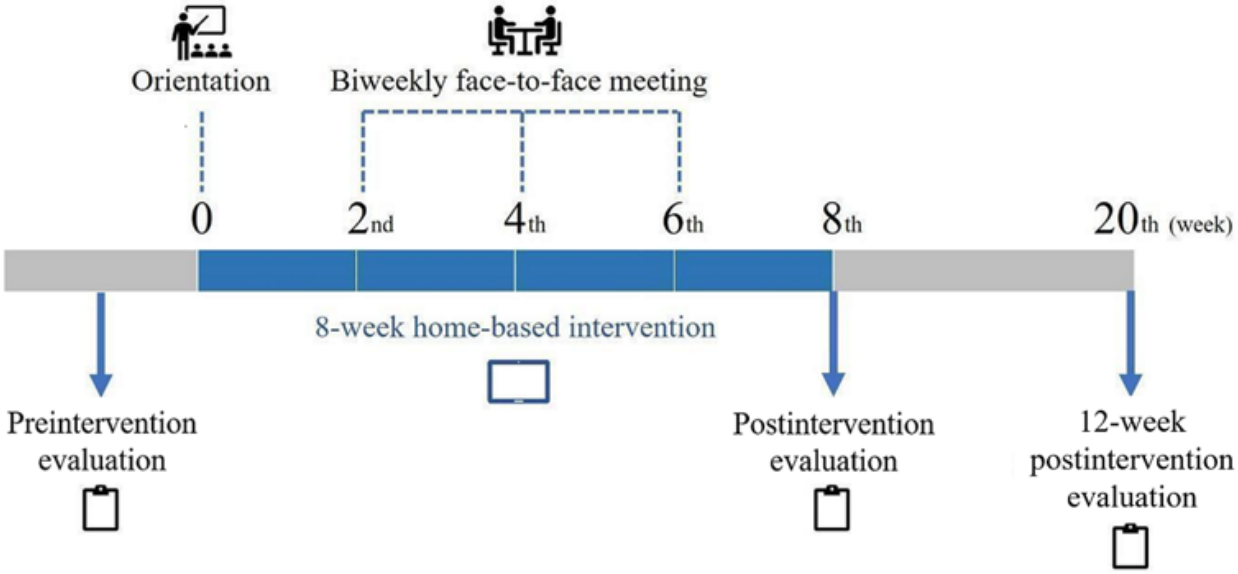
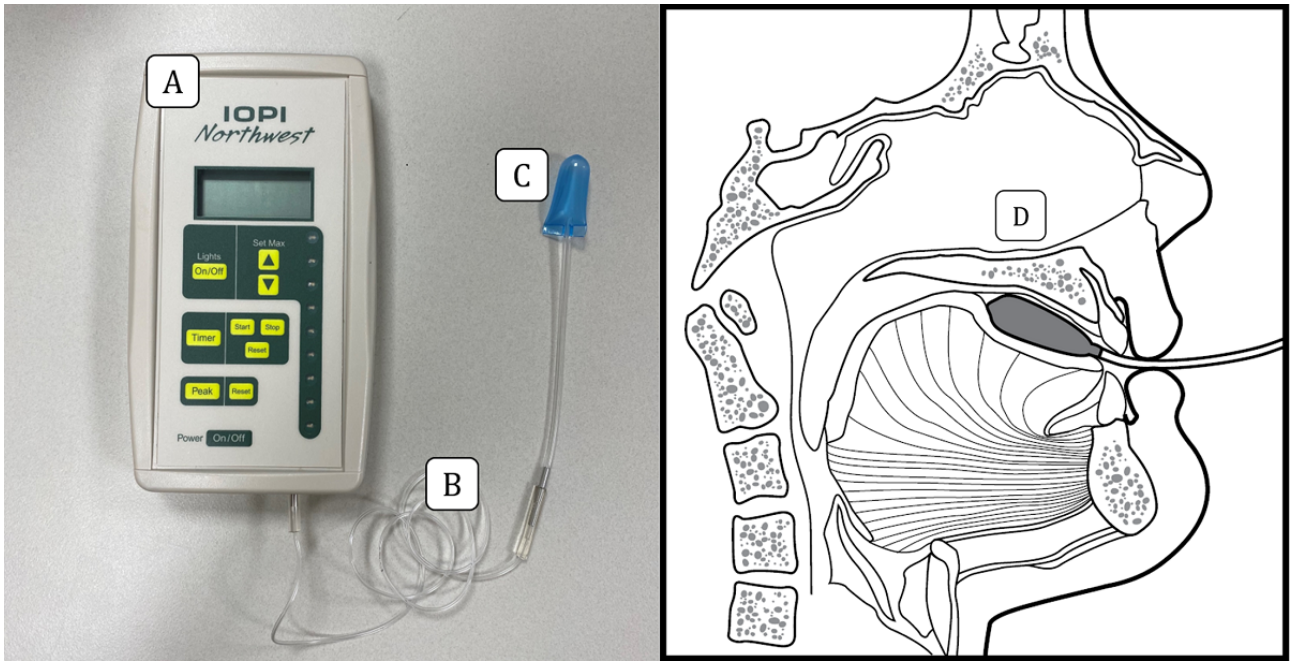


Figure 7. The Iowa Oral Performance Instrument components and the tongue bulb position. (A) Iowa Oral Performance Instrument model 2.1, (B) connecting tube, (C) air-filled tongue bulb, and (D) the position of the tongue bulb against the hard palate [39] (cited with permission).



The evaluations were conducted by the authors HK, N-BC, and SS. All participants had a verbal training session before having their perioral pressure recorded. During the tongue pressure measurement, the participants were in a comfortable and relaxed upright seated position to achieve a natural head posture, and the researchers oversaw their posture and the bulb position. When measuring the swallowing tongue pressure, the participants were asked to swallow the saliva as naturally as possible while the tongue bulb was placed between their tongue and the anterior one-third of their hard palate. When measuring the isometric tongue pressure, they were asked to push the bulb with maximum force toward their hard palate with their tongue for 2 seconds.

Each measurement was repeated 2 times. The pressure was recorded in kilopascals and the maximum value was retrieved. For swallowing tongue pressure, the participant was asked to perform a dry swallow [7]. Tongue pressure measurements were recorded before and after the intervention, as well as at 12 weeks postintervention. Immediate postintervention evaluation and 12-week postintervention evaluation findings were compared with preintervention evaluation findings to identify the short-term and long-term effects of the given therapy sessions. Statistical analysis was conducted using the Wilcoxon signed rank test (SPSS version 25; IBM Corp).

XSL•FO
RenderX

Results

Adherence to Intervention Program

The adherence of all 11 participants was evaluated at the end of the 8-week intervention program. Table 3 presents the adherence of each participant. Out of the 11 participants, 8 completed the 8-week program with the proper amount of training suggested by the researchers (adherence of 83.2%-100.8%) and were categorized as compliant participants. Only these participants were included in the data analysis. Meanwhile, 3 participants (No. 9, No. 10, and No. 11) labeled as noncompliant participants were excluded from the analysis due to failure to correctly adhere to the program protocol. Participants No. 9 and No. 10 were excluded because they completed only 35.4% and 41.1% of the required sessions, respectively. The 2 participants reported that they did not have enough time to carry out the training sessions and that they could not perform each training session in accordance with the given protocol due to the excessive amount of training. Meanwhile, participant No. 11 was excluded due to greatly exceeding (264.7%) the amount of training required. She reported that she did not understand the adherence monitoring system of the app due to unfamiliarity with the technology, resulting in an excessive amount of training.

Table 3. Adherence of each participant.

Participant classification and number	Adherence, %			
	Effortful prolonged swallowing	Effortful pitch glide	Effortful tongue rotation	Total
Compliant participants				
1	93.0	93.7	92.9	93.2
2	101.3	98.7	100.0	100.0
3	97.5	97.4	97.4	97.4
4	85.3	82.2	82.1	83.2
5	101.0	100.3	97.2	99.5
6	93.8	91.6	89.6	91.7
7	96.3	93.7	95.7	95.2
8	100.1	101.1	101.3	100.8
Noncompliant participants				
9	34.9	37.0	34.2	35.4
10	48.5	37.8	37.1	41.1
11	252.7	305.2	236.3	264.7

Intervention Effect: Preintervention Versus Postintervention

The 8 participants who properly completed the 8-week training protocol, which was mediated by 3 sessions of biweekly meetings, demonstrated statistically significant increases in swallowing tongue pressure postintervention (median 26.5 kPa) compared with the preintervention (median 17.5 kPa) values ($z=–1.992$, $P=.046$). On the other hand, maximum isometric tongue pressures were not statistically different between the postintervention (median 41 kPa) and preintervention (median 40.5 kPa) evaluation ($z=–0.632$, $P=.53$). Table 4 and 5 show ranks and test statistics of swallowing and isometric tongue pressure before and after the intervention.

XSL•FO
RenderX

Table 4. Ranks and test statistics of swallowing tongue pressure before and after the intervention.

Swallowing tongue pressure	Participants, n	Pressure (kPa), mean (SD)	Pressure (kPa), range	Pressure (kPa), median	Mean rank	Sum of ranks	*P* value[a]	z[b]
Intervention							.046	–1.992
Preintervention	8	18.3 (6.5)	7-27	17.5	N/A[c]	N/A		
Postintervention	8	27.4 (9.4)	12-41	26.5	N/A	N/A		
Rank: post-pre							N/A	N/A
Negative ranks	1[d]	N/A	N/A	N/A	1	1		
Positive ranks	5[e]	N/A	N/A	N/A	4	20.5		
Ties	2[f]	N/A	N/A	N/A	N/A	N/A		
Total	8	N/A	N/A	N/A	N/A	N/A		

[a]Wilcoxon signed ranks test.

[b]Based on negative ranks.

[c]N/A: not applicable.

[d]Postintervention<preintervention.

[e]Postintervention>preintervention.

[f]Postintervention=preintervention.

Table 5. Ranks and test statistics of isometric tongue pressure before and after the intervention.

Isometric tongue pressure	Participants, n	Pressure (kPa), mean (SD)	Pressure (kPa), range	Pressure (kPa), median	Mean rank	Sum of ranks	*P* value[a]	z[b]
Intervention							.53	–0.632
Preintervention	8	40.3 (5.1)	29-46	40.5	N/A[c]	N/A		
Postintervention	8	41.0 (8.0)	29-52	41.0	N/A	N/A		
Rank: post-pre							N/A	N/A
Negative ranks	3[d]	N/A	N/A	N/A	4.5	13.5		
Positive ranks	5[e]	N/A	N/A	N/A	4.5	22.5		
Ties	0[f]	N/A	N/A	N/A	N/A	N/A		
Total	8	N/A	N/A	N/A	N/A	N/A		

[a]Wilcoxon signed ranks test.

[b]Based on negative ranks.

[c]N/A: not applicable.

[d]Postintervention<preintervention.

[e]Postintervention>preintervention.

[f]Postintervention=preintervention.

Long-Term Effect: Preintervention Versus 12-Week Postintervention

Of the 8 participants, 6 (No. 1, No. 2, No. 3, No. 5, No. 7, No. 8) agreed to participate in the 12-week postintervention evaluation for examination of maintenance effects. Compared with preintervention (median 19 kPa), there was no statistically significant difference in swallowing tongue pressure at the 12-week postintervention (median 20.5 kPa) (z=–0.734, *P*=.46). Table 6 presents ranks and test statistics of swallowing tongue pressure measured at the preintervention and 12-week postintervention evaluation stages.

XSL•FO
RenderX

Table 6. Ranks and test statistics of swallowing tongue pressure before the intervention and 12 weeks after the intervention.

	Participants, n	Pressure (kPa), mean (SD)	Pressure (kPa), range	Pressure (kPa), median	Mean rank	Sum of ranks	*P* value[a]	z[b]
Intervention							.46	–0.734
Preintervention	6	18.8 (7.5)	7-27	19.0	N/A[c]	N/A		
12-week postintervention	6	24.8 (14.0)	12-45	20.5	N/A	N/A		
Rank: 12-week post–pre							N/A	N/A
Negative ranks	3[d]	N/A	N/A	N/A	2.33	7		
Positive ranks	3[e]	N/A	N/A	N/A	4.67	14		
Ties	0[f]	N/A	N/A	N/A	N/A	N/A		
Total	6	N/A	N/A	N/A	N/A	N/A		

[a]Wilcoxon signed ranks test.

[b]Based on negative ranks.

[c]N/A: not applicable.

[c]12-week postintervention<preintervention.

[d]12-week postintervention>preintervention.

[e]12-week postintervention=preintervention.

Discussion

Overview

To the best of our knowledge, this investigation is the first training study addressing an mHealth tongue-strengthening exercise for older adults. As tongue pressure measures have been determined to be effective indicators of swallowing function [7,8], an increase in tongue pressure may positively influence swallowing function as well. Decreases in swallowing tongue pressure generated during dry and high-viscosity liquid or bolus swallowing is prevalent in older adults [7,17]. Since such decline may be indicative of risk factors of dysphagia, looking at measures of swallowing tongue pressure in order to improve swallowing rehabilitation endeavors is a compelling pursuit [17]. In this study, a tablet-based swallowing training app with biweekly human mediation was used for 8 weeks by older adults who complained of swallowing difficulties. All 8 participants who completed the training demonstrated an increase in swallowing tongue pressure, even though therapists were not physically present during the training sessions. It should be noted that biweekly face-to-face meetings were held between the participants and the researchers for the main purpose of checking and monitoring the participants' training progress, and this process may have affected the end results (ie, increased swallowing tongue pressure) of the intervention program. There are several important issues that must be recognized regarding these study findings.

Exercise Maneuvers

Previous studies have shown that various tongue-strengthening exercises resulted in an increase in tongue strength and pressure [40-44]. Meanwhile, the current intervention contents (ASSET) adopted 3 therapeutic methods or maneuvers (ie, effortful prolonged swallowing, effortful pitch glide, and effortful tongue rotation) in combination, which led to positive effects on the swallowing tongue pressure of the participants, as indicated by the study results. However, several differences have been observed in treatment effects upon comparison of the current study to those that adopted a single therapy maneuver.

First, the current study implemented effortful prolonged swallowing, a combination of effortful swallowing training and the Mendelsohn maneuver training. In previous studies, training outcomes of effortful swallowing alone only revealed significant increases in the maximum isometric pressure of the anterior tongue, lingua-palatal pressures, and swallowing functions [41,45]. However, in this study, effortful prolonged swallowing positively affected swallowing tongue pressure. The functions of the extrinsic tongue muscles (ie, hyoglossus), which were shown to be active during the Mendelsohn maneuver [46], may have improved after therapy, thus benefiting the participants and positively affecting their swallowing tongue pressure.

Second, effortful tongue rotation is regarded as a promising training method, since swallowing tongue pressure is closely related to dysphagia symptoms such as aspiration [9] and oropharyngeal residue [47]. A previously conducted study observed a significant increase in maximum isometric tongue pressure after effortful tongue rotation [29]. Interestingly, however, only the swallowing tongue pressure and not the maximum isometric tongue pressure was enhanced in this study when the isokinetic effortful tongue rotation exercise was implemented.

Partial discrepancies observed among the previous and current results may be ascribed to the fact that the triple-combination therapy (ASSET) of effortful prolonged swallowing, effortful pitch glide, and effortful tongue rotation was employed in this study. Hence, the combination effect is not yet comparable to that of a single therapy maneuver. Consequently, the results from this study protocol are not indicative of the individual effect of each treatment method on posttreatment measures, since the treatment effects of each of these methods were not evaluated separately. Therefore, it cannot be concluded that the

XSL•FO
RenderX

additive effect of the triple-combination therapy as a whole is equal or similar to the arithmetic addition of each individual exercise's training effect. The effect that the 3 combined methods yield may be smaller than the effect that each method provides, or it may be greater [48] than the simple addition of individual effects.

Optimal Dosage

Even though a significant effect on swallowing tongue pressure was obtained from this intervention program, it remains unclear if the actual cumulative exercise dosage (ie, intensity, frequency, and duration) employed in this study is significant [49]. According to a scoping review, a tongue exercise dosage of 3 times a day for 7 days with a frequency of 30 repetitions for each exercise was the most widely implemented regimen for studies including single tongue exercises [50]. However, for studies that adopted combined exercises (ie, more than one therapy method), treatment dosage, frequency, and duration varied widely, and thus, the improvements in various outcome measures could not be concluded to be attributable to the use of particular treatment dosages [50].

Determining the optimal exercise dosage of a treatment program is crucial, as overtraining may contribute to exercise-induced fatigue [51] and high dropout rates [52], whereas undertraining may not result in any benefit [53]. In fact, some participants in this study complained of the excessive training load and stated that they could not complete the intervention program.

The "active ingredients" of interventions need to be identified in order to determine the optimal intervention dosage and intensity [49]. The active ingredients are components in an intervention that can influence the treatment results. Client-related variables, including life situations or circumstances, or therapy-related variables, such as the complexity of treatments and the treatment duration [54], may also influence the optimal therapeutic dosages. Since the dosage stated in repetitions (eg, 20 exercise repetitions per session) indicates how many times an "active ingredient" appears per training session, it may be important to represent exercise dosage in such a form when determining the optimal intervention dosage for speech language pathology–related fields [49].

Choosing to adopt a distributed intervention or an intensive intervention is also critical in determining the ideal intervention dosage and intensity. Previous studies have adopted intervention programs for various speech and language disorders with different levels of intensity [55-57]. One study that compared the intervention results of a distributed (8 weeks, 6 hours per week, for a total of 48 hours) and intensive (3 weeks, 16 hours per week, for a total of 48 hours) aphasia intervention showed that the distributed intervention derived greater therapeutic gains compared with the intensive treatment [55]. The protocol duration implemented by this study is comparable to the distributed intervention adopted in the aforementioned study because the total intervention duration of this mHealth program was also 8 weeks (20 repetitions of each exercise, totaling 60 repetitions/session, 3 sessions/day, 5 times/week). However, the exact dose per session, stated as the number of repetitions, was not indicated in the aphasia intervention study. Thus, an exact one-to-one comparison is not possible.

Long-Term Effects

The current study measured the swallowing tongue pressure at baseline and at 2 follow-up time points, including at 20 weeks after the intervention commenced (ie, 12-week postintervention), in order to investigate long-term intervention effects. Tracking changes in gained muscle (ie, tongue) pressure after the removal of an intervention is of particular interest, since a consideration of detraining effects may be necessary for the effective planning of aftercare [30]. In the current study, the training effect in the swallowing tongue pressure observed at the termination of the 8-week intervention program did not remain upon measuring it at 12 weeks postintervention. Despite the clinical implications of detraining effects, however, studies that examine the long-term effects of swallowing interventions are largely lacking [30,58]. Past studies show inconsistent results in posttreatment decay of lingual strength. Discrepancies in the findings from previous studies may be ascribed to several methodological differences, including (1) the postintervention measurement points, (2) the age of the participants, (3) the baseline health conditions of the participants, (4) the intervention delivery method, and (5) the type of outcome measure used.

The first factor is the postintervention measurement points. When the current results are compared with those of past studies, the long-term efficacy of tongue-strengthening exercises does not seem to show a consistent correlation with the measurement point alone. For instance, significant decreases in tongue strength have been observed after 2 to 4 weeks [30] and after 12 weeks (in the current study) of detraining, while no changes or decreases have been found at 4 weeks [59] and 28 weeks [40] postintervention. Since participant details and intervention dosages varied across the studies, the sole effect of postintervention measurement points needs to be addressed in a separate controlled study.

The second and third factors are the age and the baseline health conditions of the participants. A common observation is that older adults are more vulnerable to posttreatment decay [60,61]. Moreover, detraining effects are known to be larger for those who exhibit more impaired swallowing than those with more functional swallowing at baseline [58]. Hence, we hypothesize that the relatively high age and the swallowing difficulties of this study's participants may have contributed to their markedly decreased posttreatment performance compared with the aforementioned studies, which targeted healthy adults without swallowing problems.

The fourth factor is the intervention delivery method. This study employed a tablet-based mHealth app for a swallowing intervention, which fundamentally differs from traditional methods, as telepractice requires participants' self-management with the technology and their active engagement [62] for higher performance. Hence, it is possible that the long-term effects of telepractice interventions are susceptible to individual-dependent aspects (eg, technology learning rate, engagement rate).

The fifth factor is the type of outcome measure used (ie, isometric tongue pressure vs swallowing tongue pressure). As seen from the results, the comparison between isometric and swallowing tongue pressure demonstrated that only the latter was significantly improved in the posttreatment performance.

XSL•FO
RenderX

In addition, both tongue-related measures returned to baseline after the detraining period, providing no direct clues regarding the correlation between a specific type of measure and its long-term effect. Although the maintenance of different types of measures is rarely found in the previous literature, the discrepancy between the posttreatment performance of isometric versus swallowing tongue pressure in this study indirectly suggests that the chosen outcome measure may be one possible factor that needs to be considered when interpreting long-term postintervention effects.

mHealth and Self-Management

The mHealth platform can enhance the delivery of exercise maneuvers through improvements in training accuracy, economic efficiency, safety continuity, and self-management compared with traditional clinician-to-client intervention [63-66]. In this study, we used a biofeedback system involving a mirror function (in effortful prolonged swallowing and effortful tongue rotation) and a real-time pitch graph (in effortful pitch glide), preventing improper or incorrect performance, which may counteract the purpose of the maneuvers and become entrenched [67,68]. In this way, the training accuracy is maintained. Although there is no sufficient evidence that biofeedback benefits overall swallowing functions among patients with dysphagia, in studies that used tongue manometry as biofeedback, significant increases in maximum isometric and swallowing tongue pressure were observed in patients who received biofeedback compared with those who did not [69]. In addition, a previous study has suggested that home-based therapy can be as effective as office-based therapy if a biofeedback system is applied [70]. In addition to increasing training accuracy, the home-based mHealth training app may enable the participants to attend various training programs (eg, the 8-week program adopted in this study) without the burden of visiting clinics as frequently or the need to pay extra costs [71,72]. Thus, the time- and cost-effectiveness of mHealth programs with tongue-strengthening exercises may represent strong points for promoting their widespread use. Furthermore, older adults, including community-dwelling and patient groups, have higher risks for falls and disease complications, which highlights the need to promote their safety by reducing the frequency of their clinic visits [65,73]. The mHealth service could also allow for continuation of therapy sessions even during unexpected circumstances, such as pandemics (eg, COVID-19) [74]. A nondirect clinician-to-client intervention service could serve as a valid alternative to direct interaction.

In addition, it is highly necessary to construct mHealth tools that provide the support needed to promote participants' adherence to the program. In general, the training adherence rate among the older population with various chronic diseases tends to decrease significantly for long-term training [75]. To address this, the tablet-based app used in this study included reminder functions, such as a calendar, a progress bar, and alarms, to promote participants' self-monitoring of their adherence level. For instance, if a participant did not meet the training goal in a session, an incomplete red (or orange) circle would appear on the calendar and the progress bar would not be filled completely in green. The participant could then check their progress through visual biofeedback and finish the training accordingly. Such reminder features are commonly adopted in other training apps [76,77] and may be considered a crucial component of a self-therapy–related app. App features aimed at promoting users' adherence (eg, reminders, feedback), however, may not be designed to help users easily maintain long-term engagement without external support [78-80]. This is especially true for older adults who are not familiar with digital devices and apps, including mHealth devices [34,81]. Therefore, a hybrid method of intervention delivery (eg, mHealth app intervention with human mediation) may be necessary to help older adults maintain their adherence and thus potentially maximize the intended therapeutic effect, even when using the mHealth app as a primary means of swallowing health care. In this study, researchers checked the participants' progress face-to-face once every 2 weeks to monitor their adherence to the intervention program using the mHealth app. A total of 8 of 11 participants successfully completed the program with optimal intended adherence, and it is possible that the overall adherence of the participants may have been lower in the absence of human mediation (ie, without the biweekly face-to-face meeting). A further study is warranted to see if similar therapeutic results would be observed for the mHealth intervention without external support.

Furthermore, in order to compensate for the possible issues caused by face-to-face interactions and to develop the current app platform into a more convenient means of intervention, it is suggested that real-time monitoring and control solutions be added into the app to allow clinicians to remotely keep track of the participants' adherence and performance.

Limitations and Future Directions

The findings from the study indicate that, using the mHealth app, the intervention was able to help older adults increase their swallowing tongue pressures. However, this study is not without limitations. This study involves a small number of participants, which does not permit the generalization of its findings. Accordingly, we could not classify participants by severity and make group comparisons. Furthermore, only 6 participants agreed to take part in the 12-week postintervention evaluation. The results from these participants may not have been sufficient to derive a generalized long-term effect of the given protocol. Thus, an implementation of the mHealth app with a larger sample size is proposed. Finally, although the effect of biweekly face-to-face monitoring on the increased swallowing tongue pressure measured at the postintervention phase is not negligible, the current study did not address the independent effect of the face-to-face monitoring on mHealth intervention effects.

We suggest additional issues that may be useful to address in future studies. First, although the participants of the current study did not receive any retraining during the 8 weeks of rest, it is likely that subsequent retraining may help participants retain their swallowing tongue pressures. Hence, a randomized controlled trial is further warranted to see if a treatment group (which follows a maintenance protocol after the intervention) exhibits higher long-term effects of training compared with a control group (which does not receive any retraining after the intervention). In addition, it may be necessary to take into consideration the frequency of postintervention measurements

XSL•FO
RenderX

and the measurement points (eg, 8 weeks versus 12 weeks postintervention) to see if the detraining period influences the maintenance of swallowing functions.

Furthermore, the training intensity may have negatively affected the participants' overall adherence. Several participants in our study expressed dissatisfaction toward the training intensity and frequency, claiming that 3 sessions a day is excessive, given that they have other plans during the day. Although the 365 Healthy Swallowing Coach app originally provided a tailored training setting for users (ie, users are able to customize the training schedule and training amount), this feature was made unavailable in the current study to measure the therapeutic effect of the fixed amount of training (ie, 3 times/day, 5 days/week). Hence, the reported problems related to the intensity of the protocol call for additional research to systematically investigate the optimal dosage of the mHealth swallowing intervention by comparing the effect of intensive (eg, 3 times/day, 7 days/week for 2 weeks) versus distributed (eg, 1 time/day, 3 days/week for 12 weeks) intervention methods.

Conclusion

Swallowing tongue pressure is known to be closely related to dysphagia symptoms. This is the first study to demonstrate the effectiveness of the combined methods of effortful prolonged swallowing, effortful pitch glide, and effortful tongue rotation using mobile app training accompanied by biweekly human mediation in improving swallowing tongue pressure in older adults. The fixed amount of the training dosage (ie, 3 times/day, 5 days/week) was monitored biweekly to maintain the participants' optimal adherence to the training program. As a result, a significant increase in swallowing tongue pressure was observed after 8 weeks of the intervention. This mHealth pilot study will help guide future intervention studies for older adults who have limited access to health care services. The mHealth app program is a promising method of providing therapy through the execution of intensive and repetitive tongue exercises for vulnerable older adults. In order to observe therapy efficacy with a larger sample size and to investigate long-term effects of the intervention program, further studies are warranted. Moreover, further studies using videofluoroscopic swallowing studies are needed to investigate the effects that the app-based therapy program may have on the various stages of swallowing.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT for convergent research in development program for convergence R&D over Science and Technology Liberal Arts (NRF-2017M3C1B6070665, NRF-2017M3C1B6070526). We would like to express our sincerest gratitude to all of the participants for the time and effort they invested by participating in this research. We would also like to acknowledge the contribution of two senior community centers for helping in the recruitment of participants.

Authors' Contributions

All authors have made a substantial contribution to this study. HK, HY, and SIN obtained research funding for the study. HK contributed to the app contents, study concept and design, data analysis and interpretation, and drafting and revision of the manuscript. N-BC and SS contributed to data collection. KMK and M Kang contributed to drafting and editing the manuscript. SIN contributed to participant recruitment and revision of the manuscript. HY, YC, M Kim, and JK contributed to app design and development and revision of the manuscript. All authors testify that they have reviewed and approved the final version of the manuscript being submitted. The paper is the authors' original work and is not under consideration for publication elsewhere.

Conflicts of Interest

None declared.

References

1. Fei T, Polacco RC, Hori SE, Molfenter SM, Peladeau-Pigeon M, Tsang C, et al. Age-related differences in tongue-palate pressures for strength and swallowing tasks. Dysphagia 2013 Dec;28(4):575-581 [FREE Full text] [doi: 10.1007/s00455-013-9469-6] [Medline: 23677389]
2. Utanohara Y, Hayashi R, Yoshikawa M, Yoshida M, Tsuga K, Akagawa Y. Standard values of maximum tongue pressure taken using newly developed disposable tongue pressure measurement device. Dysphagia 2008 Sep;23(3):286-290. [doi: 10.1007/s00455-007-9142-z] [Medline: 18574632]
3. Feng X, Todd T, Lintzenich CR, Ding J, Carr JJ, Ge Y, et al. Aging-related geniohyoid muscle atrophy is related to aspiration status in healthy older adults. J Gerontol A Biol Sci Med Sci 2013 Jul;68(7):853-860 [FREE Full text] [doi: 10.1093/gerona/gls225] [Medline: 23112114]
4. Machida N, Tohara H, Hara K, Kumakura A, Wakasugi Y, Nakane A, et al. Effects of aging and sarcopenia on tongue pressure and jaw-opening force. Geriatr Gerontol Int 2017 Feb;17(2):295-301. [doi: 10.1111/ggi.12715] [Medline: 26800427]
5. Maeda K, Akagi J. Decreased tongue pressure is associated with sarcopenia and sarcopenic dysphagia in the elderly. Dysphagia 2015 Feb;30(1):80-87. [doi: 10.1007/s00455-014-9577-y] [Medline: 25248988]

XSL•FO
RenderX

6. Morris ME, Aguilera A. Mobile, Social, and Wearable Computing and the Evolution of Psychological Practice. Prof Psychol Res Pr 2012 Dec;43(6):622-626. [doi: 10.1037/a0029041] [Medline: 25587207]
7. Choi SH, Kim H, Choi C, Seo HN, Park CR. Characteristics of Tongue Pressures Based on Swallowing Tasks in Korean Healthy Older Adults. Audiol Speech Res 2018 Jul;14(3):194-203. [doi: 10.21848/asr.2018.14.3.194]
8. Steele CM. The influence of tongue strength on oral viscosity discrimination acuity. J Texture Stud 2018 Jun;49(3):249-255 [FREE Full text] [doi: 10.1111/jtxs.12318] [Medline: 29280136]
9. Nakamori M, Hosomi N, Ishikawa K, Imamura E, Shishido T, Ohshita T, et al. Prediction of Pneumonia in Acute Stroke Patients Using Tongue Pressure Measurements. PLoS One 2016;11(11):e0165837 [FREE Full text] [doi: 10.1371/journal.pone.0165837] [Medline: 27802333]
10. Adkins C, Takakura W, Spiegel BMR, Lu M, Vera-Llonch M, Williams J, et al. Prevalence and Characteristics of Dysphagia Based on a Population-Based Survey. Clin Gastroenterol Hepatol 2020 Aug;18(9):1970-1979.e2 [FREE Full text] [doi: 10.1016/j.cgh.2019.10.029] [Medline: 31669055]
11. Igarashi K, Kikutani T, Tamura F. Survey of suspected dysphagia prevalence in home-dwelling older people using the 10-Item Eating Assessment Tool (EAT-10). PLoS One 2019;14(1):e0211040 [FREE Full text] [doi: 10.1371/journal.pone.0211040] [Medline: 30673750]
12. Kawashima K, Motohashi Y, Fujishima I. Prevalence of dysphagia among community-dwelling elderly individuals as estimated using a questionnaire for dysphagia screening. Dysphagia 2004;19(4):266-271. [doi: 10.1007/s00455-004-0013-6] [Medline: 15667063]
13. Kim H, Lee H, Pyo HY, Kim JW, Choi SH, Choi H, et al. Commun Sci Disord 2019 Jun 30;24(2):507-517. [doi: 10.12963/csd.19599]
14. Leder SB, Suiter DM, Murray J, Rademaker AW. Can an oral mechanism examination contribute to the assessment of odds of aspiration? Dysphagia 2013 Sep;28(3):370-374. [doi: 10.1007/s00455-012-9442-9] [Medline: 23292501]
15. Sakai K, Hirano H, Watanabe Y, Tohara H, Sato E, Sato K, et al. An examination of factors related to aspiration and silent aspiration in older adults requiring long-term care in rural Japan. J Oral Rehabil 2016 Feb;43(2):103-110. [doi: 10.1111/joor.12349] [Medline: 26432521]
16. Yoshida M, Kikutani T, Tsuga K, Utanohara Y, Hayashi R, Akagawa Y. Decreased tongue pressure reflects symptom of dysphagia. Dysphagia 2006 Jan;21(1):61-65. [doi: 10.1007/s00455-005-9011-6] [Medline: 16544085]
17. Robbins J, Humpal NS, Banaszynski K, Hind J, Rogus-Pulia N. Age-Related Differences in Pressures Generated During Isometric Presses and Swallows by Healthy Adults. Dysphagia 2016 Feb;31(1):90-96. [doi: 10.1007/s00455-015-9662-x] [Medline: 26525059]
18. Furuya J, Nakamura S, Ono T, Suzuki T. Tongue pressure production while swallowing water and pudding and during dry swallow using a sensor sheet system. J Oral Rehabil 2012 Sep;39(9):684-691. [doi: 10.1111/j.1365-2842.2012.02319.x] [Medline: 22642615]
19. McCullough GH, Kamarunas E, Mann GC, Schmidley JW, Robbins JA, Crary MA. Effects of Mendelsohn maneuver on measures of swallowing duration post stroke. Top Stroke Rehabil 2012;19(3):234-243 [FREE Full text] [doi: 10.1310/tsr1903-234] [Medline: 22668678]
20. Fukuoka T, Ono T, Hori K, Tamine K, Nozaki S, Shimada K, et al. Effect of the effortful swallow and the Mendelsohn maneuver on tongue pressure production against the hard palate. Dysphagia 2013 Dec;28(4):539-547. [doi: 10.1007/s00455-013-9464-y] [Medline: 23576155]
21. Doeltgen SH, Ong E, Scholten I, Cock C, Omari T. Biomechanical Quantification of Mendelsohn Maneuver and Effortful Swallowing on Pharyngoesophageal Function. Otolaryngol Head Neck Surg 2017 Nov;157(5):816-823. [doi: 10.1177/0194599817708173] [Medline: 28608778]
22. Balou M, Herzberg EG, Kamelhar D, Molfenter SM. An intensive swallowing exercise protocol for improving swallowing physiology in older adults with radiographically confirmed dysphagia. Clin Interv Aging 2019;14:283-288 [FREE Full text] [doi: 10.2147/CIA.S194723] [Medline: 30804667]
23. Malandraki GA, Rajappa A, Kantarcigil C, Wagner E, Ivey C, Youse K. The Intensive Dysphagia Rehabilitation Approach Applied to Patients With Neurogenic Dysphagia: A Case Series Design Study. Arch Phys Med Rehabil 2016 Apr;97(4):567-574. [doi: 10.1016/j.apmr.2015.11.019] [Medline: 26711168]
24. Kim J, Kim Y, Lee H, Kim K, Kim S, Kim T, et al. Effect of the combination of Mendelsohn maneuver and effortful swallowing on aspiration in patients with dysphagia after stroke. J Phys Ther Sci 2017 Nov;29(11):1967-1969 [FREE Full text] [doi: 10.1589/jpts.29.1967] [Medline: 29200636]
25. Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res 2008 Feb;51(1):S225-S239. [doi: 10.1044/1092-4388(2008/018)] [Medline: 18230848]
26. Robbins J, Butler SG, Daniels SK, Diez Gross R, Langmore S, Lazarus CL, et al. Swallowing and dysphagia rehabilitation: translating principles of neural plasticity into clinically oriented evidence. J Speech Lang Hear Res 2008 Feb;51(1):S276-S300. [doi: 10.1044/1092-4388(2008/021)] [Medline: 18230851]
27. Pearson WG, Hindson DF, Langmore SE, Zumwalt AC. Evaluating swallowing muscles essential for hyolaryngeal elevation by using muscle functional magnetic resonance imaging. Int J Radiat Oncol Biol Phys 2013 Mar 01;85(3):735-740 [FREE Full text] [doi: 10.1016/j.ijrobp.2012.07.2370] [Medline: 22995662]

XSL•FO
RenderX

28. Murakami K, Hori K, Minagi Y, Uehara F, Salazar SE, Ishihara S, et al. Coordination of tongue pressure production, hyoid movement, and suprahyoid muscle activity during squeezing of gels. Arch Oral Biol 2020 Mar;111:104631. [doi: 10.1016/j.archoralbio.2019.104631] [Medline: 31869725]
29. Arakawa I, Koide K, Takahashi M, Mizuhashi F. Effect of the tongue rotation exercise training on the oral functions in normal adults - Part 1 investigation of tongue pressure and labial closure strength. J Oral Rehabil 2015 Jun;42(6):407-413. [doi: 10.1111/joor.12271] [Medline: 25640889]
30. Clark HM, O'Brien K, Calleja A, Corrie SN. Effects of directional exercise on lingual strength. J Speech Lang Hear Res 2009 Aug;52(4):1034-1047. [doi: 10.1044/1092-4388(2009/08-0062)] [Medline: 19641080]
31. Fitzpatrick AL, Powe NR, Cooper LS, Ives DG, Robbins JA. Barriers to health care access among the elderly and who perceives them. Am J Public Health 2004 Oct;94(10):1788-1794. [doi: 10.2105/ajph.94.10.1788] [Medline: 15451751]
32. Murata C, Yamada T, Chen C, Ojima T, Hirai H, Kondo K. Barriers to health care among the elderly in Japan. Int J Environ Res Public Health 2010 Apr;7(4):1330-1341 [FREE Full text] [doi: 10.3390/ijerph7041330] [Medline: 20617033]
33. Bene BA, O'Connor S, Mastellos N, Majeed A, Fadahunsi KP, O'Donoghue J. Impact of mobile health applications on self-management in patients with type 2 diabetes mellitus: protocol of a systematic review. BMJ Open 2019 Jun 25;9(6):e025714. [doi: 10.1136/bmjopen-2018-025714] [Medline: 31243029]
34. Kim H, Lee S, Cho N, You H, Choi T, Kim J. User-Dependent Usability and Feasibility of a Swallowing Training mHealth App for Older Adults: Mixed Methods Pilot Study. JMIR Mhealth Uhealth 2020 Jul 27;8(7):e19585 [FREE Full text] [doi: 10.2196/19585] [Medline: 32663161]
35. Jang HJ, Leigh JH, Seo HG, Han TR, Oh BM. Effortful swallow enhances vertical hyolaryngeal movement and prolongs duration after maximal excursion. J Oral Rehabil 2015 Oct;42(10):765-773. [doi: 10.1111/joor.12312] [Medline: 26013277]
36. Mendelsohn MS, McConnel FM. Function in the pharyngoesophageal segment. Laryngoscope 1987 Apr;97(4):483-489. [doi: 10.1288/00005537-198704000-00014] [Medline: 3561135]
37. Miloro KV, Pearson WG, Langmore SE. Effortful pitch glide: a potential new exercise evaluated by dynamic MRI. J Speech Lang Hear Res 2014 Aug;57(4):1243-1250 [FREE Full text] [doi: 10.1044/2014_JSLHR-S-13-0168] [Medline: 24686494]
38. Völter C, Schirmer C, Hinsen D, Roeber M, Dazert S, Bilda K. Therapist-Guided Telerehabilitation for Adult Cochlear Implant Users: Developmental and Feasibility Study. JMIR Rehabil Assist Technol 2020 May 28;7(1):e15843 [FREE Full text] [doi: 10.2196/15843] [Medline: 32255434]
39. IOPI User Manual. IOPI Medical LLC. URL: http://ww.iopimedical.com/Images/PDFs/800-2301-02%20LS1,%20EN,%20IOPI%202.3%20User%20Manual,%20WEB.pdf [accessed 2020-06-15]
40. Oh J. Effects of Tongue Strength Training and Detraining on Tongue Pressures in Healthy Adults. Dysphagia 2015 Jun;30(3):315-320. [doi: 10.1007/s00455-015-9601-x] [Medline: 25840786]
41. Park H, Oh D, Yoon T, Park J. Effect of effortful swallowing training on tongue strength and oropharyngeal swallowing function in stroke patients with dysphagia: a double-blind, randomized controlled trial. Int J Lang Commun Disord 2019 May;54(3):479-484. [doi: 10.1111/1460-6984.12453] [Medline: 30693627]
42. Robbins J, Gangnon RE, Theis SM, Kays SA, Hewitt AL, Hind JA. The effects of lingual exercise on swallowing in older adults. J Am Geriatr Soc 2005 Sep;53(9):1483-1489. [doi: 10.1111/j.1532-5415.2005.53467.x] [Medline: 16137276]
43. Yano J, Yamamoto-Shimizu S, Yokoyama T, Kumakura I, Hanayama K, Tsubahara A. Effects of anterior tongue strengthening exercises on posterior tongue strength in healthy young adults. Arch Oral Biol 2019 Feb;98:238-242. [doi: 10.1016/j.archoralbio.2018.11.028] [Medline: 30522043]
44. Yeates EM, Steele CM, Pelletier CA. Tongue pressure and submental surface electromyography measures during noneffortful and effortful saliva swallows in healthy women. Am J Speech Lang Pathol 2010 Aug;19(3):274-281. [doi: 10.1044/1058-0360(2010/09-0040)] [Medline: 20543016]
45. Clark HM, Shelton N. Training effects of the effortful swallow under three exercise conditions. Dysphagia 2014 Oct;29(5):553-563. [doi: 10.1007/s00455-014-9544-7] [Medline: 24913837]
46. Gassert RB, Pearson WG. Evaluating muscles underlying tongue base retraction in deglutition using muscular functional magnetic resonance imaging (mfMRI). Magn Reson Imaging 2016 Feb;34(2):204-208 [FREE Full text] [doi: 10.1016/j.mri.2015.10.029] [Medline: 26523657]
47. Ono T, Kumakura I, Arimoto M, Hori K, Dong J, Iwata H, et al. Influence of bite force and tongue pressure on oro-pharyngeal residue in the elderly. Gerodontology 2007 Sep;24(3):143-150. [doi: 10.1111/j.1741-2358.2007.00172.x] [Medline: 17696891]
48. Jenks J, Pitts LL. Effects of an Intensive Exercise-Based Swallowing Program for Persons With Parkinson's Disease and Complex Medical History: A Single-Case Experiment. Am J Speech Lang Pathol 2019 Aug 09;28(3):1268-1274. [doi: 10.1044/2019_AJSLP-18-0168] [Medline: 31335179]
49. Baker E. Optimal intervention intensity. Int J Speech Lang Pathol 2012 Oct;14(5):401-409. [doi: 10.3109/17549507.2012.700323] [Medline: 22916999]
50. Krekeler BN, Rowe LM, Connor NP. Dose in Exercise-Based Dysphagia Therapies: A Scoping Review [in press]. Dysphagia 2020 Mar 05. [doi: 10.1007/s00455-020-10104-3] [Medline: 32140905]
51. Ament W, Verkerke GJ. Exercise and fatigue. Sports Med 2009;39(5):389-422. [doi: 10.2165/00007256-200939050-00005] [Medline: 19402743]

XSL•FO
RenderX

52. Bakheit AMO, Shaw S, Barrett L, Wood J, Carrington S, Griffiths S, et al. A prospective, randomized, parallel group, controlled study of the effect of intensity of speech and language therapy on early recovery from poststroke aphasia. Clin Rehabil 2007 Oct;21(10):885-894. [doi: 10.1177/0269215507078486] [Medline: 17981847]
53. Baker E. Optimal intervention intensity in speech-language pathology: discoveries, challenges, and unchartered territories. Int J Speech Lang Pathol 2012 Oct;14(5):478-485. [doi: 10.3109/17549507.2012.717967] [Medline: 22974107]
54. Jin J, Sklar GE, Min Sen Oh V, Chuen Li S. Factors affecting therapeutic compliance: A review from the patient's perspective. Ther Clin Risk Manag 2008 Feb;4(1):269-286 [FREE Full text] [doi: 10.2147/tcrm.s1458] [Medline: 18728716]
55. Dignam J, Copland D, McKinnon E, Burfein P, O'Brien K, Farrell A, et al. Intensive Versus Distributed Aphasia Therapy: A Nonrandomized, Parallel-Group, Dosage-Controlled Study. Stroke 2015 Aug;46(8):2206-2211. [doi: 10.1161/STROKEAHA.115.009522] [Medline: 26106114]
56. O'Brien SR, Zhang N. Association Between Therapy Intensity and Discharge Outcomes in Aged Medicare Skilled Nursing Facilities Admissions. Arch Phys Med Rehabil 2018 Jan;99(1):107-115 [FREE Full text] [doi: 10.1016/j.apmr.2017.07.012] [Medline: 28860096]
57. Wambaugh JL, Wright S, Boss E, Mauszycki SC, DeLong C, Hula W, et al. Effects of Treatment Intensity on Outcomes in Acquired Apraxia of Speech. Am J Speech Lang Pathol 2018 Mar;27(1S):306-322. [doi: 10.1044/2017_ajslp-16-0188]
58. Troche MS, Rosenbek JC, Okun MS, Sapienza CM. Detraining outcomes with expiratory muscle strength training in Parkinson disease. J Rehabil Res Dev 2014;51(2):305-310 [FREE Full text] [doi: 10.1682/JRRD.2013.05.0101] [Medline: 24933728]
59. Van den Steen L, Schellen C, Verstraelen K, Beeckman A, Vanderwegen J, De Bodt M, et al. Tongue-Strengthening Exercises in Healthy Older Adults: Specificity of Bulb Position and Detraining Effects. Dysphagia 2018 Jun;33(3):337-344. [doi: 10.1007/s00455-017-9858-3] [Medline: 29052051]
60. Ivey FM, Tracy BL, Lemmer JT, NessAiver M, Metter EJ, Fozard JL, et al. Effects of strength training and detraining on muscle quality: age and gender comparisons. J Gerontol A Biol Sci Med Sci 2000 Mar;55(3):B152-7; discussion B158. [doi: 10.1093/gerona/55.3.b152] [Medline: 10795719]
61. Toraman NF. Short term and long term detraining: is there any difference between young-old and old people? Br J Sports Med 2005 Aug;39(8):561-564. [doi: 10.1136/bjsm.2004.015420] [Medline: 16046344]
62. Cartmill B, Wall LR, Ward EC, Hill AJ, Porceddu SV. Computer Literacy and Health Locus of Control as Determinants for Readiness and Acceptability of Telepractice in a Head and Neck Cancer Population. Int J Telerehabil 2016;8(2):49-60 [FREE Full text] [doi: 10.5195/ijt.2016.6203] [Medline: 28775801]
63. Faust-Christmann CA, Taetz B, Zolynski G, Zimmermann T, Bleser G. A Biofeedback App to Instruct Abdominal Breathing (Breathing-Mentor): Pilot Experiment. JMIR Mhealth Uhealth 2019 Sep 30;7(9):e13703 [FREE Full text] [doi: 10.2196/13703] [Medline: 31573919]
64. Griffin M, Bentley J, Shanks J, Wood C. The effectiveness of Lee Silverman Voice Treatment therapy issued interactively through an iPad device: A non-inferiority study. J Telemed Telecare 2018 Apr;24(3):209-215. [doi: 10.1177/1357633X17691865] [Medline: 28147896]
65. Karlsen C, Ludvigsen MS, Moe CE, Haraldstad K, Thygesen E. Experiences of community-dwelling older adults with the use of telecare in home care services: a qualitative systematic review. JBI Database System Rev Implement Rep 2017 Dec;15(12):2913-2980. [doi: 10.11124/JBISRIR-2017-003345] [Medline: 29219874]
66. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: Implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Mar 20:1357633X20916567. [doi: 10.1177/1357633X20916567] [Medline: 32196391]
67. Giggins OM, Persson UM, Caulfield B. Biofeedback in rehabilitation. J Neuroeng Rehabil 2013 Jun 18;10:60 [FREE Full text] [doi: 10.1186/1743-0003-10-60] [Medline: 23777436]
68. Vose AK, Marcus A, Humbert I. Kinematic Visual Biofeedback Improves Accuracy of Swallowing Maneuver Training and Accuracy of Clinician Cues During Training in Stroke Patients with Dysphagia. PM R 2019 Nov;11(11):1159-1169. [doi: 10.1002/pmrj.12093] [Medline: 30701691]
69. Benfield JK, Everton LF, Bath PM, England TJ. Does Therapy With Biofeedback Improve Swallowing in Adults With Dysphagia? A Systematic Review and Meta-Analysis. Arch Phys Med Rehabil 2019 Mar;100(3):551-561. [doi: 10.1016/j.apmr.2018.04.031] [Medline: 29859178]
70. Rao SSC, Valestin JA, Xiang X, Hamdy S, Bradley CS, Zimmerman MB. Home-based versus office-based biofeedback therapy for constipation with dyssynergic defecation: a randomised controlled trial. Lancet Gastroenterol Hepatol 2018 Nov;3(11):768-777 [FREE Full text] [doi: 10.1016/S2468-1253(18)30266-8] [Medline: 30236904]
71. Abelson JS, Kaufman E, Symer M, Peters A, Charlson M, Yeo H. Barriers and benefits to using mobile health technology after operation: A qualitative study. Surgery 2017 Sep;162(3):605-611. [doi: 10.1016/j.surg.2017.05.007] [Medline: 28651777]
72. Bowser D, Shepard D, Nandakumar A, Okunogbe A, Morrill T, Halasa-Rappell Y, et al. Cost Effectiveness of Mobile Health for Antenatal Care and Facility Births in Nigeria. Ann Glob Health 2018 Nov 05;84(4):592-602 [FREE Full text] [doi: 10.9204/aogh.2364] [Medline: 30779506]

XSL•FO
RenderX

73. Chaudhuri S, Kneale L, Le T, Phelan E, Rosenberg D, Thompson H, et al. Older Adults' Perceptions of Fall Detection Devices. J Appl Gerontol 2017 Aug;36(8):915-930 [FREE Full text] [doi: 10.1177/0733464815591211] [Medline: 26112030]
74. Zhou X, Snoswell CL, Harding LE, Bambling M, Edirippulige S, Bai X, et al. The Role of Telehealth in Reducing the Mental Health Burden from COVID-19. Telemed J E Health 2020 Apr;26(4):377-379. [doi: 10.1089/tmj.2020.0068] [Medline: 32202977]
75. Nicolson PJA, Bennell KL, Dobson FL, Van Ginckel A, Holden MA, Hinman RS. Interventions to increase adherence to therapeutic exercise in older adults with low back pain and/or hip/knee osteoarthritis: a systematic review and meta-analysis. Br J Sports Med 2017 May;51(10):791-799. [doi: 10.1136/bjsports-2016-096458] [Medline: 28087567]
76. Fishbein JN, Nisotel LE, MacDonald JJ, Amoyal PN, Jacobs JM, Flanagan C, et al. Mobile Application to Promote Adherence to Oral Chemotherapy and Symptom Management: A Protocol for Design and Development. JMIR Res Protoc 2017 Apr 20;6(4):e62 [FREE Full text] [doi: 10.2196/resprot.6198] [Medline: 28428158]
77. Palazzo C, Klinger E, Dorner V, Kadri A, Thierry O, Boumenir Y, et al. Barriers to home-based exercise program adherence with chronic low back pain: Patient expectations regarding new technologies. Ann Phys Rehabil Med 2016 Apr;59(2):107-113 [FREE Full text] [doi: 10.1016/j.rehab.2016.01.009] [Medline: 27050664]
78. Damschroder LJ, Buis LR, McCant FA, Kim HM, Evans R, Oddone EZ, et al. Effect of Adding Telephone-Based Brief Coaching to an mHealth App (Stay Strong) for Promoting Physical Activity Among Veterans: Randomized Controlled Trial. J Med Internet Res 2020 Aug 04;22(8):e19216 [FREE Full text] [doi: 10.2196/19216] [Medline: 32687474]
79. Juarascio AS, Goldstein SP, Manasse SM, Forman EM, Butryn ML. Perceptions of the feasibility and acceptability of a smartphone application for the treatment of binge eating disorders: Qualitative feedback from a user population and clinicians. Int J Med Inform 2015 Oct;84(10):808-816 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.06.004] [Medline: 26113461]
80. Ly KH, Janni E, Wrede R, Sedem M, Donker T, Carlbring P, et al. Experiences of a guided smartphone-based behavioral activation therapy for depression: A qualitative study. Internet Interventions 2015 Mar;2(1):60-68. [doi: 10.1016/j.invent.2014.12.002]
81. Lu S, Wen T, Chang P. The Study of Smartphone Usage Competency Assessment and Training for the Elderly. Stud Health Technol Inform 2017;245:161-165. [Medline: 29295073]

Abbreviations

ASSET: A Successful Swallow with Effortful Trainings
IOPI: Iowa Oral Performance Instrument
mHealth: mobile health

Edited by G Eysenbach; submitted 06.07.20; peer-reviewed by MDG Pimentel; comments to author 04.08.20; revised version received 07.08.20; accepted 03.10.20; published 16.10.20.

Please cite as:
Kim H, Cho NB, Kim J, Kim KM, Kang M, Choi Y, Kim M, You H, Nam SI, Shin S
Implementation of a Home-Based mHealth App Intervention Program With Human Mediation for Swallowing Tongue Pressure Strengthening Exercises in Older Adults: Longitudinal Observational Study
JMIR Mhealth Uhealth 2020;8(10):e22080
URL: http://mhealth.jmir.org/2020/10/e22080/
doi:10.2196/22080
PMID:33012704

©HyangHee Kim, Nam-Bin Cho, Jinwon Kim, Kyung Min Kim, Minji Kang, Younggeun Choi, Minjae Kim, Heecheon You, Seok In Nam, Soyeon Shin. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 16.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mobile Apps for Speech-Language Therapy in Adults With Communication Disorders: Review of Content and Quality

Atiyeh Vaezipour[1], PhD; Jessica Campbell[2], PhD; Deborah Theodoros[1], PhD; Trevor Russell[1], PhD

[1]RECOVER Injury Research Centre, Faculty of Health and Behavioural Sciences, The University of Queensland, Brisbane, Australia

[2]Queensland Aphasia Research Centre, Faculty of Health and Behavioural Sciences, The University of Queensland, Brisbane, Australia

Corresponding Author:
Atiyeh Vaezipour, PhD
RECOVER Injury Research Centre
Faculty of Health and Behavioural Sciences
The University of Queensland
288 Herston Road
Brisbane, 4006
Australia
Phone: 61 7 3365 5560
Email: a.vaezipour@uq.edu.au

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2020/12/e26309/

Abstract

Background: Worldwide, more than 75% of people with acquired brain injury (ABI) experience communication disorders. Communication disorders are impairments in the ability to communicate effectively, that is, sending, receiving, processing, and comprehending verbal and nonverbal concepts and symbols. Such disorders may have enduring impacts on employment, social participation, and quality of life. Technology-enabled interventions such as mobile apps have the potential to increase the reach of speech-language therapy to treat communication disorders. However, ensuring that apps are evidence-based and of high quality is critical for facilitating safe and effective treatment for adults with communication disorders.

Objective: The aim of this review is to identify mobile apps that are currently widely available to adults with communication disorders for speech-language therapy and to assess their content and quality using the validated Mobile App Rating Scale (MARS).

Methods: Google Play Store, Apple App Store, and webpages were searched to identify mobile apps for speech-language therapy. Apps were included in the review if they were designed for the treatment of adult communication disorders after ABI, were in English, and were either free or for purchase. Certified speech-language pathologists used the MARS to assess the quality of the apps.

Results: From a total of 2680 apps identified from Google Play Store, Apple App Store, and web searches, 2.61% (70/2680) apps met the eligibility criteria for inclusion. Overall, 61% (43/70) were available for download on the iPhone Operating System (iOS) platform, 20% (14/70) on the Android platform, and 19% (13/70) on both iOS and Android platforms. A content analysis of the apps revealed 43 apps for *language*, 17 apps for *speech*, 8 apps for *cognitive communication*, 6 apps for *voice*, and 5 apps for *oromotor function* or *numeracy*. The overall MARS mean score was 3.7 out of 5, SD 0.6, ranging between 2.1 and 4.5, with *functionality* being the highest-scored subscale (4.3, SD 0.6), *followed by aesthetics* (3.8, SD 0.8), *information* (3.4, SD 0.6), *and engagement* (3.3, SD 0.6). The top 5 apps were *Naming Therapy* (4.6/5), *Speech Flipbook Standard* (4.6/5), *Number Therapy* (4.5/5), *Answering Therapy*, and *Constant Therapy* (4.4/5).

Conclusions: To our knowledge, this is the first study to systematically identify and evaluate a broad range of mobile apps for speech-language therapy for adults with communication disorders after sustaining ABI. We found a lack of interactive and engaging elements in the apps, a critical factor in sustaining self-managed speech-language therapy. More evidence-based apps with a focus on human factors, user experience, and a patient-led design approach are required to enhance effectiveness and long-term use.

XSL•FO
RenderX

(JMIR Mhealth Uhealth 2020;8(10):e18858) doi:10.2196/18858

KEYWORDS

communication disorders; speech therapy; language therapy; ergonomics; rehabilitation; mobile health; mHealth

Introduction

Acquired brain injury (ABI) is a life-changing health condition that can result from trauma, cerebrovascular events, and brain tumors [1]. The population affected by ABI is large and growing, with 69 million individuals sustaining a traumatic brain injury (TBI) globally each year [2], and the global incidence of first stroke is expected to rise from 16 million in 2005 to 23 million in 2030 [3]. More than 75% of people experience a communication disorder after ABI [4]. A communication disorder may involve speech impairment characterized by slurred and indistinct speech (dysarthria or apraxia of speech), specific language impairments characterized by difficulties with comprehension or expression of language (aphasia), communication difficulties associated with cognitive disorders, and impaired social communication skills [5-7]. Post-ABI communication disorders can impact a person's social integration and participation in their school, work, and community [8]. The quality of life and mood can also be reduced for the affected person and their family members [9]. In addition, communication difficulties may represent significant stressors that influence family and/or caregiver burden [10].

Evidence supports the delivery of rehabilitation by speech-language pathologists (SLPs) for adults after sustaining ABI to treat language [11], motor speech [12], and social communication skill [13] impairments. As the availability of mobile technology increases, apps designed for mobile phones and tablets are increasingly of interest for such therapy. Apps have been developed to identify the presence of aphasia (language impairment) and improve language outcomes [14,15], facilitate homework completion in adults with stroke and TBI [16,17], and improve cognitive skills in adults with acquired cognitive disorders [18]. In a recent cohort study by Munoz et al [16], adults with stroke and TBI used an app targeting speech, language, and cognitive skills. Usage of this app was reported to be higher in geographical areas with limited access to SLP clinics, regardless of demographics such as age. Such apps have the potential to increase the reach of allied health interventions by making therapy available anywhere that a mobile device can be used. The recent global COVID-19 pandemic has highlighted the potential for digital health technologies to provide health care support at a distance [19]. Specifically, mobile therapy apps may increase customization, ease of access, engagement with therapy, and optimize therapy dosage, which could assist in reducing the effects of social stigmas associated with communication impairment [20]. Mobile therapy apps may also offer greater opportunities for generalizing therapy goals to real-world settings and provide additional ways for clients to receive valuable feedback to reinforce positive behaviors and enhance performance [21,22].

Despite the availability and potential benefits of mobile apps for adults with communication disorders after ABI, there is little published evidence regarding their quality beyond the app star rating allocated by some consumers in app store platforms and web-based reviews. A recent analysis of apps for children with speech-language disorders found that most apps were of average quality and that app cost did not always correlate with therapeutic quality [23]. A systematic review of apps targeting general rehabilitation found that some may have a positive impact on outcomes in exercise or gait training or self-management or may be effective as measurement tools [24]. However, only 3 apps targeting communication were included in this general review. The review categorized app functionality without evaluating app qualities such as usability, consumer interaction, or engagement.

In recent years, the term gamification has become increasingly popular in digital health technologies as an underlying element to enhance individuals' engagement with mobile health technologies. Gamification has been described by Detering et al [25] as the "use of game design elements in nongame contexts," that is, the use of game design, game playing techniques, and game mechanisms to engage users and motivate positive behavior [25,26]. Previous systematic reviews have reported the positive effects of gamification on health-related interventions [27,28]. Therefore, the design and development of mobile health technologies should focus on relevant, evidence-based therapy goals, and app functionality as well as target positive user experience and sustained engagement (eg, by using gamification principles). To our knowledge, no systematic quality evaluation of apps for adults with communication disorders has been conducted to date. Therefore, an in-depth evaluation of the quality of mobile apps for adults with communication disorders after ABI is needed.

This study aims to (1) identify the available apps designed for adults with communication disorders and (2) evaluate the quality of the available apps using the Mobile App Rating Scale (MARS) [29], a validated tool that has been used to evaluate various medical apps [23,30]. The outcomes of this study will have implications for the design and development of mobile apps as a clinical rehabilitation tool for speech-language therapy.

Methods

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines were adopted for this study. These have been used successfully for evaluating speech-language therapy apps in children with speech disorders [23].

App Eligibility Criteria

Apps were included in this review if they were primarily designed for adults (18 years or older) with communication disorders secondary to ABI, were for the provision of speech-language therapy (eg, naming drills to improve word-retrieval skills), were in English, were either free or for purchase, were compatible with Android or iPhone Operating

XSL•FO
RenderX

System (iOS), and were available on mobile phones and/or tablets.

Apps were excluded if they were primarily designed for children, were not designed for speech-language therapy (eg, designed to teach English as a second language), provided assessment only without speech-language therapy, and were speech-to-text/text-to-speech apps or other augmentative and alternative communication apps.

App Identification and Search Strategy

To identify mobile apps, a search on Google Play Store (74.6% of the phone and 41.4% tablet market) and Apple App Store (24.8% of the phone and 58.5% of tablet market) was conducted using the following keywords: "*aphasia*," "*apraxia*," "*dysarthria*," "*dyspraxia*," "*dysphasia*," "*speech, articulation*," "*speech therapy*," "*speech pathology*," "*language therapy*," "*speech-language pathology*," "*speech rehabilitation*," and "*language rehabilitation*". The search terms were generated in consultation with 4 certified and experienced SLPs. One researcher entered the terms into the search fields of both the Google Play Store and Apple App Store. Boolean operators were not used as they were not supported by these platforms. In addition, webpages reviewing multiple apps for adults with communication disorders after ABI were reviewed for apps that had not been identified previously. The titles, app platform (ie, iOS or Android), and marketing description of all resulting apps were extracted into a spreadsheet and duplicates were removed. The last search was conducted in November 2019.

App Screening and Extraction

App screening was performed by an Australian certified SLP using the marketing description of the apps against a list of eligibility criteria. If the SLP could not decide the eligibility, a second reviewer was consulted. The included apps were downloaded to either an Android (Samsung Galaxy) or iOS (iPhone/iPad) device, depending on compatibility, for further evaluation and quality appraisal.

To identify the functionality of the apps for adults with communication disorders following ABI, a coding sheet was developed to categorize the apps based on the description of the primary therapeutic function contained in the app description on the Apple App Store or Google Play Store (ie, communication skills that the app is designed to improve). The coding process was a two-step analysis in which 2 independent researchers reviewed the coding sheet and extracted relevant categories.

App Quality Appraisal

A total of 3 certified SLPs scored the quality of the included apps using the MARS [29]. The MARS allows raters to assign a 5-point Likert scale rating (1=inadequate to 5=excellent) across 6 categories. The categories were (1) *engagement*, including individual items for entertainment, interest, customization, interactivity, and whether the app was engaging for the target users; (2) *functionality*, including performance, ease of use, navigation, and gestural design; (3) *aesthetics*, including layout, graphics, and visual appeal; and (4) *information quality*, including accuracy of app description; whether the app had specific, measurable, and achievable goals; quality of information; quantity of information; visual information; credibility; and whether the app was evidence-based. Furthermore, 2 additional categories, *subjective quality* and *perceived impact*, captured the rater's overall impressions. *Subjective quality* included whether the rater would recommend the app (1=definitely not to 5=definitely yes); time predicted by the rater during which the app would be used in the next 12 months, if it was relevant (1=none to 5=>50 times); rater willingness to pay for the app (1=definitely not to 5=definitely yes); and the overall star rating (1=one of the worst apps I have used to 5=one of the best apps I have used). *Perceived impact* was also rated using a 5-point Likert scale (1=strongly disagree to 5=strongly agree) and addressed the app's perceived potential to increase users' awareness, knowledge, attitudes, and motivation to complete the desired health behavior, encourage help-seeking, and improve user communication.

As recommended by the MARS developers, each rater viewed the training video by Stoyanov and Hides [29,31]. The raters practiced rating before using the MARS tool and discussed their ratings to agree on the relevance of the MARS items to the SLP apps and to establish a consensus on the MARS items. Each app was trialed for at least 10 min before the rating was completed. When reviewing the apps with the MARS tool, the raters also reviewed app descriptions and developer websites for the availability of published evidence reporting app effectiveness. As there are potential risks associated with mobile health apps, such as inaccurate or out-of-date content [32], the SLP noted any potential safety issues (eg, app provided incorrect feedback to users) or other issues that could potentially hamper therapy progress.

A MARS mean score was calculated for each app in each category, and then a total MARS mean score was calculated from the first 4 categories, with 5 being the highest score possible. The final MARS score was the average score of the 2 raters for each app. The overall MARS score demonstrated a good [33] level of interrater reliability (two-way mixed intraclass correlation coefficient (ICC) 0.61, 95% CI 0.43-0.74) between raters. MARS subscale correlations included engagement, ICC 0.52 (95% CI 0.32-0.68); functionality, ICC 0.60 (95% CI 0.39-0.72); aesthetics, ICC 0.50 (95% CI 0.29-0.66); and information quality, ICC 0.62 (95% CI 0.45-0.75). According to the MARS tool guideline, the total MARS mean scores were used for comparison with consumer app star ratings in the app stores. All correlation data analyses were conducted using SPSS (IBM Corporation).

Results

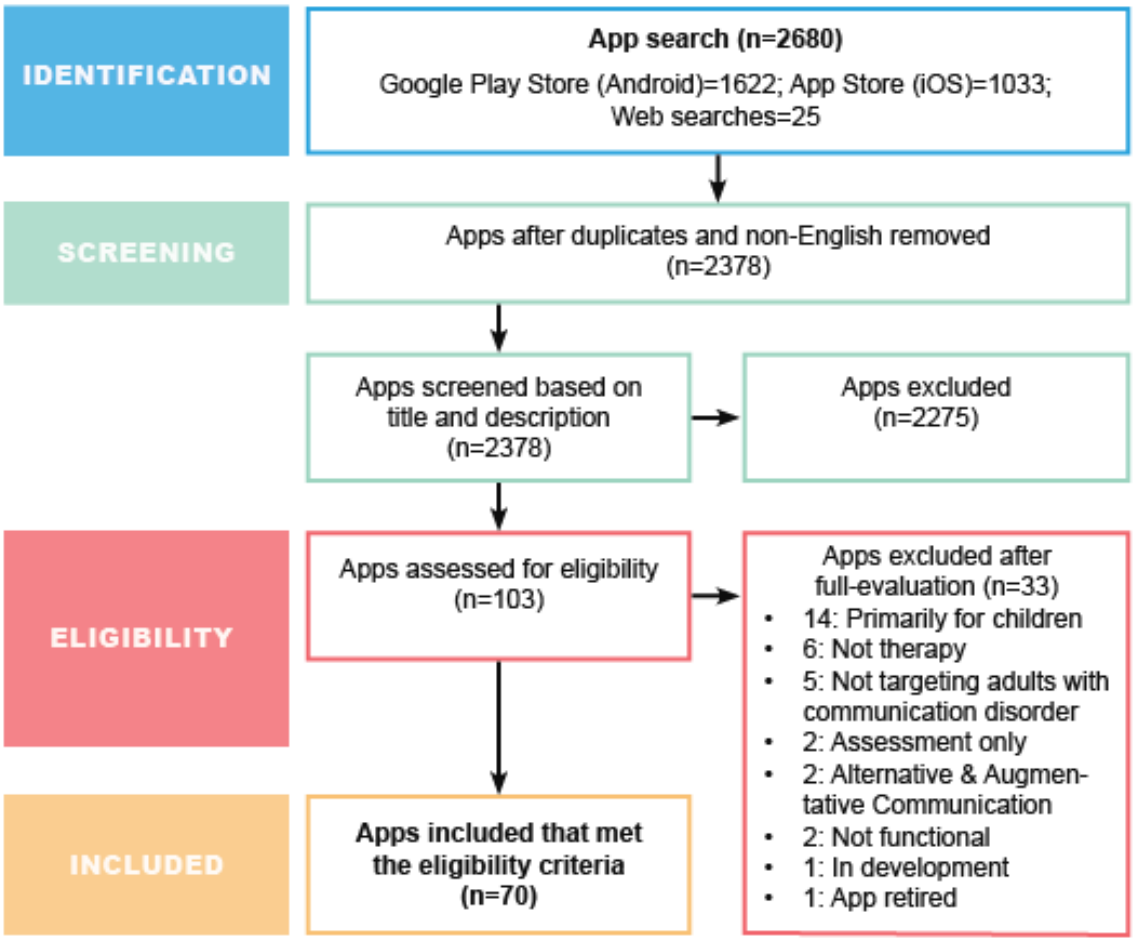
App Overview

From a total of 2680 apps identified from Google Play Store, Apple App Store, and web searches, 70 apps met the eligibility criteria. Figure 1 shows the search and selection process. Of the 70 apps included, 43 (61%) were available on the iOS platform, 14 (20%) were available on the Android platform, and 13 (19%) were available on both iOS and Android platforms. A total of 20% (14/70) of apps were completely free to use; 40% (28/70)

XSL•FO
RenderX

offered a free or lite version with the option for additional purchase or upgrade; and 34% (24/70) apps needed to be purchased, with prices ranging from Aus $1.99 to $57.99 (US $1.42 to $41.26). Overall, 26% (18/70) apps were available for purchase as a bundle with other apps for a discounted price. Furthermore, 50% (35/70) apps were developed by or in collaboration with SLPs. A total of 29 apps did not have a rating because of an insufficient number of users rating the app. The available star ratings ranged from 1 to 5, with a mean user rating of 3.7 out of 5 (SD 1.2). The Pearson product-moment correlation coefficient analysis was performed to assess the relationship between consumer star ratings and overall MARS ratings. The relationship was not significant (r=.052; n=37; *P*=.76). Multimedia Appendix 1 provides a summary of our findings, and Multimedia Appendix 2 lists the details of all the included apps.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) flow diagram of app selection.



App Content

The apps identified as primarily targeting adults with communication disorders after ABI were organized by therapeutic purposes (Multimedia Appendix 3). Overall, 43 apps targeted *language therapy*, defined as therapy for comprehending and producing written and spoken words and sentences. Among these language apps, 19 targeted spoken language expression, 15 targeted spoken language comprehension, 13 targeted reading, and 8 targeted writing. A total of 17 apps were designed for *speech therapy*, defined as therapy to improve perception and production of speech sounds and speech segments. This included 10 apps for articulation (individual speech sound production), 6 for motor speech (addressing problems with motor speech planning), and 2 for speech rate. Furthermore, 6 apps were used for *voice therapy*, including 5 addressing vocal loudness and 3 addressing vocal pitch. Overall, 8 apps were designed for *cognitive communication therapy* or cognitive skills underlying communication, such as problem solving, reasoning, inferencing, and executive functions. Finally, 5 apps were designated as *others*, including 3 apps for numbers/numeracy and 1 app for oromotor function (ie, function of oral musculature).

App Quality

Of the 70 apps, 66 (94%) were rated using the MARS tool. A total of 5 *talk around me apps* were rated together once, as they differed only with regard to the vocabulary items used. A summary of the MARS scores is included in Multimedia Appendix 4.

XSL•FO
RenderX

The average MARS rating was 3.7 out of 5 (SD 0.6; range 2.1-4.5). Overall, 27 apps achieved ratings between 4.0 and 4.9 (27/66, 41%), 30 apps were rated between 3.0 and 3.9 (30/66, 45%), whereas 11% (7/66) apps received ratings between 2.0 and 2.9. No app received a rating of 1 (inadequate). Overall, 86% (57/66) apps were rated as acceptable (3 out of 5) or higher, indicating favorable perceptions of the apps as perceived by certified SLPs.

The top 5 apps were *Naming Therapy*, *Speech Flipbook Standard* (4.6/5), *Number Therapy* (4.5/5), *Answering Therapy*, and *Constant Therapy* (4.4/5). The lowest-rated apps were *Aphasia: Start Talking Again* (2.2/5), *Aphasia Speech Therapy* (2.3/5), *HelpMeTalk*, and *Speech therapy Logopedic Free* (2.5/5). Across all apps identified in the review, *functionality* was the highest-rated subscale (mean 4.3, SD 0.6), followed by *aesthetics* (mean 3.8, SD 0.8), *information* (mean 3.4, SD 0.6), and *engagement* (mean 3.3, SD 0.6).

Engagement

The top-rated app on the engagement subscale was *Naming Therapy* (4.4/5), as it had a progress bar to introduce an element of gamification, a range of interesting cues, and a number of customization options (eg, ability to alter the number of trials, syllables, add categories/items). *Naming Therapy* also allowed the recording and emailing of the audio recording, provided self-rating options and cues on request, and utilized a simple layout and visual icons suitable for adults with communication disorders. In contrast, the lowest-rated app regarding engagement was *Aphasia Speech Therapy* (1.6/5), which lacked entertaining and engaging elements, did not have a setting page or method for customizing targets, and relied on written targets with no other visual elements to support adults with communication disorders. In addition, the voice recognition function of the *Aphasia Speech Therapy* app did not work at the time of rating (although it may have worked on other devices).

Functionality

A total of 13 apps received an excellent (5) rating on the functionality subscale for timely app performance, absence of technical issues, being easy and intuitive to use, employing a logical and clear layout, and consistent and intuitive gestural design. The lowest-rated app regarding functionality was *Aphasia: Start Talking Again* (2.1/5), as the speech recognition function was poor and failed to recognize many accurate speech productions. Therefore, the accuracy of the recorded data was poor. In addition, the feedback for speech evaluations was very general in nature and did not provide specific feedback.

Aesthetic

A total of 2 apps received an excellent (5) rating on the aesthetic subscale. An example of this was *Think Therapy*, which had a simple layout, high-quality graphics, a very attractive color scheme and visual motifs, and seamless animations. *Speech therapy Logopedic Free* was the lowest-rated app on this subscale (2/5). This rating was given because of the unusual layout involving menu items distributed across the screen, activities embedded within activities, its low quality, stylistically inconsistent graphics, and garish color scheme.

Information

The top 2 apps on the information subscale were *Naming Therapy* (4.5/5) and *Speech Flipbook Standard* (4.4/5). These contained detailed information about how to use the app that incorporated visual information (images of icons and buttons). The information contained in the *Naming Therapy* app is linked to evidence-based techniques, for example, cueing hierarchies and semantic feature analysis [34-36]. Only 4 apps referred to published empirical studies assessing app effectiveness, including one joint trial of *Naming Therapy*, *Reading Therapy*, *Writing Therapy*, and *Comprehension Therapy* and one trial of *Constant Therapy* [17,37]. It is worth mentioning that *Constant Therapy* was the only app to describe the use of a big data approach to collect patient data to personalize therapy and optimize outcomes based on what worked for patients with similar demographics and diagnoses [38]. In contrast, *Aphasia: Start Talking Again*, *HelpMeTalk*, and *Cognitive Rehabilitation 1-3* were rated lowest on this subscale because of poor description of purpose and goals.

Subjective Quality and Perceived Impact

Only 1 app, *Number Therapy*, received an excellent *subjective quality* rating. Other highly rated apps were *Apraxia Therapy*, *Category Therapy*, *Comprehension Therapy*, *Naming Therapy*, and *Talk Around It* (4.75/5). These apps share a combination of appealing qualities, including multiple ways of interacting with the app (eg, tapping a multiple-choice answer, recording yourself, emailing the recording to others, self-rating, requesting hints), clear intuitive layout, navigation and gestural design, high-quality attractive visuals, and easy ways to track progress and obtain accurate feedback.

Only 2 apps, *Category Therapy* and *Cognifit–Test and Brain Games*, received an excellent *perceived impact* rating. Other highly rated apps included *Naming Therapy*, *Advanced Comprehension Therapy*, *Advanced Naming Therapy*, *Advanced Reading Therapy*, *and Asking Therapy* (4.67/5). With regard to *perceived impact*, in general, apps did not receive ratings for awareness, attitudes, or help-seeking. Rather than building public awareness of communication disorders, changing user attitudes, or encouraging help-seeking, most apps were designed to assist individuals in working on specific communication skills and gaining awareness of their own strengths and weaknesses. Furthermore, apps designed to facilitate repetitive practice of specific communication skills appeared best suited for use as part of a more holistic therapy program directed by an SLP.

Apps that were rated highly on *perceived impact* generally functioned well, accurately rewarded target performance, provided multiple cues/prompts, and offered multiple options for presentation and response modalities. For example, while using *Advanced Comprehension Therapy*, the user can select auditory stimuli, written stimuli, or both; request hints; or request that the stimulus be repeated slowly. In contrast, low-rated apps provided inaccurate feedback (*Aphasia: Start Talking Again*) or rewarded incorrect productions (*Speech therapy Logopedic Free*). Other low-rated apps offered features that were not functional on the device used at the time of the review (eg, *vowel recognition* in *VowelViz Pro* practice function in *Aphasia Word*).

XSL•FO
RenderX

Discussion

Principal Findings

This study was designed to identify speech-language therapy mobile apps available in English for adults with communication disorders and evaluate their content and quality using the MARS tool [29]. A total of 70 apps were reviewed on iOS and/or Android platforms in the areas of *speech*, *language*, *voice*, *cognitive communication*,*oromotor function*, and/or *numeracy*. A high proportion of the reviewed apps had an average MARS rating of 3.7 out of 5, which is similar to that from previous research by Furlong et al [23] who reviewed 132 apps for children with speech disorders and reported the same average MARS rating of 3.7 out of 5.

This review revealed a lack of clear evidence to demonstrate the clinical benefits of speech-language therapy mobile apps currently available in Google Play Store or Apple App Store for adults with communication disorders. At instances where claims were made for clinical effectiveness, there was a lack of high-quality clinical trials to support these assertions. Overall, 3 of the top 10 apps claimed to be designed and developed based on evidence-based therapy techniques (*Naming Therapy*, *Number Therapy*, and *Apraxia Therapy)*. Despite these claims, only 2 published studies were found that evaluated the clinical effectiveness of these apps. A small pilot, crossover design study evaluated the use of 4 Tactus Therapy apps (*Naming Therapy*, *Comprehension Therapy*, *Reading Therapy*, and *Writing Therapy)* in patients with chronic expressive aphasia. This study found small but significant improvements in language outcomes as measured by a standardized aphasia battery and a narrative discourse measure. However, the number of participants who completed the study was small (n=10), and participants differed significantly in terms of aphasia severity at baseline [37]. In another nonrandomized study [17] involving 51 individuals with aphasia due to stroke or TBI, an iPad-based software platform, *Constant Therapy*, was trialed to ascertain its effects on specific therapy tasks and overall language and cognitive skills. Both the experimental group (n=42) and the control group (n=9) received individual face-to-face clinic sessions once a week for 10 weeks, which involved clinician-assisted delivery of *Constant Therapy* tasks. The experimental group was also asked to use *Constant Therapy* to practice language activities at home, whereas the control group did not do home practice. Small but significant improvements on a standardized aphasia battery were reported in both participant groups, with more significant gains evident in the experimental group. However, as the experimental group spent more time on therapy tasks, it is not clear whether the improvement was the result of more opportunities to use *Constant Therapy* or simply the result of more time spent practicing language. In addition, as the participant groups were not matched for severity level at baseline or with regard to time after onset of stroke/TBI, the results may have been confounded. Typically, a degree of improvement over time is expected, especially for individuals still within the first 12 months after onset. Therefore, none of the apps identified in this review had high-level evidence of clinical effectiveness. These findings have important implications for further research with a focus on evaluating the clinical effectiveness of mobile apps. In particular, with the recent worldwide COVID-19 pandemic, it is crucial to offer health care support remotely via digital health technologies [19].

This review showed that the included apps appeared to favor functionality (mean score 4.3, SD 0.6) over aesthetics (mean 3.8, SD 0.8), information (mean 3.4, SD 0.6) and engagement (mean 3.3, SD 0.6). This was surprising given the long-term engagement in speech-language therapy commonly required to gain real health benefits [39]. The idea that patients must be fully engaged in the rehabilitation process to achieve targeted outcomes has been considered analogous to patient participation in compliance with or adherence to this process. Engagement is also facilitated through the relationship and communication between the patient and clinician [39], which can vary considerably depending on individuals' communication impairment profiles. A possible explanation for the low scores regarding how well apps were able to achieve participant *engagement* might be the lack of human-centered theory-driven approaches [40,41]. Although prior studies on speech therapy mobile apps targeting children have used theory-driven or co-design approaches to improve engagement in therapy [42-44], this is lacking in the design of apps targeting adults with speech-language therapy. In particular, the apps should be tailored to suit individuals' treatment goals and demographics with a high level of attention to *human factors* [45] to achieve optimal outcomes. It may be challenging for SLPs, unfamiliar with these approaches and methodologies, to develop apps that achieve the desired levels of engagement. The results of this study are consistent with previous research on patient acceptance of consumer health information technology [46], which concluded that there is a need for developers and those who implement the systems to carefully consider the underlying reasons (eg, physical, psychological, and social) for patient acceptance and engagement with technology. A recent review of rehabilitation technology acceptance in adults with TBI also found limited research that comprehensively evaluated usability and user acceptance [47].

A multidisciplinary approach should be taken from the early stages of speech-language therapy app design and development and should involve SLPs, human-computer interaction researchers, user experience researchers, developers, and individuals with communication disorders. This approach would ensure the consideration of both human-centered design [40,48] and evidence-based therapy techniques with attention to the level of engagement, functionality, aesthetics, and information. Further research should be undertaken to investigate how people with communication disorders engage or disengage with mobile-based therapy apps over time in relation to their health goals. There are still unanswered questions about the long-term success of SLP apps accessed by adults with communication disorders.

This review identified limited gamification elements (eg, progress bar, scoring system, self-rating capabilities, ability to request cues/hints) in the included apps. Further research should be undertaken to investigate other game design elements to enhance engagement with rehabilitation technologies by adults with communication disorders. In particular, attention must be

XSL•FO
RenderX

paid to inclusive and accessible game design for people with different communication needs and to the inclusion of game elements that are relevant for users of different ages (ie, young vs older adults), experience with technology, cultures, languages, and regional dialects. Future research should also include non-English apps that could be used for bilingual/multilingual patients who should have the opportunity to be treated in their primary or preferred language, irrespective of its popularity. Finally, this review found a lack of relationship between the MARS rating and the in-app consumer star ratings. Similar findings were reported in the study by Knitza et al [30], where MARS had been used to evaluate apps in rheumatology. This is not surprising given that the MARS tool was developed to provide an objective and reliable multidimensional measure of the app quality of health-related apps, that is, engagement, functionality, aesthetics, information, and subjective quality. However, consumer star rating lacks objectivity, and its focus is on usability and popularity among consumers, with little indication of apps' clinical effectiveness [30].

Recommendations

The findings of this study have number of recommendations and practical implications.

- More research using a co-design process involving all stakeholders (ie, individuals with communication disorders, SLPs, caregivers) who will use or prescribe the end product is essential for developing apps targeting specific speech-language and communication impairments.
- It is essential to tailor apps to suit targeted patient group profiles by incorporating the capabilities to customize or adjust the content of the apps to simplify or increase complexity to match patient skill levels. For example, incorporation of gamification elements may be a priority to engage young patients with TBI, whereas for older patients affected by stroke, it may be more important to present age-appropriate content in an aphasia-friendly way through the use of visually distinct and appealing pictorial information and simplified language/text and formatting [49].
- Subject matter expert ratings for mobile health apps must be included to provide a more reliable measure of the app's quality, which could include ratings of engagement, functionality, aesthetics, information, and subjective quality.
- Owing to the increasing number of new app releases in Google Play Store and Apple App Store globally, it is recommended that professional bodies such as Speech-Language Pathology associations establish a database where app developers could register their apps. This could assist with the long-term management of mobile apps and support SLPs in decision making to recommend particular apps to their patients.

Limitations

A limitation of this review is that details regarding the number of app downloads and demographics of current users were not available, which would have been beneficial for in-depth analysis. Another limitation of this review is that the content and quality of the apps were evaluated solely from the perspective of SLPs and did not include consumers. However, most of these apps were designed to deliver speech-language therapy, where the guidance of SLPs would be instrumental in influencing user uptake. Further research should evaluate the usability of speech-language therapy app and the level of adoption by the target population. In addition, this paper did not review experimental apps currently reported only on journal articles, conference proceedings, or developer websites. Although such apps are not widely available, they may have more theory-driven designs and could show more promise of efficacy for use by adults with communication disorders. Future research focusing on these experimental mobile apps may be worthwhile.

Conclusions

The results of this study revealed a limited evidence base for speech-language therapy apps and a lack of highly engaging elements such as gamification techniques. Therefore, there is a need for apps to be developed to complement traditional speech-language therapy, utilizing interactive and engaging design elements to enhance user experience and optimize sustainable technology uptake. Furthermore, from the early stages, apps should be designed and developed by a multidisciplinary team of experts, including speech pathologists, human-computer interaction experts, user experience designers, and app developers to ensure a balance between clinically suitable content and positive user experience. Future studies should also ensure the design and development of speech-language therapy apps using patient-led co-design principles by involving adults with communication disorders as codevelopers. If this is done, mobile apps can have the potential to positively enhance the effectiveness and reach of long-term speech-language therapy for adults with communication disorders. Finally, adequately powered randomized controlled trials are needed to assess the efficacy of apps developed for long-term use in the management of adult communication disorders.

Acknowledgments

The authors would like to acknowledge Louise Cahill and Brooke-Mai Whelan for their input in the search strategy, Ms Jane Crombie for assisting with a quality appraisal, and Ms Danielle Aldridge for her valuable input in the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Summary of findings.
[PNG File , 480 KB - mhealth_v8i10e18858_app1.png]

Multimedia Appendix 2
Summary of the included apps.
[DOCX File , 41 KB - mhealth_v8i10e18858_app2.docx]

Multimedia Appendix 3
List of apps based on their therapeutic purpose.
[DOCX File , 30 KB - mhealth_v8i10e18858_app3.docx]

Multimedia Appendix 4
Mobile App Rating Scale Score by speech pathologists.
[DOCX File , 38 KB - mhealth_v8i10e18858_app4.docx]

References

1. Kat L, Schipper K, Knibbe J, Abma TA. Acquired brain injury. Br Med J 2010 Feb 19;340:c808. [doi: 10.1136/bmj.c808] [Medline: 20172926]
2. Dewan MC, Rattani A, Gupta S, Baticulon RE, Hung YC, Punchak M, et al. Estimating the global incidence of traumatic brain injury. J Neurosurg 2018 Apr 1:1-18. [doi: 10.3171/2017.10.JNS17352] [Medline: 29701556]
3. Strong K, Mathers C, Bonita R. Preventing stroke: saving lives around the world. Lancet Neurol 2007 Feb;6(2):182-187. [doi: 10.1016/S1474-4422(07)70031-5] [Medline: 17239805]
4. MacDonald S. Introducing the model of cognitive-communication competence: a model to guide evidence-based communication interventions after brain injury. Brain Inj 2017;31(13-14):1760-1780. [doi: 10.1080/02699052.2017.1379613] [Medline: 29064304]
5. Definitions of communication disorders and variations. Ad Hoc Committee on Service Delivery in the Schools. American Speech-Language-Hearing Association. ASHA Suppl 1993 Mar;35(3 Suppl 10):40-41. [Medline: 8097650]
6. Ponsford JL, Downing MG, Olver J, Ponsford M, Acher R, Carty M, et al. Longitudinal follow-up of patients with traumatic brain injury: outcome at two, five, and ten years post-injury. J Neurotrauma 2014 Jan 1;31(1):64-77. [doi: 10.1089/neu.2013.2997] [Medline: 23889321]
7. Flowers HL, Skoretz SA, Silver FL, Rochon E, Fang J, Flamand-Roze C, et al. Poststroke aphasia frequency, recovery, and outcomes: a systematic review and meta-analysis. Arch Phys Med Rehabil 2016 Dec;97(12):2188-201.e8. [doi: 10.1016/j.apmr.2016.03.006] [Medline: 27063364]
8. Struchen MA, Pappadis MR, Sander AM, Burrows CS, Myszka KA. Examining the contribution of social communication abilities and affective/behavioral functioning to social integration outcomes for adults with traumatic brain injury. J Head Trauma Rehabil 2011;26(1):30-42. [doi: 10.1097/HTR.0b013e3182048f7c] [Medline: 21209561]
9. Worrall L, Ryan B, Hudson K, Kneebone I, Simmons-Mackie N, Khan A, et al. Reducing the psychosocial impact of aphasia on mood and quality of life in people with aphasia and the impact of caregiving in family members through the Aphasia Action Success Knowledge (Aphasia ASK) program: study protocol for a randomized controlled trial. Trials 2016 Mar 22;17:153 [FREE Full text] [doi: 10.1186/s13063-016-1257-9] [Medline: 27005901]
10. Savundranayagam MY, Hummert ML, Montgomery RJ. Investigating the effects of communication problems on caregiver burden. J Gerontol B Psychol Sci Soc Sci 2005 Jan;60(1):S48-S55. [doi: 10.1093/geronb/60.1.s48] [Medline: 15643047]
11. Brady MC, Kelly H, Godwin J, Enderby P, Campbell P. Speech and language therapy for aphasia following stroke. Cochrane Database Syst Rev 2016 Jun 1(6):CD000425. [doi: 10.1002/14651858.CD000425.pub4] [Medline: 27245310]
12. Ballard KJ, Wambaugh JL, Duffy JR, Layfield C, Maas E, Mauszycki S, et al. Treatment for acquired apraxia of speech: a systematic review of intervention research between 2004 and 2012. Am J Speech Lang Pathol 2015 May 1;24(2):316-337. [doi: 10.1044/2015_AJSLP-14-0118] [Medline: 25815778]
13. Finch E, Copley A, Cornwell P, Kelly C. Systematic review of behavioral interventions targeting social communication difficulties after traumatic brain injury. Arch Phys Med Rehabil 2016 Aug;97(8):1352-1365. [doi: 10.1016/j.apmr.2015.11.005] [Medline: 26679234]
14. Choi Y, Park HK, Paik N. A telerehabilitation approach for chronic aphasia following stroke. Telemed J E Health 2016 May;22(5):434-440. [doi: 10.1089/tmj.2015.0138] [Medline: 26544535]
15. Choi Y, Park HK, Ahn K, Son Y, Paik N. A telescreening tool to detect aphasia in patients with stroke. Telemed J E Health 2015 Sep;21(9):729-734. [doi: 10.1089/tmj.2014.0207] [Medline: 25942492]
16. Munsell M, De Oliveira E, Saxena S, Godlove J, Kiran S. Closing the digital divide in speech, language, and cognitive therapy: cohort study of the factors associated with technology usage for rehabilitation. J Med Internet Res 2020 Feb 7;22(2):e16286 [FREE Full text] [doi: 10.2196/16286] [Medline: 32044752]

XSL•FO
RenderX

17. Des Roches CA, Balachandran I, Ascenso EM, Tripodis Y, Kiran S. Effectiveness of an impairment-based individualized rehabilitation program using an iPad-based software platform. Front Hum Neurosci 2014;8:1015 [FREE Full text] [doi: 10.3389/fnhum.2014.01015] [Medline: 25601831]
18. Cardullo S, Gamberini L, Milan S, Mapelli D. Rehabilitation tool: a pilot study on a new neuropsychological interactive training system. Stud Health Technol Inform 2015;219:168-173. [Medline: 26799901]
19. Pérez Sust P, Solans O, Fajardo JC, Medina Peralta M, Rodenas P, Gabaldà J, et al. Turning the crisis into an opportunity: digital health strategies deployed during the covid-19 outbreak. JMIR Public Health Surveill 2020 May 4;6(2):e19106 [FREE Full text] [doi: 10.2196/19106] [Medline: 32339998]
20. Dickson S, Barbour RS, Brady M, Clark AM, Paton G. Patients' experiences of disruptions associated with post-stroke dysarthria. Int J Lang Commun Disord 2008;43(2):135-153. [doi: 10.1080/13682820701862228] [Medline: 18283594]
21. Muñoz ML, Hoffman LM, Brimo D. Be smarter than your phone: a framework for using apps in clinical practice. Contemp Issues Commun Sci Disord 2013 Oct;40(Fall):138-150. [doi: 10.1044/cicsd_40_f_138]
22. Furlong LM, Morris ME, Erickson S, Serry TA. Quality of mobile phone and tablet mobile apps for speech sound disorders: protocol for an evidence-based appraisal. JMIR Res Protoc 2016 Nov 29;5(4):e233 [FREE Full text] [doi: 10.2196/resprot.6505] [Medline: 27899341]
23. Furlong L, Morris M, Serry T, Erickson S. Mobile apps for treatment of speech disorders in children: an evidence-based analysis of quality and efficacy. PLoS One 2018;13(8):e0201513 [FREE Full text] [doi: 10.1371/journal.pone.0201513] [Medline: 30092095]
24. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. Arch Phys Med Rehabil 2019 Jan;100(1):115-127. [doi: 10.1016/j.apmr.2018.07.439] [Medline: 30171827]
25. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification: Using Game-design Elements in Non-Gaming Contexts. In: CHI '11 Extended Abstracts on Human Factors in Computing Systems. 2011 Presented at: CHI EA'11; May 7-12, 2011; Vancouver, BC, Canada. [doi: 10.1145/1979742.1979575]
26. Deterding S, Björk S, Nacke L, Dixon D, Lawley E, editors. Designing Gamification: Creating Gameful and Playful Experiences. In: CHI'13 Extended Abstracts on Human Factors in Computing Systems. 2013 Presented at: CHI EA'13; April 27-May 2, 2013; Paris, France. [doi: 10.1145/2468356.2479662]
27. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
28. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Jul;71:31-48 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
29. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
30. Knitza J, Tascilar K, Messner E, Meyer M, Vossen D, Pulla A, et al. German mobile apps in rheumatology: review and analysis using the mobile application rating scale (MARS). JMIR Mhealth Uhealth 2019 Aug 5;7(8):e14991 [FREE Full text] [doi: 10.2196/14991] [Medline: 31381501]
31. MARS Training Video. YouTube. 2016. URL: https://www.youtube.com/watch?v=25vBwJQIOcE [accessed 2020-05-05]
32. Lewis TL, Wyatt JC. mHealth and mobile medical Apps: a framework to assess risk and promote safer use. J Med Internet Res 2014 Sep 15;16(9):e210 [FREE Full text] [doi: 10.2196/jmir.3133] [Medline: 25223398]
33. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment 1994;6(4):284-290. [doi: 10.1037/1040-3590.6.4.284]
34. Boyle M. Semantic feature analysis treatment for aphasic word retrieval impairments: what's in a name? Top Stroke Rehabil 2010;17(6):411-422. [doi: 10.1310/tsr1706-411] [Medline: 21239365]
35. Leonard C, Rochon E, Laird L. Treating naming impairments in aphasia: findings from a phonological components analysis treatment. Aphasiology 2008 Sep;22(9):923-947. [doi: 10.1080/02687030701831474]
36. Wisenburn B, Mahoney K. A meta-analysis of word-finding treatments for aphasia. Aphasiology 2009 Nov 10;23(11):1338-1352. [doi: 10.1080/02687030902732745]
37. Stark BC, Warburton EA. Improved language in chronic aphasia after self-delivered iPad speech therapy. Neuropsychol Rehabil 2018 Jul;28(5):818-831. [doi: 10.1080/09602011.2016.1146150] [Medline: 26926872]
38. Kiran S. How does severity of aphasia influence individual responsiveness to rehabilitation? Using big data to understand theories of aphasia rehabilitation. Semin Speech Lang 2016 Feb;37(1):48-60. [doi: 10.1055/s-0036-1571358] [Medline: 26882364]
39. Bright FA, Kayes NM, McPherson KM, Worrall LE. Engaging people experiencing communication disability in stroke rehabilitation: a qualitative study. Int J Lang Commun Disord 2018 Sep;53(5):981-994. [doi: 10.1111/1460-6984.12409] [Medline: 30003629]
40. MAGUIRE M. Methods to support human-centred design. Int J Hum-Comput Stud 2001 Oct;55(4):587-634. [doi: 10.1006/ijhc.2001.0503]

XSL•FO
RenderX

41. Norman D, Draper S. User Centered System Design: New Perspectives on Human-computer Interaction. Hillsdale, NJ: Erlbaum Associates; 1986.
42. Ahmed B, Monroe P, Hair A, Tan CT, Gutierrez-Osuna R, Ballard KJ. Speech-driven mobile games for speech therapy: user experiences and feasibility. Int J Speech Lang Pathol 2018 Nov;20(6):644-658. [doi: 10.1080/17549507.2018.1513562] [Medline: 30301384]
43. Hair A, Monroe P, Ahmed B, Ballard K, Gutierrez-Osuna R. Apraxia World: a Speech Therapy Game for Children With Speech Sound Disorders. In: 17th ACM Conference on Interaction Design and Children. 2018 Presented at: IDC'18; June 1-4, 2018; Trondheim, Norway. [doi: 10.1145/3202185.3202733]
44. Lopes M, Magalhães J, Cavaco S. A Voice-controlled Serious Game for the Sustained Vowel Exercise. In: 13th International Conference on Advances in Computer Entertainment Technology. 2016 Presented at: ACET'16; November 9-12, 2018; Osaka, Japan. [doi: 10.1145/3001773.3001807]
45. Carayon P. Human factors of complex sociotechnical systems. Appl Ergon 2006 Jul;37(4):525-535. [doi: 10.1016/j.apergo.2006.04.011] [Medline: 16756937]
46. Or CK, Karsh B. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.M2888] [Medline: 19390112]
47. Vaezipour A, Whelan B, Wall K, Theodoros D. Acceptance of rehabilitation technology in adults with moderate to severe traumatic brain injury, their caregivers, and healthcare professionals: a systematic review. J Head Trauma Rehabil 2019;34(4):E67-E82. [doi: 10.1097/HTR.0000000000000462] [Medline: 30608310]
48. Putnam C, Reiner A, Ryou E, Caputo M, Cheng J, Allen M, et al. Human-centered design in practice. J Tech Writ Commun 2016 Aug 18;46(4):446-470. [doi: 10.1177/0047281616653491]
49. Togher L, Power E, Tate R, McDonald S, Rietdijk R. Measuring the social interactions of people with traumatic brain injury and their communication partners: The adapted Kagan scales. Aphasiology 2010 Feb 03;24(6-8):914-927. [doi: 10.1080/02687030903422478]

Abbreviations

ABI: acquired brain injury
ICC: intraclass correlation coefficient
iOS: iPhone operating system
MARS: Mobile App Rating Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
SLP: speech-language pathologist
TBI: traumatic brain injury

Edited by G Eysenbach; submitted 23.03.20; peer-reviewed by DC Burns, P Ballard, A Rezaei Aghdam, M Pascoe, B Kane; comments to author 12.06.20; revised version received 02.09.20; accepted 27.09.20; published 29.10.20.

Please cite as:
Vaezipour A, Campbell J, Theodoros D, Russell T
Mobile Apps for Speech-Language Therapy in Adults With Communication Disorders: Review of Content and Quality
JMIR Mhealth Uhealth 2020;8(10):e18858
URL: http://mhealth.jmir.org/2020/10/e18858/
doi:10.2196/18858
PMID:33118953

©Atiyeh Vaezipour, Jessica Campbell, Deborah Theodoros, Trevor Russell. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 29.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Utility of Smartphone-Based Sensor Assessments in Persons With Multiple Sclerosis in the Real-World Using an App (elevateMS): Observational, Prospective Pilot Digital Health Study

Abhishek Pratap[1], PhD; Daniel Grant[2], PhD; Ashok Vegesna[2], PharmD, MS, RPh; Meghasyam Tummalacherla[1], MS; Stanley Cohan[3], MD, PhD; Chinmay Deshpande[2], PhD; Lara Mangravite[1], PhD; Larsson Omberg[1], PhD

[1]Sage Bionetworks, Seattle, WA, United States

[2]Novartis Pharmaceuticals Corporation, East Hanover, NJ, United States

[3]Providence Multiple Sclerosis Center, Providence St Vincent Medical Center, Portland, OR, United States

Corresponding Author:
Larsson Omberg, PhD
Sage Bionetworks
2901 Third Avenue, Suite 330
Seattle, WA, 98121
United States
Phone: 1 (206) 928 8260
Email: larsson.omberg@sagebionetworks.org

Abstract

Background: Multiple sclerosis (MS) is a chronic neurodegenerative disease. Current monitoring practices predominantly rely on brief and infrequent assessments, which may not be representative of the real-world patient experience. Smartphone technology provides an opportunity to assess people's daily-lived experience of MS on a frequent, regular basis outside of episodic clinical evaluations.

Objective: The objectives of this study were to evaluate the feasibility and utility of capturing real-world MS-related health data remotely using a smartphone app, "elevateMS," to investigate the associations between self-reported MS severity and sensor-based active functional tests measurements, and the impact of local weather conditions on disease burden.

Methods: This was a 12-week, observational, digital health study involving 3 cohorts: self-referred participants who reported an MS diagnosis, clinic-referred participants with neurologist-confirmed MS, and participants without MS (controls). Participants downloaded the elevateMS app and completed baseline assessments, including self-reported physical ability (Patient-Determined Disease Steps [PDDS]), as well as longitudinal assessments of quality of life (Quality of Life in Neurological Disorders [Neuro-QoL] Cognitive, Upper Extremity, and Lower Extremity Function) and daily health (MS symptoms, triggers, health, mobility, pain). Participants also completed functional tests (finger-tapping, walk and balance, voice-based Digit Symbol Substitution Test [DSST], and finger-to-nose) as an independent assessment of MS-related cognition and motor activity. Local weather data were collected each time participants completed an active task. Associations between self-reported baseline/longitudinal assessments, functional tests, and weather were evaluated using linear (for cross-sectional data) and mixed-effects (for longitudinal data) regression models.

Results: A total of 660 individuals enrolled in the study; 31 withdrew, 495 had MS (n=359 self-referred, n=136 clinic-referred), and 134 were controls. Participation was highest in clinic-referred versus self-referred participants (median retention: 25.5 vs 7.0 days). The top 5 most common MS symptoms, reported at least once by participants with MS, were fatigue (310/495, 62.6%), weakness (222/495, 44.8%), memory/attention issues (209/495, 42.2%), and difficulty walking (205/495, 41.4%), and the most common triggers were high ambient temperature (259/495, 52.3%), stress (250/495, 50.5%), and late bedtime (221/495, 44.6%). Baseline PDDS was significantly associated with functional test performance in participants with MS (mixed model–based estimate of most significant feature across functional tests [β]: finger-tapping: β=–43.64, P<.001; DSST: β=–5.47, P=.005; walk and balance: β=–.39, P=.001; finger-to-nose: β=.01, P=.01). Longitudinal Neuro-QoL scores were also significantly associated with functional tests (finger-tapping with Upper Extremity Function: β=.40, P<.001; walk and balance with Lower Extremity

XSL•FO
RenderX

Function: β=–99.18, *P*=.02; DSST with Cognitive Function: β=1.60, *P*=.03). Finally, local temperature was significantly associated with participants' test performance (finger-tapping: β=–.14, *P*<.001; DSST: β=–.06, *P*=.009; finger-to-nose: β=–53.88, *P*<.001).

Conclusions: The elevateMS study app captured the real-world experience of MS, characterized some MS symptoms, and assessed the impact of environmental factors on symptom severity. Our study provides further evidence that supports smartphone app use to monitor MS with both active assessments and patient-reported measures of disease burden. App-based tracking may provide unique and timely real-world data for clinicians and patients, resulting in improved disease insights and management.

(JMIR Mhealth Uhealth 2020;8(10):e22108) doi:10.2196/22108

KEYWORDS

multiple sclerosis; digital health; real-world data; real-world evidence; remote monitoring; smartphone; mobile phone; neurodegeneration

Introduction

Multiple Sclerosis

Multiple sclerosis (MS) is a chronic neurodegenerative disease that affects more than 2 million people worldwide, with prevalence rates equating to greater than 400,000 cases in the United States [1,2]. Symptoms of MS can affect motor function, sensation, cognition, and mood [3], and substantially impact quality of life (QoL) [4]. The combination of MS symptoms, their severity, and the course of disease varies between individual patients and can be affected by environmental factors, such as temperature, vitamin D stores, stress [1,5-10], and comorbidities, including depression, diabetes, cardiovascular disease, cancer, and autoimmune conditions [11,12]. Despite the heterogeneity of MS, certain symptoms and triggers are common among patients; for example, fatigue is reported as a symptom in approximately 75% of patients [13], and elevated temperature is estimated to cause transient symptom worsening in up to 80% [9].

Patient Monitoring and Digital Tools

Routine clinical care for MS typically involves brief assessments performed during infrequent neurologist visits and, hence, often relies on retrospective self-reporting of symptoms and treatment responses, which can be subject to recall bias or affected by MS-associated cognitive impairment [14-17]. As such, this approach may fail to fully capture an individual's day-to-day experience of living with MS and contribute to reduced accuracy and timeliness in detecting changes in symptom burden, disease severity/relapses, therapeutic outcomes, and the need for timely therapeutic agent change [17,18], Furthermore, while patient-reported outcomes (PROs) are increasingly used in MS clinical care, there is a lack of universal guidance on MS-specific PROs, and their usage and interpretation can differ between individual clinicians [17]. These challenges in monitoring MS highlight an unmet need for more effective, patient-centered tools that are able to capture the daily lived experience of disease outside of episodic clinic visits [19].

While a number of studies have used web-based tools to collect MS health data through patient diaries and electronic PROs [20,21], there is a growing need to tailor these digital health tools to the needs of patients with MS; this includes developing sensor-based assessments of MS disease severity and evaluating the impact of environmental factors, such as weather, stress, and sleep impairment on MS burden [22,23]. By bridging the gap between episodic clinical observation and the real-world experience of MS, these digital tools have the potential to improve self-reporting and disease monitoring, and provide a more comprehensive assessment of disease trajectories. In turn, this may support clinicians in making disease management recommendations to patients with MS, as well as other diseases that have heterogeneous and variable symptoms over time [24].

The ubiquity of smartphones with built-in sensors provides an opportunity to address the growing need for real-time disease monitoring. A number of previous studies have demonstrated the feasibility of smartphones in collecting health data in a real-time, real-world setting from patients across a range of disease areas including asthma, diabetes, depression, and Parkinson disease [25-30]. In addition to monitoring symptoms and triggers, these studies have identified geographic and environmental factors related to disease severity, and have been reported by patients to have a positive impact on their disease management [26,27,29]. Several digital health studies have already been undertaken in patients with MS with encouraging results; data suggest that smartphone technology can be effectively leveraged to monitor MS symptom severity, QoL, and medication usage, enabling patients to play an active role in disease management [23,31-35]. The recent FLOODLIGHT study has also shown that smartphone-based active testing can be used to remotely monitor motor function and capture MS symptoms, thus providing a more accurate assessment of MS in the real world [23,34]. However, a number of these studies have involved partially remote designs that include scheduled clinic visits at predetermined time points [32,34], which may limit widespread usage and participation. To this end, additional studies are required to build on these existing data and further assess digital health tools in a large, remote population of patients with MS.

Objective

The main objective of this study was to evaluate the feasibility and utility of gathering MS-related health information from a large, remotely enrolled cohort using a dedicated smartphone app and to monitor study participants over a 12-week period. The study app, "elevateMS," was developed through a user-centered design process. Secondary objectives were to examine the relationship between disease severity and QoL, measured via PROs and performance in sensor-based active functional tests, and to investigate the impact of local weather conditions on variations in MS symptoms and severity.

XSL•FO
RenderX

Methods

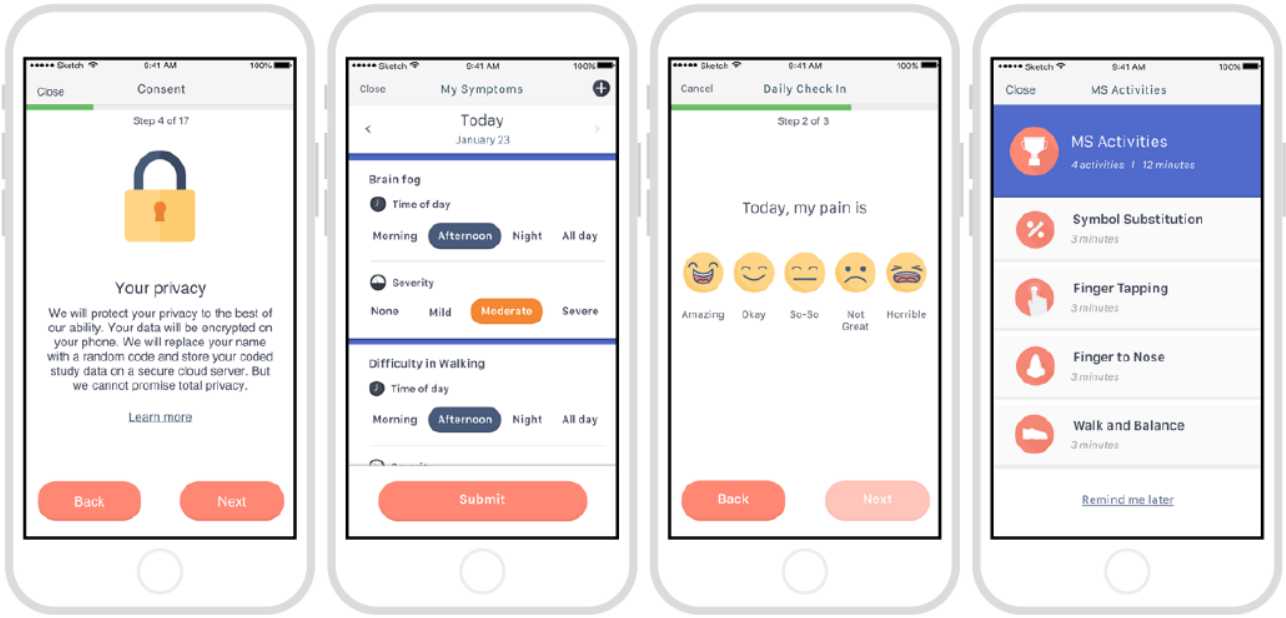
Study Design

This was a 12-week observational, prospective pilot digital health study using data collected from a dedicated smartphone app, elevateMS. The app was developed using a patient-centered design process with MS patient advisors (Multimedia Appendix 1) and was freely available to download from the Apple App Store. Enrollment spanned from August 2017 to October 2019, and participants were required to be aged 18 years or older, reside in the United States, and use an iPhone 5 or newer device.

Study Participants

Participants were openly recruited through word of mouth, press releases, online advertisements, and the study website [36] and grouped into 2 cohorts: individuals without MS (controls) and individuals who self-reported diagnosis of MS (self-referred). A third, "clinic-referred" cohort, with a neurologist-confirmed MS diagnosis, was also recruited through information flyers distributed at 3 MS treatment centers. Ethical approval was granted by the Western Institutional Review Board, and enrollment, informed consent, and data collection were carried out electronically through the study app (Figure 1) [37].

Figure 1. Example screenshots from the elevateMS study app.



Participants had to make an active choice to complete the consent process, and no default option was presented. Participants were also given the option to share their data only with the elevateMS study team and partners (share narrowly), or more broadly with qualified researchers worldwide [38].

Data Collection

elevateMS primarily targeted collection of real-world data from participants with MS. This included self-reported measures of symptoms and health via optional "check-in" surveys, and independent assessments of motor function via sensor-based active functional tests. Participants were encouraged to complete check-in surveys on a daily basis and were notified to perform more comprehensive functional tests once a week. Local weather data were collected every time an assessment was performed. The data collected through elevateMS and the frequency at which each element was recorded are summarized in Table 1. With the exception of overall physical ability, which was a baseline-only assessment, all data were collected longitudinally at various intervals over the 12-week study duration.

XSL•FO
RenderX

Table 1. Data collected through the elevateMS study app.

Data source	Timeline of data collection
Baseline demographics	
Sociodemographic data (age, gender, race, education, health insurance, employment status, geographic location)	Day 2
MS[a] disease characteristics (diagnosis, medication, family history)	Day 1
Patient-reported outcomes	
Overall physical ability[b]	Day 2
Check-in survey: MS symptoms and triggers	Daily
Check-in survey: health, mobility, pain[c]	Daily
Short-form Neuro-QoL[d] domains	Every third functional test targeting that domain (Cognitive Function domain: DSST; Upper Extremity Function domain: finger-tapping/finger-to-nose; Lower Extremity Function domain: walk and balance)
Active functional test	
Finger-tapping	Weekly[f]
Walk and balance	Weekly[f]
DSST[e]	Weekly[f]
Finger-to-nose	Weekly[f]
Local weather data	
Temperature	Every time a test was performed
Humidity	Every time a test was performed
Cloud coverage	Every time a test was performed
Atmospheric pressure	Every time a test was performed

[a]MS: multiple sclerosis.

[b]Based on a truncated 4-point Patient-Determined Disease Steps scale, administered at baseline to all patients with MS (Normal, Mild Disability, Moderate Disability, Gait Disability).

[c]Based on a 5-point Likert scale (Health: Amazing, Okay, So-So, Not great, Horrible; Mobility: Excellent, Very good, Good, Not great, Horrible; Pain: None, Mild, Moderate, Severe, Horrible).

[d]Neuro-QoL: Quality of Life in Neurological Disorders.

[e]DSST: Digit Symbol Substitution Test.

[f]With option for participants to complete more frequently.

Participants with MS completed optional daily check-in surveys to record their individual symptoms and triggers, and to assess their health, mobility, and pain using a 5-point Likert scale. Self-reported disease severity was determined using a truncated 4-point Patient-Determined Disease Steps (PDDS) scale to assess overall physical ability (Normal, Mild Disability, Moderate Disability, Gait Disability) [39]. The impact of MS on daily function was also assessed using 3 short-form domains of the previously validated Quality of Life in Neurological Disorders measurement tool (Neuro-QoL; Cognitive Function, Upper Extremity Function, and Lower Extremity Function; Multimedia Appendix 2) [40-42]. Raw Neuro-QoL scores were converted to standardized Theta (T) scores using standard scoring protocols [42], and subsequently classified into discrete, clinically relevant categories (Normal, Mild, Moderate, and Severe) based on past research [43-45]. All participants carried out active functional tests, which used smartphone sensors as a proxy for traditional symptom measurements. These tests included:

1. The finger-tapping test, where participants repeatedly tapped between 2 circles with alternating fingers as fast as possible for 20 seconds, to measure dexterity, speed, and abnormality in movement.
2. The walk and balance test, where participants walked for 20 seconds with their iPhone in their pocket, then stood still for 10 seconds, to assess gait, posture, stability, and balance;
3. The voice-controlled Digit Symbol Substitution Test (DSST) [46], where participants used the microphone to record answers to measure cognitive function.
4. The finger-to-nose test, where participants extended their arm while holding their iPhone and then touched the phone screen to their nose repeatedly, to measure kinetic tremor and dysmetria in each hand.

XSL•FO
RenderX

Raw data collected from sensor-based active functional tests were transformed using the mhealthtools package, an open-source feature engineering pipeline [47]. This process generated features for each functional test related to different aspects of a participant's health state; for example, the finger-tapping test comprised over 40 features, including number of taps, frequency, and location drift. See Multimedia Appendix 3 for specific examples and the elevateMS Feature Definitions webpage for a full list of features [48]. In addition to extracting features, the mhealthtools pipeline also filtered out records lacking data [47].

The data contributed by participants, as well as scheduling of in-app active functional tests, were managed using an open-source platform developed and maintained by Sage Bionetworks [49].

Statistical Analyses

Descriptive statistics were used to summarize and compare the baseline demographics and MS disease–related characteristics in the study cohort.

User-engagement data were collected and analyzed to understand participant demographics, retention, and compliance in the study. Retention analysis was performed under the definition that participants were considered active in the study if they completed at least one survey or sensor-based test in a given week. The total duration in the study was determined by the number of days between the first and last test. Participants' weekly compliance was assessed using a more stringent cut off than weekly retention; a participant was considered minimally compliant if he/she completed at least one out of four weekly sensor-based active functional tests. Overall retention (ie, total duration a participant remained in the study) was examined across the 3 cohorts (controls, self-referred participants with MS, and clinic-referred participants with MS) using Kaplan–Meier plots. A log-rank test was used to compare the retention difference between the 3 cohorts. The impact of baseline demographics and MS disease characteristics on participant retention was assessed using a Cox proportional hazards model. Each covariate of interest was tested independently, including an interaction term for the clinical referral status. The assumption of the proportional hazard model was tested using scaled Schoenfeld residuals. Finally, per the study protocol, participants were expected to remain in the study for 12 weeks, and thus all user-engagement analysis was limited to the first 12 weeks of study participation.

Linear regression models were used to test for association between participants' self-reported demographics, baseline physical ability (collected once during the onboarding), and the median value of all features generated for each of the 4 sensor-based active functional tests. To test for association between longitudinal PRO assessments (ie, Neuro-QoL results and daily check-ins) and active functional test performance, as well as the potential impact of local weather conditions, a linear mixed-effects (LME) modeling approach was used to account for the subject-level heterogeneity. Prior to modeling, the PRO data were aligned with more frequently administered functional tests by aggregating the average value for all features per week. LME models were fit using the R package lme4, version 1.1-23 [50] with combinations of fixed and random effects. Because of a significant amount of missing responses in sociodemographic information for participants (see Multimedia Appendix 4), some LME models did not converge, and therefore a simpler LME model accounting for participant-level random effect only was used. Statistical significance (*P*-values) for LME models was determined using the Satterthwaite degrees of freedom method through the lmerTest package (version 3.1-1) [51]. For analysis conducted using LME models, we report estimates (β) of all fixed effects covariates along with *P*-values. In addition, *P*-values from ANOVA test conducted to assess significant differences in LME model fixed effect estimates (regardless of the individual factor levels) are also reported. In case a functional test was due to be completed by both hands, the LME models also accounted for variations due to left- and right-hand differences. As participants with MS completed active functional tests at different frequencies, participation rates, sample size, and number of data points varied between analyses. We also conducted sensitivity analyses to evaluate the impact of extreme Neuro-QoL and health, mobility, and pain categories on association results; this involved excluding functional test scores that mapped to the Severe Neuro-QoL category and excluding health, mobility, and pain scores that mapped to the Horrible category. All *P*-values were corrected for multiple testing and false positives using the Benjamini–Hochberg procedure. All analyses were performed using open-source statistical analysis framework R (version 3.5.2; R Foundation for Statistical Computing) [52].

Data Availability

Complete results from this analysis are available online through the accompanying elevateMS study portal [53]. Additionally, individual user-level raw data for those participants who consented to share their data broadly with qualified researchers worldwide is also available under controlled access through the study portal [54].

Results

Study Population

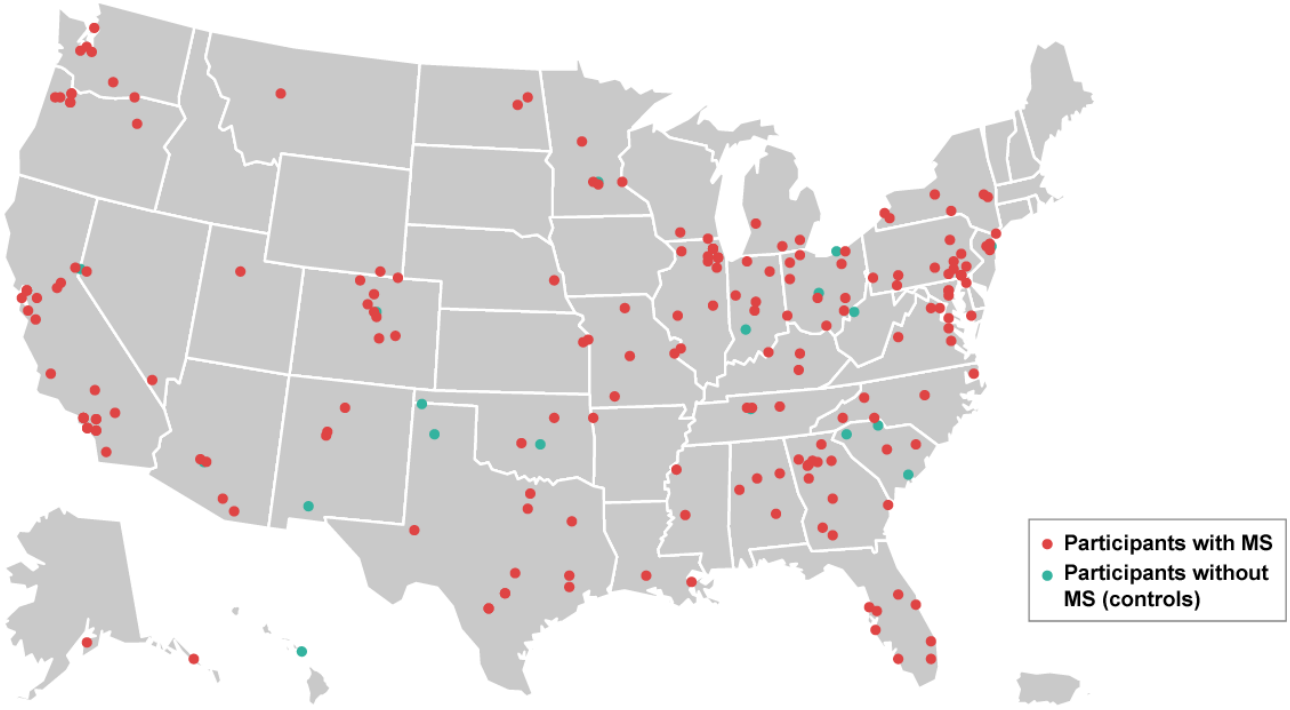
The elevateMS app was released in August 2017 through the Apple App Store and enrolled participants on a rolling basis until October 2019. A total of 660 participants enrolled in the study, of which 31 selected to withdraw with no reason provided. Of the remaining 629 participants, 134 (21.3%) were controls (self-reported as not having MS) and 495 (78.7%) were participants with MS. Of the 495 participants with MS, 359 (72.5%) self-referred to the study with a self-reported MS diagnosis and 136 (27.5%) were referred from 3 clinical sites and had a neurologist-confirmed MS diagnosis. Participants were located across the United States (Figure 2), with a mean (SD) age of 39.34 (11.41), 45.20 (11.64), and 48.93 (11.20) years in the control, self-referred, and clinic-referred cohorts, respectively. A summary of baseline sociodemographic data is presented in Table 2. Further information on missing responses in demographic data is presented in Multimedia Appendix 4.

Baseline disease characteristics for the self-referred and clinic-referred participants with MS are shown in Table 3. Most participants reported relapsing–remitting MS; this included

XSL•FO
RenderX

83.6% (300/359) of the self-referred cohort and 90.4% (123/136) of the clinic-referred cohort. Infusion disease-modifying therapy was the most common treatment received by both self-referred (116/359, 32.3%) and clinic-referred (64/136, 47.1%) participants.

Figure 2. Geographic locations of participants. Dots (n=329) represent the location of those participants who continued in the study beyond initial enrollment and provided the first three digits of their zip code during the collection of demographic information on Day 2. One dot is included for each location, with participants with the same first three digits of the zip code shown under the same dot.



XSL•FO
RenderX

Table 2. Baseline sociodemographic characteristics of study participants.[a]

Characteristic	Controls (N=134)	Participants with MS[b] (self-referred; N=359)	Participants with MS (clinic-referred; N=136)
Age (years), mean (SD)	39.34 (11.41)	45.20 (11.64)	48.93 (11.20)
Gender			
Female	27 (64.3)	154 (73.3)	78 (84.8)
Male	15 (35.7)	56 (26.7)	14 (15.2)
Race			
Asian	6 (13.6)	4 (1.9)	0 (0.0)
Black African	1 (2.3)	13 (6.1)	9 (9.8)
Caucasian	26 (59.1)	182 (85.4)	74 (80.4)
Latino Hispanic	5 (11.4)	9 (4.2)	5 (5.4)
Other	6 (13.6)	5 (2.3)	4 (4.3)
Education			
College degree	16 (37.2)	123 (57.5)	55 (60.4)
High-school diploma/GED[c]	4 (9.3)	16 (7.5)	7 (7.7)
Postgraduate degree	23 (53.5)	70 (32.7)	29 (31.9)
Other	0 (0.0)	5 (2.3)	0 (0.0)
Health insurance			
Government insurance	3 (7.0)	65 (30.5)	19 (20.7)
Employer insurance	30 (69.8)	100 (46.9)	55 (59.8)
No insurance	1 (2.3)	2 (0.9)	2 (2.2)
Other	9 (20.9)	46 (21.6)	16 (17.4)
Employment status			
Full-time	30 (69.8)	93 (43.5)	41 (44.6)
Part-time	3 (7.0)	17 (7.9)	11 (12.0)
Retired	3 (7.0)	18 (8.4)	10 (10.9)
Disabled	4 (9.3)	62 (29.0)	19 (20.7)
Unemployed	0 (0.0)	10 (4.7)	3 (3.3)
Other	3 (7.0)	14 (6.5)	8 (8.7)

[a]All data shown are n (%), unless otherwise stated. Percentages were calculated based on the total number of participants who provided a response and excluded missing information. See Multimedia Appendix 4 for further details on missing results, including the number and proportion of participants who did not provide responses.

[b]MS: multiple sclerosis.

[c]GED: General Educational Development.

XSL•FO
RenderX

Table 3. Baseline disease characteristics of study participants with MS.

Characteristic[a]	Participants with MS (self-referred; N=359)	Participants with MS (clinic referred; N=136)
MS[b] diagnosis		
Relapsing–remitting	300 (83.6)	123 (90.4)
Primary progressive	34 (9.5)	6 (4.4)
Secondary progressive	25 (7.0)	5 (3.7)
Not sure	0 (0.0)	2 (1.5)
Current DMT[c]		
Infusion	116 (32.3)	64 (47.1)
Injection	83 (23.1)	24 (17.6)
Oral	114 (31.8)	40 (29.4)
None	46 (12.8)	6 (4.4)
Missing	0 (0.0)	2 (1.5)
MS family history		
Yes	77 (21.4)	23 (16.9)
No	251 (69.9)	104 (76.5)
Not sure	31 (8.6)	9 (6.6)
Overall physical ability[d]		
Normal	101 (28.1)	52 (38.2)
Gait disability	85 (23.7)	24 (17.6)
Mild disability	104 (29.0)	38 (27.9)
Moderate disability	69 (19.2)	20 (14.7)
Missing	0 (0.0)	2 (1.5)
Duration of disease		
Years since diagnosis, mean (SD)	11.14 (8.86)	14.29 (8.89)
Duration of treatment		
Years since first DMT, mean (SD)	10.09 (7.97)	13.07 (7.92)

[a]All data shown are n (%), unless otherwise stated.

[b]MS: multiple sclerosis.

[c]DMT: disease-modifying therapy.

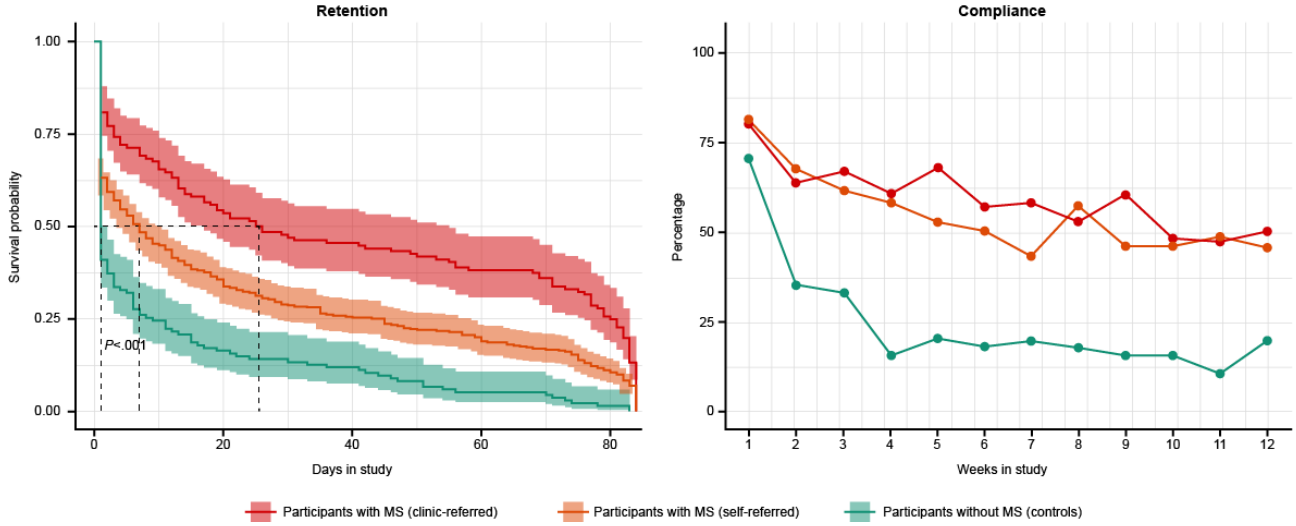
[d]Based on truncated 4-point Patient Determined Disease Steps scale.

Participant Engagement

Median study retention was significantly higher in clinic-referred participants with MS (25.5 days [95% CI 17.0-55.0]) compared with self-referred participants with MS (7.0 days [95% CI 4.0-11.0]) and controls (1.0 day [95% CI 1.0-2.0 days]; $P<.001$; Figure 3). Compliance, defined in this study as the completion of at least one sensor-based active functional test per week, decreased over time in all cohorts; from Week 1 to Week 12, compliance fell from 80.2% to 50.0% for the clinic-referred cohort, from 81.1% at Week 1 to 46.1% for the self-referred cohort, and from 70.9% to 20.0% in the control cohort (Figure 3). Given the lack of ongoing engagement in the control group and the fact that elevateMS primarily targeted participants with MS, the control cohort was not included in subsequent data analyses, and results from self-referred and clinic-referred participants with MS were pooled for analysis.

XSL•FO
RenderX

Figure 3. elevateMS user engagement. Participant retention (median number of days in the study) and compliance (completion of at least one out of four sensor-based active functional tests per week) across the three study cohorts. MS: multiple sclerosis.



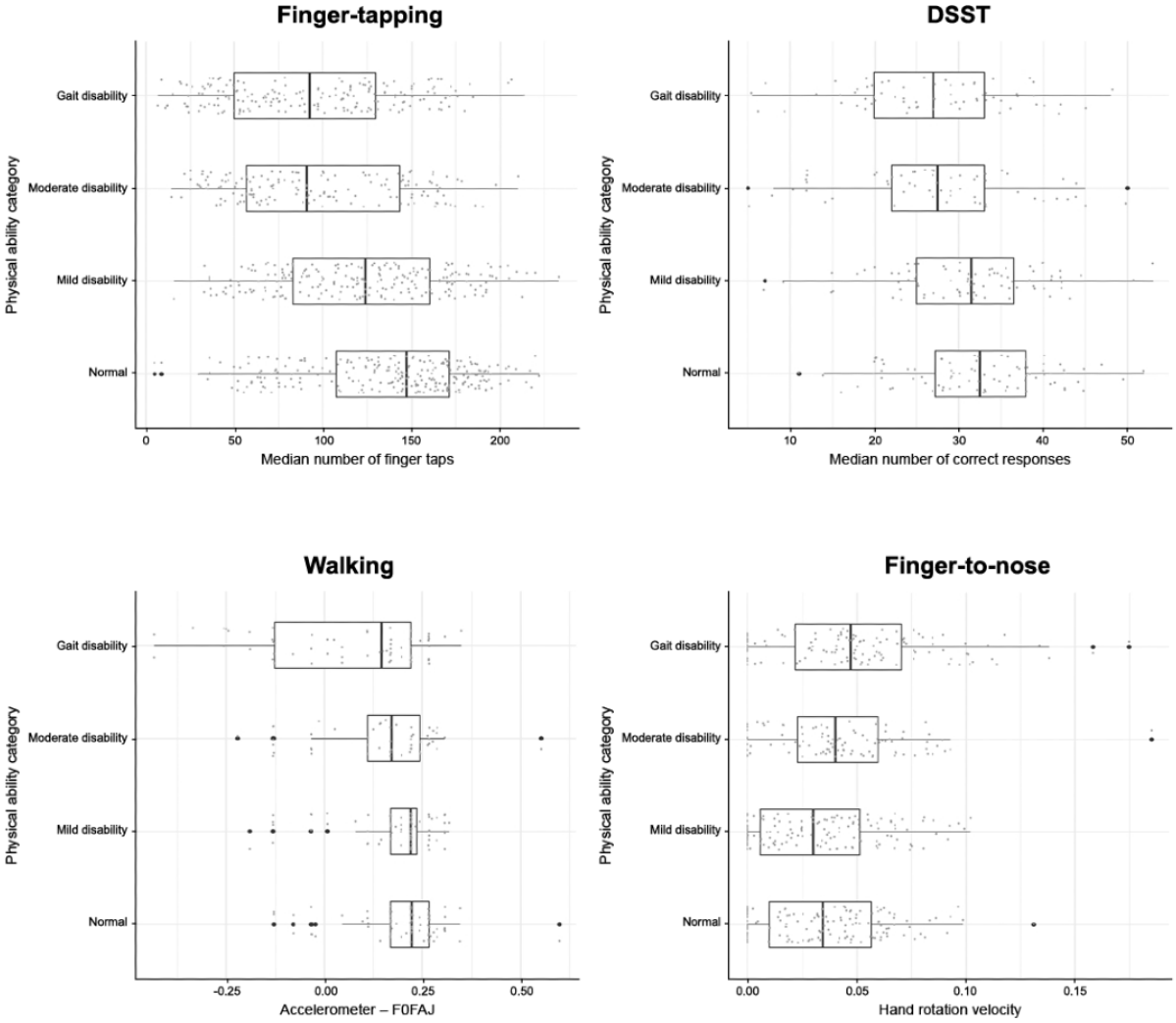
For patients with MS, participation in all tasks (both sensor-based active functional tests and daily check-in surveys) was highest in Weeks 1 and 2, then decreased over time (Multimedia Appendix 5). When using the elevateMS app, participants with MS completed the active functional tests (median 40% of overall individual activity; IQR 30.3) the most, followed by reporting MS symptoms and triggers (median 33.3% of overall individual activity; IQR 33) and completing daily check-in surveys (median 22.3% of overall individual activity; IQR 18.2; Multimedia Appendix 6). The most common self-reported symptoms were fatigue, weakness, memory/attention issues, and difficulty walking, and the most common self-reported triggers were high ambient temperature, emotional stress, and going to bed late, all of which were experienced on at least one occasion by more than half of participants with MS (Multimedia Appendix 7). Data collected from sensor-based active tests were transformed into features and filtered for validity using the mhealthtools pipeline [47] (see Multimedia Appendix 8 for further details).

Relationship Between Baseline Physical Ability and Performance in Active Functional Tests

Higher physical disability at baseline was associated with significantly worse performance (P<.001; top associations listed below) in multiple sensor-based active functional tests. For each active functional test, a subset of the most significantly associated features is presented here (Figure 4 and Multimedia Appendix 9); the full list of per-feature results is available via Synapse, an online data repository managed by Sage Bionetworks [55]. For the finger-tapping assay, a reduced number of finger taps was statistically significantly associated with baseline physical ability category ($\beta_{\text{gait disability vs normal}}$=–43.64, P<.001). Similarly, fewer correct responses in the voice-based DSST task was significantly associated with baseline physical ability ($\beta_{\text{gait disability vs normal}}$=–5.47, P=.005). Worse performance in the walking test, based on a feature derived from the device accelerometer (F0FAJ) [48], was statistically significantly associated with low baseline physical ability ($\beta_{\text{gait disability vs normal}}$=–0.39, P=.001). For the finger-to-nose test, a tremor feature derived from capturing the hand rotation velocity was also found to be significantly associated with baseline physical ability ($\beta_{\text{gait disability vs normal}}$=0.01, P=.01). Balance features were not significantly associated with baseline physical ability (data not shown). Notably, the feature most associated with performance from the finger-tapping assay (median number of finger taps) was also significantly associated with several sociodemographic characteristics in participants with MS, including age group (P<.001), education (P=.001), duration of treatment (P=.004), and duration of disease (P=.009; Multimedia Appendix 9). Additionally, baseline physical ability was significantly associated with duration of treatment (P=.003), duration of disease, employment status, type of health insurance, and age group (all P<.001) in participants with MS (Multimedia Appendix 10).

XSL•FO
RenderX

Figure 4. Association between baseline characteristics and functional test performance in participants with MS. DSST (Digit Symbol Substitution Test): decrease in number of correct DSST responses with increased baseline physical disability; F0FAJ: frequency at which the maximum peak of the Lomb-Scargle periodogram occurred for the average acceleration series, with frequencies limited to 0.2-5 Hz; Finger-tapping: decrease in median number of finger taps with increased baseline physical disability; Finger-to-nose: increase in hand rotation velocity tremor feature with increased baseline physical disability; MS: multiple sclerosis; PDDS: Patient-Determined Disease Steps; Walking: decrease in F0FAJ accelerometer results with increased baseline physical disability.

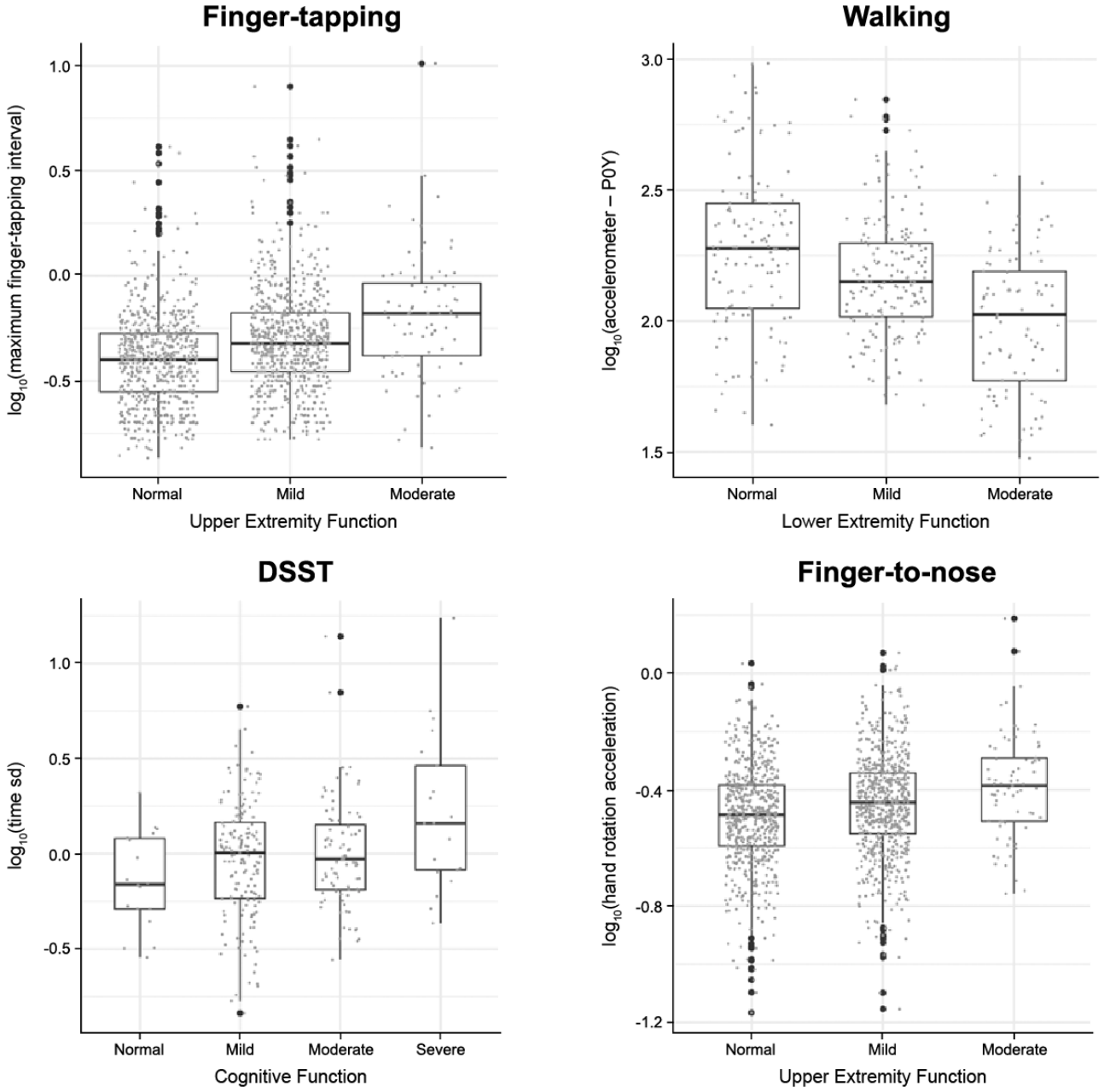


Relationship Between Neuro-QoL Domains and Performance in Active Functional Tests

Performance in sensor-based active functional tests was also significantly associated (P<.001; top associations listed below) with the 3 short-form Neuro-QoL surveys (Cognitive Function, Upper Extremity Function, and Lower Extremity Function) that were administered to participants at several points during the course of the study. For each test, a subset of the top most significantly associated features are presented here (Figure 5 and Multimedia Appendix 11), with the full list of per-feature comparisons available online via Synapse data repository [56]. Performance in the finger-tapping test was significantly associated with Upper Extremity Function domain scores ($\beta_{\text{moderate vs normal}}$=0.40 seconds, P<.001) of the Neuro-QoL. For the walking test, a feature derived from the device accelerometer (P0Y) [48] was found to be significantly associated with the Neuro-QoL Lower Extremity Function domain scores ($\beta_{\text{moderate vs normal}}$=–99.18 seconds, P=.02). Increased severity in the Neuro-QoL Cognitive and Lower Extremity Function domains was also significantly associated with the variation in time taken to complete the voice-based DSST task (Cognitive Function: $\beta_{\text{severe vs normal}}$=1.60 seconds, P=.03; Lower Extremity Function: $\beta_{\text{severe vs normal}}$=10.31 seconds, P<.001). Finally, for the finger-to-nose test, a feature derived from capturing the hand rotation acceleration using the device gyroscope was significantly associated with the Neuro-QoL Upper Extremity Function domain ($\beta_{\text{moderate vs normal}}$=.11, P=.003). Additional sensitivity analyses were performed to assess the impact of Neuro-QoL scores that mapped to the Severe Neuro-QoL category. While excluding Neuro-QoL scores in severe category did not change the main findings for the finger-tapping, walk and balance, and finger-to-nose tests, the analysis showed that the significant association of Cognitive Function and Lower Extremity Function with DSST was mainly driven by those participants reporting severe Neuro-QoL outcomes (Multimedia Appendix 11).

XSL•FO
RenderX

Figure 5. Association between Neuro-QoL™ domains and functional test performance in participants with MS. Neuro-QoL categories comprising <5% of total participants were not plotted. DSST (Digit Symbol Substitution Test): increase in DSST response time with increased severity in Neuro-QoL Cognition domain. Finger-tapping: increase in maximum finger tapping interval with increased severity in the Neuro-QoL Upper Extremity Function domain; Finger-to-nose: increase in hand rotation acceleration tremor feature with increased severity in Neuro-QoL Upper Extremity Function domain; MS: multiple sclerosis; Neuro-QoL: Quality of Life in Neurological Disorders; P0Y: maximum power in the inspected frequency interval of the Lomb–Scargle periodogram for the Y acceleration series (0.2–5 Hz); Walking: decrease in accelerometer-derived feature (P0Y) with increased severity in Neuro-QoL Lower Extremity Function domain.

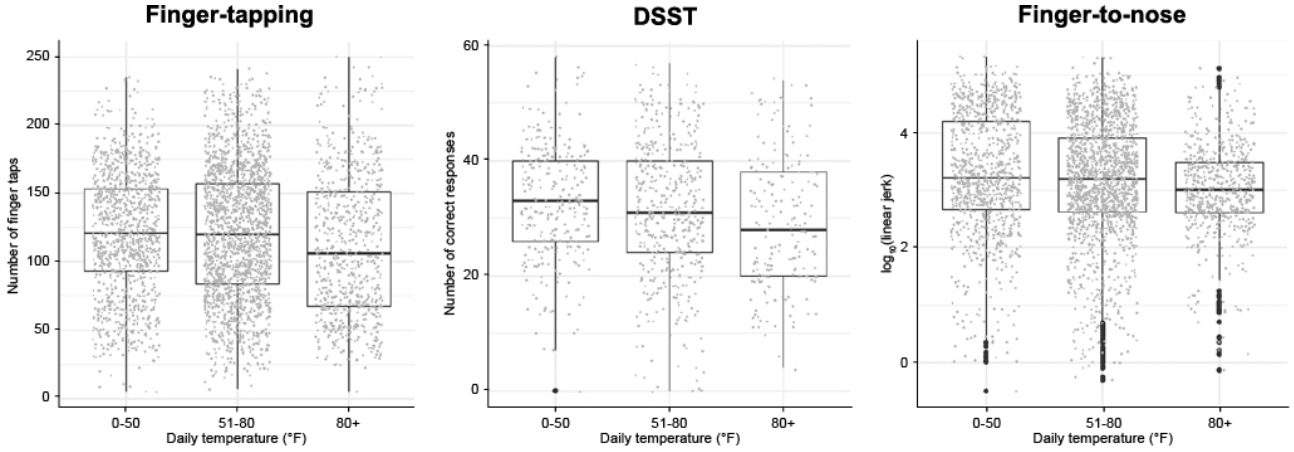


Impact of Local Weather Conditions on Performance in Active Functional Tests and Daily Check-In PROs

Local weather data were captured each time a participant completed a functional test and were found to be significantly associated (P<.001; top associations listed below) with respective test performance (Figure 6 and Multimedia Appendix 12) [57]. The local temperature elevations at the time of test completion negatively impacted participant's performance to the greatest extent, with a significant difference in performance in the finger-tapping test observed (β=–.14, P<.001). For example, with a 30°F increase in temperature (ie, from 50°F to 80°F), the participant's performance dropped by an average of 4.3 finger-taps (β × 30°F). Similarly, performance in the voice-based DSST and finger-to-nose tests was significantly associated with increased temperature (β=–.06, P=.009 and β=–53.88, P<.001, respectively). In contrast to active functional test features, PROs recorded from daily check-in surveys were only moderately associated with local weather conditions (Multimedia Appendix 13) [58]. Sensitivity analysis further showed that the association between PROs and weather features was mainly driven by a small proportion of participants reporting extreme outcomes on the Likert scale (Multimedia Appendix 14) [59].

XSL•FO
RenderX

Figure 6. Association between daily temperature and functional test performance in participants with MS. Finger-tapping: decrease in number of finger taps with increased temperature. DSST (Digit Symbol Substitution Test): decrease in number of correct DSST responses with increased temperature. Finger-to-nose: decrease in accelerometer-derived linear jerk tremor feature derived with increased temperature. MS: multiple sclerosis.



Discussion

Principal Results

The results from this observational, remote data collection study demonstrate the feasibility and utility of a decentralized method to gather real-world data about participant's real-time life experience of MS through a digital health app, elevateMS. Compared with previous digital health studies in MS, elevateMS enrolled one of the largest remote cohorts with a self-reported or neurologist-confirmed diagnosis of MS from across the United States [23,31,32,34,35]. The sociodemographic characteristics of the enrolled patient cohort were broadly similar to the wider population of patients with MS [60,61]. Participation in elevateMS was also generally consistent with that of digital health studies in other disease areas [62]. However, to our knowledge, elevateMS is one of the first remote, digital studies to show the significant impact of clinical referral on overall engagement in a remote population [22]; clinic-referred participants with MS remained active in our study almost 3 times longer than self-referred participants with MS. Importantly, the digitally measured functional activity correlated with clinical outcomes and QoL.

Through the longitudinal collection of PROs, active functional test results, and local weather data, the elevateMS study demonstrates the importance of frequent, real-world assessments of MS disease manifestations outside of episodic clinical evaluations. Tracking of self-reported data identified the most common disease symptoms and triggers in patients with MS, as well as significant associations (P<.001) between performance in active functional tests and disease severity, measured by both PDDS and Neuro-QoL subdomain scores. Although PROs failed to capture the impact of local weather conditions, participants' performance in various active functional test features was found to be significantly associated (P<.001) with local temperature patterns, with the worst performance observed at temperatures over 80°F. This supports the well-established link between increased temperature and MS [7-9] and demonstrates the sensitivity of these sensor-based tests in patient monitoring. Together, these results show the potential utility of active functional tests in capturing measurements of MS-related motor activity and assessing the impact of local environmental factors on disease symptoms and severity in a real-world setting.

A major strength of the elevateMS study is that it collected self-reported and self-administered measurements of MS health data from patients remotely. PROs are increasingly recognized as valid and meaningful clinical measures for disease monitoring and patient care across a range of therapeutic areas [63-72]. Within the MS field, improved self-assessment and QoL reporting, particularly using electronic/digital tools such as elevateMS in a remote, unsupervised setting, has the potential to benefit both patients and clinicians by enhancing communication and understanding of individual patient needs, thus improving the overall patient experience and informing therapy selection [17,24,73,74]. This is particularly important given that a low level of concordance has been observed between patients and neurologists in recognizing MS relapses, assessing health status, and identifying QoL parameters that are of greatest concern to the individual patient [17,24,75].

Another strength of this study is that both PROs and active measurements were collected contemporaneously in a real-time and real-world setting, with a frequency far exceeding that obtainable in routine MS clinical care. This differentiates elevateMS from the episodic and retrospective periods of data collection that are characteristic of traditional, clinic-based care and often subject to recall bias. By capturing a more comprehensive body of data, such as the frequency of triggers, variability of symptoms, and effect of environmental conditions, elevateMS could enable the interaction between daily life stressors and MS severity to be better evaluated. Furthermore, by leveraging frequent and low-burden assessments, elevateMS and other digital health tools may facilitate regular patient monitoring between clinic visits. This could complement the existing clinical practice, by helping patients to record their symptoms, relapses, and medication usage more accurately and have an active role in their disease management; in turn, this has the potential to provide a more thorough assessment of disease, improve communication with health care providers, and ultimately support clinicians in developing personalized treatment plans [19,27,29,32,34,35]. More broadly and universally applied, this technology may also provide novel insights into the course of chronic and progressive conditions,

XSL•FO
RenderX

as demonstrated by previous digital health studies in MS, asthma, and Parkinson disease [25-27,34]. Finally, as shown by elevateMS, digital health tools can be utilized to gather data from large, remote populations and could, therefore, offer unique opportunities to track and evaluate drug efficacy in a continuous, real-world setting through decentralized clinical trials [34].

Limitations

Some potential limitations need to be considered when interpreting the results from this study. Given that participation required a specific smartphone, the study population may be subject to selection bias [62]; however, the sociodemographic characteristics available from the enrolled patient cohort are broadly representative of the general population with MS [60,61]. As this study largely focused on patients with MS, participation by the control cohort was low, and these data were not included in our analyses. Approximately two-thirds of the total MS cohort self-reported their disease status, and these data may be vulnerable to inaccuracies. However, prior studies have shown that self-reported MS diagnoses in the Pacific Northwest MS Registry were subsequently confirmed by neurology health care providers in more than 98% of cases [76]. In a second study, levels of disability were accurately self-reported by patients and comparable to neurologist ratings in 73% of cases [77]. Within the elevateMS study, data from both self-referred and clinic-referred patients with MS were pooled for analysis, thus precluding between-group comparisons. Future studies could evaluate these 2 patient cohorts separately to enable results to be analyzed in relation to clinical referral status.

We also observed a large degree of missing baseline demographic data within this pilot study, which may reflect the fact that this information was not collected until Day 2 of participation, in an attempt to reduce initial onboarding participant burden. This missing data could impact statistical inference, in particular the robustness of the LME models, as the random effect of many sociodemographic characteristics could not be accounted for. Given that a significant proportion of participants in remote app-based studies drop out during the first week, with the majority leaving on Day 1-2 [62], future studies should prioritize immediate collection of baseline demographic data and make this a compulsory step upon enrollment. This would ensure that demographic information can be fully evaluated in relation to app study results.

Given the limited, 12-week study participation period, we were not able to assess longer-term variations in MS symptoms or severity in this pilot study. Furthermore, this short window of observation meant that disease severity could be assessed in relation only to daily, and not seasonal, fluctuations. Going forward, longer study durations would provide greater opportunities for longitudinal disease monitoring and assessment of the impact of external lifestyle and environmental factors. This is particularly important in chronic diseases such as MS, where there is a risk of progression and unpredictable variability of symptoms or relapses over time [5].

The sample size and engagement of the elevateMS patient cohort were generally consistent with or higher than those in previously published remote digital studies undertaken across different therapy areas [28-30,62]. However, the overall participation in elevateMS was still low compared with digital studies in MS that included an in-person clinic visit [31,32]. The greatest proportion of activity within the elevateMS study app was spent on sensor-based active functional tests, followed by symptom and trigger surveys, and check-in surveys. Participation in specific app activities may have been affected by the time, effort, or frequency associated with that activity; for example, each of the 4 sensor-based active tests were administered on a weekly basis and could be completed in less than 1 minute, whereas each survey was administered daily, potentially making them more burdensome despite their short length. Future studies could further align the study protocol with participant needs and busy schedules employing user-centered co-design techniques. Review of compliance patterns also demonstrated that patients referred from and currently under the care of a provider were the most adherent participants, and future studies should focus on building studies around this cohort. Doing so will not only increase the likelihood of patient retention, but will also provide physician verification of the validity of responses. Additional studies are necessary to develop strategies, including the use of incentives to increase involvement and adherence of people who self-identify and self-refer, to expand the number of patient participants with greater assurance of long-term adherence.

Within the elevateMS study, clinic-referred participants demonstrated greater retention and compliance than the self-referred cohort, which suggests that participants are more likely to engage if they are encouraged by a clinician or aware of how their data may be used to inform and personalize their care. Based on these results, future digital health studies should be incorporated, as exploratory outcomes initially, within clinical trials to assess the applicability and utility of digital monitoring. Finally, it is known that comorbidities, as well as concurrent medication use, have a significant impact on MS patients [11,12,78,79]; details regarding comorbidities were not collected in this pilot study, but should be included in future long-term assessments.

Comparison With Prior Work

In line with previous digital health feasibility studies, our results demonstrate that smartphone technology can be used to collect both sensor-based active measurements and passive data related to disease symptoms and severity in patients with MS [23,31-35]. In contrast to previous studies, elevateMS collected data from a large, geographically diverse, remote, and unsupervised population, independent of scheduled clinic visits. Although a significant number of individuals did not participate beyond enrollment on the first day of the elevateMS study, our user engagement data are consistent with previous digital health studies that recruited broadly from the general population, with no scheduled in-clinic touchpoints or incentives associated with the app usage [62]. As with previous digital health studies [26,28,34], elevateMS relied on arbitrary measures of retention and compliance. In order to fully assess participant engagement and enable comparisons between different studies, these parameters need to be defined in more specific terms. For example, the BEST (Biomarkers, EndpointS, and other Tools) Resource, created by the US Food and Drug Administration and the National Institutes of Health [80], could be expanded to include clear and unambiguous definitions of retention and

XSL•FO
RenderX

compliance in digital health studies, creating standardized measures that could be utilized across the field.

Clinical validation of elevateMS data was beyond the scope of this study; however, it is reassuring that the symptoms and triggers most commonly reported in the app, such as fatigue, weakness, temperature and stress, are already well-documented in traditional studies of MS [7-10,81]. This is further reinforced by our results showing worse performance in active functional tests, such as finger-tapping and DSST responses, in increased ambient temperatures, reflecting the well-known heat sensitivity experienced by patients with MS [7-9]. Furthermore, by comparing self-reported PDDS and Neuro-QoL results, we have shown that it is possible to use smartphone-based motor measurements to assess both disease severity and QoL, providing internal validation of elevateMS results.

Conclusions

In contrast to current, episodic disease monitoring practices, this study demonstrates the value and utility of frequently assessing the real-world, live patient experience of MS using a digital health app. By providing a more comprehensive and representative assessment of patients outside of the clinic, elevateMS and other disease-tracking apps have the potential to enhance the understanding of MS, facilitate patient–clinician communication, and support personalization of disease management plans.

Acknowledgments

The authors thank all MS patient advisors for their time and feedback during the design of the elevateMS app, as well as the clinicians and staff of the clinical sites that helped recruit participants to the elevateMS study. The authors are grateful to individuals who downloaded the elevateMS study app and voluntarily contributed data. The authors also thank the Sage Bionetworks engineering, design, and governance teams, in particular Vanessa Barone, Christine Suver, Woody MacDuffy, and Mike Kellen, for designing and developing the elevateMS app and providing support to the research participants. We also thank Roshani Shah of Novartis Pharmaceuticals Corporation for assisting with manuscript development. The elevateMS study was run by Sage Bionetworks, a nonprofit institution based in Seattle, and sponsored by Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA. The authors received editorial support for this manuscript from Lauren Chessum of Fishawack Communications Ltd., funded by Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA.

Authors' Contributions

LO, DG, AV, and CD were involved in the conception, study design, and execution of the investigation. LM acted as investigator and reviewed the results and manuscript. AP oversaw the data curation and featurization and conducted the analysis. LO wrote the initial analytical plan and oversaw the study. MT was responsible for feature extraction from raw sensor-level data as well as the analysis and curation of study data for public release. SC provided expert knowledge to contextualize results from a clinically meaningful perspective and reviewed the results and manuscript. AP and LO co-wrote the first draft of the manuscript and contributed substantially to the final draft.

Conflicts of Interest

SC received consulting fees for research/speaking honoraria/advisory boards from AbbVie, Alexion, Atara Biotherapeutics, Biogen, EMD Serono, Novartis, Roche/Genentech, Sanofi Genzyme, MedDay, and Pear Therapeutics. AP, LM, MT, and LO are employees of Sage Bionetworks. AV and CD are employees of Novartis Pharmaceuticals Corporation; DG is employed by Alcon.

Multimedia Appendix 1
Overview of user-centered design process used to create the elevateMS study app.
[DOCX File , 60 KB - mhealth_v8i10e22108_app1.docx]

Multimedia Appendix 2
Example elevateMS survey questions.
[DOCX File , 39 KB - mhealth_v8i10e22108_app2.docx]

Multimedia Appendix 3
Example features from active functional performance tests.
[DOCX File , 27 KB - mhealth_v8i10e22108_app3.docx]

Multimedia Appendix 4
Missing responses in baseline sociodemographic characteristics of study participants.
[DOCX File , 27 KB - mhealth_v8i10e22108_app4.docx]

Multimedia Appendix 5

XSL•FO
RenderX

Summary of activity-specific compliance across the 12-week study duration.
[DOCX File , 39 KB - mhealth_v8i10e22108_app5.docx]

Multimedia Appendix 6
Distribution of overall activity type per user for participants with multiple sclerosis.
[PNG File , 36 KB - mhealth_v8i10e22108_app6.png]

Multimedia Appendix 7
Participant-reported symptoms and triggers.
[DOCX File , 35 KB - mhealth_v8i10e22108_app7.docx]

Multimedia Appendix 8
Example assessment of participant performance in the walk and balance sensor-based active functional test.
[DOCX File , 146 KB - mhealth_v8i10e22108_app8.docx]

Multimedia Appendix 9
Association between baseline characteristics and active functional test performance in participants with MS (top features for each test).
[DOCX File , 45 KB - mhealth_v8i10e22108_app9.docx]

Multimedia Appendix 10
Heatmap showing association between all recorded baseline characteristics in participants with multiple sclerosis.
[DOCX File , 61 KB - mhealth_v8i10e22108_app10.docx]

Multimedia Appendix 11
Association between Neuro-QoL™ domains and functional test performance in participants with MS (top features for each test).
[DOCX File , 48 KB - mhealth_v8i10e22108_app11.docx]

Multimedia Appendix 12
Association between local weather conditions and functional test performance in participants with MS.
[DOCX File , 39 KB - mhealth_v8i10e22108_app12.docx]

Multimedia Appendix 13
Association between local weather conditions and PROs in participants with MS.
[DOCX File , 61 KB - mhealth_v8i10e22108_app13.docx]

Multimedia Appendix 14
Association between local weather conditions and PROs in participants with MS (sensitivity analysis).
[DOCX File , 58 KB - mhealth_v8i10e22108_app14.docx]

References

1. GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Neurol 2017 Nov;16(11):877-897 [FREE Full text] [doi: 10.1016/S1474-4422(17)30299-5] [Medline: 28931491]
2. Zwibel HL, Smrtka J. Improving quality of life in multiple sclerosis: an unmet need. Am J Manag Care 2011 May;17 Suppl 5:S139-S145 [FREE Full text] [Medline: 21761952]
3. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983 Nov;33(11):1444-1452. [Medline: 6685237]
4. Campbell JD, Ghushchyan V, Brett McQueen R, Cahoon-Metzger S, Livingston T, Vollmer T, et al. Burden of multiple sclerosis on direct, indirect costs and quality of life: National US estimates. Mult Scler Relat Disord 2014 Mar;3(2):227-236. [doi: 10.1016/j.msard.2013.09.004] [Medline: 25878010]
5. Reich DS, Lucchinetti CF, Calabresi PA. Multiple Sclerosis. N Engl J Med 2018 Jan 11;378(2):169-180 [FREE Full text] [doi: 10.1056/NEJMra1401483] [Medline: 29320652]
6. Munger KL, Zhang SM, O'Reilly E, Hernán MA, Olek MJ, Willett WC, et al. Vitamin D intake and incidence of multiple sclerosis. Neurology 2004 Jan 13;62(1):60-65. [doi: 10.1212/01.wnl.0000101723.79681.38] [Medline: 14718698]

XSL•FO
RenderX

7. Hämäläinen P, Ikonen A, Romberg A, Helenius H, Ruutiainen J. The effects of heat stress on cognition in persons with multiple sclerosis. Mult Scler 2012 Apr;18(4):489-497. [doi: 10.1177/1352458511422926] [Medline: 21914688]
8. Romberg A, Ikonen A, Ruutiainen J, Virtanen A, Hämäläinen P. The effects of heat stress on physical functioning in persons with multiple sclerosis. J Neurol Sci 2012 Aug 15;319(1-2):42-46. [doi: 10.1016/j.jns.2012.05.024] [Medline: 22647586]
9. Davis SL, Wilson TE, White AT, Frohman EM. Thermoregulation in multiple sclerosis. J Appl Physiol (1985) 2010 Nov;109(5):1531-1537. [doi: 10.1152/japplphysiol.00460.2010] [Medline: 20671034]
10. Mohr DC, Hart SL, Julian L, Cox D, Pelletier D. Association between stressful life events and exacerbation in multiple sclerosis: a meta-analysis. BMJ 2004 Mar 27;328(7442):731 [FREE Full text] [doi: 10.1136/bmj.38041.724421.55] [Medline: 15033880]
11. Marrie RA, Cohen J, Stuve O, Trojano M, Sørensen PS, Reingold S, et al. A systematic review of the incidence and prevalence of comorbidity in multiple sclerosis: overview. Mult Scler 2015 Mar;21(3):263-281 [FREE Full text] [doi: 10.1177/1352458514564491] [Medline: 25623244]
12. Capkun G, Dahlke F, Lahoz R, Nordstrom B, Tilson HH, Cutter G, et al. Mortality and comorbidities in patients with multiple sclerosis compared with a population without multiple sclerosis: An observational study using the US Department of Defense administrative claims database. Mult Scler Relat Disord 2015 Nov;4(6):546-554 [FREE Full text] [doi: 10.1016/j.msard.2015.08.005] [Medline: 26590661]
13. Lerdal A, Celius EG, Krupp L, Dahl AA. A prospective study of patterns of fatigue in multiple sclerosis. Eur J Neurol 2007 Dec;14(12):1338-1343. [doi: 10.1111/j.1468-1331.2007.01974.x] [Medline: 17903208]
14. Hassan E. Recall Bias can be a Threat to Retrospective and Prospective Research Designs. IJE 2006 Jan;3(2):1-7. [doi: 10.5580/2732]
15. McPhail S, Haines T. Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients. Health Qual Life Outcomes 2010 Jul 10;8:65 [FREE Full text] [doi: 10.1186/1477-7525-8-65] [Medline: 20618978]
16. Schmier JK, Halpern MT. Patient recall and recall bias of health state and health status. Expert Rev Pharmacoecon Outcomes Res 2004 Apr;4(2):159-163. [doi: 10.1586/14737167.4.2.159] [Medline: 19807519]
17. D'Amico E, Haase R, Ziemssen T. Review: Patient-reported outcomes in multiple sclerosis care. Mult Scler Relat Disord 2019 Aug;33:61-66. [doi: 10.1016/j.msard.2019.05.019] [Medline: 31154262]
18. Solomon AJ, Weinshenker BG. Misdiagnosis of multiple sclerosis: frequency, causes, effects, and prevention. Curr Neurol Neurosci Rep 2013 Dec;13(12):403. [doi: 10.1007/s11910-013-0403-y] [Medline: 24142849]
19. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. Acad Emerg Med 2015 Jun;22(6):754-756 [FREE Full text] [doi: 10.1111/acem.12692] [Medline: 25997375]
20. Jongen PJ, Ter Veen G, Lemmens W, Donders R, van Noort E, Zeinstra E. The Interactive Web-Based Program MSmonitor for Self-Management and Multidisciplinary Care in Persons With Multiple Sclerosis: Quasi-Experimental Study of Short-Term Effects on Patient Empowerment. J Med Internet Res 2020 Mar 09;22(3):e14297 [FREE Full text] [doi: 10.2196/14297] [Medline: 32149713]
21. Ziemssen T, Piani-Meier D, Bennett B, Johnson C, Tinsley K, Trigg A, et al. A Physician-Completed Digital Tool for Evaluating Disease Progression (Multiple Sclerosis Progression Discussion Tool): Validation Study. J Med Internet Res 2020 Feb 12;22(2):e16932 [FREE Full text] [doi: 10.2196/16932] [Medline: 32049062]
22. Giunti G, Kool J, Rivera RO, Dorronzoro ZE. Exploring the Specific Needs of Persons with Multiple Sclerosis for mHealth Solutions for Physical Activity: Mixed-Methods Study. JMIR Mhealth Uhealth 2018 Feb 09;6(2):e37 [FREE Full text] [doi: 10.2196/mhealth.8996] [Medline: 29426814]
23. Creagh AP, Simillion C, Scotland A, Lipsmeier F, Bernasconi C, Belachew S, et al. Smartphone-based remote assessment of upper extremity function for multiple sclerosis using the Draw a Shape Test. Physiol Meas 2020 Jun 19;41(5):054002. [doi: 10.1088/1361-6579/ab8771] [Medline: 32259798]
24. Ysrraelit MC, Fiol MP, Gaitán MI, Correale J. Quality of Life Assessment in Multiple Sclerosis: Different Perception between Patients and Neurologists. Front Neurol 2017;8:729 [FREE Full text] [doi: 10.3389/fneur.2017.00729] [Medline: 29375468]
25. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar 03;3:160011 [FREE Full text] [doi: 10.1038/sdata.2016.11] [Medline: 26938265]
26. Chan YY, Wang P, Rogers L, Tignor N, Zweig M, Hershman SG, et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. Nat Biotechnol 2017 Mar 13 [FREE Full text] [doi: 10.1038/nbt.3826] [Medline: 28288104]
27. 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]
28. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using Mobile Apps to Assess and Treat Depression in Hispanic and Latino Populations: Fully Remote Randomized Clinical Trial. J Med Internet Res 2018 Aug 09;20(8):e10130 [FREE Full text] [doi: 10.2196/10130] [Medline: 30093372]

XSL•FO
RenderX

29. Crouthamel M, Quattrocchi E, Watts S, Wang S, Berry P, Garcia-Gancedo L, et al. Using a ResearchKit Smartphone App to Collect Rheumatoid Arthritis Symptoms From Real-World Participants: Feasibility Study. JMIR Mhealth Uhealth 2018 Sep 13;6(9):e177. [doi: 10.2196/mhealth.9656] [Medline: 30213779]
30. McConnell MV, Shcherbina A, Pavlovic A, Homburger JR, Goldfeder RL, Waggot D, et al. Feasibility of Obtaining Measures of Lifestyle From a Smartphone App: The MyHeart Counts Cardiovascular Health Study. JAMA Cardiol 2017 Jan 01;2(1):67-76. [doi: 10.1001/jamacardio.2016.4395] [Medline: 27973671]
31. Bove R, White CC, Giovannoni G, Glanz B, Golubchikov V, Hujol J, et al. Evaluating more naturalistic outcome measures: A 1-year smartphone study in multiple sclerosis. Neurol Neuroimmunol Neuroinflamm 2015 Dec;2(6):e162 [FREE Full text] [doi: 10.1212/NXI.0000000000000162] [Medline: 26516627]
32. D'hooghe M, Van GG, Kos D, Bouquiaux O, Cambron M, Decoo D, et al. Improving fatigue in multiple sclerosis by smartphone-supported energy management: The MS TeleCoach feasibility study. Mult Scler Relat Disord 2018 May;22:90-96 [FREE Full text] [doi: 10.1016/j.msard.2018.03.020] [Medline: 29649789]
33. Genentech Inc. Floodlight Open. 2020. URL: https://floodlightopen.com/en-US/ [accessed 2020-10-08]
34. Midaglia L, Mulero P, Montalban X, Graves J, Hauser SL, Julian L, et al. Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study. J Med Internet Res 2019 Aug 30;21(8):e14863 [FREE Full text] [doi: 10.2196/14863] [Medline: 31471961]
35. Limmroth V, Hechenbichler K, Müller C, Schürks M. Assessment of Medication Adherence Using a Medical App Among Patients With Multiple Sclerosis Treated With Interferon Beta-1b: Pilot Digital Observational Study (PROmyBETAapp). J Med Internet Res 2019 Jul 29;21(7):e14373 [FREE Full text] [doi: 10.2196/14373] [Medline: 31359863]
36. elevateMS. Sage Bionetworks, Novartis Pharmaceuticals Corporation. URL: http://www.elevatems.org/ [accessed 2020-05-01]
37. Doerr M, Maguire Truong A, Bot BM, Wilbanks J, Suver C, Mangravite LM. Formative Evaluation of Participant Experience With Mobile eConsent in the App-Mediated Parkinson mPower Study: A Mixed Methods Study. JMIR Mhealth Uhealth 2017 Feb 16;5(2):e14 [FREE Full text] [doi: 10.2196/mhealth.6521] [Medline: 28209557]
38. Grayson S, Suver C, Wilbanks J, Doerr M. Open Data Sharing in the 21st Century: Sage Bionetworks' Qualified Research Program and Its Application in mHealth Data Release. Social Science Research Network. 2020. URL: https://ssrn.com/abstract=3502410 [accessed 2020-10-08]
39. Hohol MJ, Orav EJ, Weiner HL. Disease steps in multiple sclerosis: a simple approach to evaluate disease progression. Neurology 1995 Feb;45(2):251-255. [Medline: 7854521]
40. Cella D, Lai J, Nowinski CJ, Victorson D, Peterman A, Miller D, et al. Neuro-QOL: brief measures of health-related quality of life for clinical research in neurology. Neurology 2012 Jun 05;78(23):1860-1867 [FREE Full text] [doi: 10.1212/WNL.0b013e318258f744] [Medline: 22573626]
41. Cella D, Nowinski C, Peterman A, Victorson D, Miller D, Lai J, et al. The neurology quality-of-life measurement initiative. Arch Phys Med Rehabil 2011 Oct;92(10 Suppl):S28-S36 [FREE Full text] [doi: 10.1016/j.apmr.2011.01.025] [Medline: 21958920]
42. User Manual for the Quality of Life in Neurological Disorders (Neuro-QoL) Measures, Version 2. Bethesda, MD: National Institute of Neurological Disorders and Stroke (NINDS); 2015. URL: http://www.healthmeasures.net/images/neuro_qol/Neuro-QOL_User_Manual_v2_24Mar2015.pdf [accessed 2020-10-09]
43. Cook KF, Victorson DE, Cella D, Schalet BD, Miller D. Creating meaningful cut-scores for Neuro-QOL measures of fatigue, physical functioning, and sleep disturbance using standard setting with patients and providers. Qual Life Res 2015 Mar;24(3):575-589. [doi: 10.1007/s11136-014-0790-9] [Medline: 25148759]
44. Neuro-QoL™ Score Cut Points. 2020. URL: http://www.healthmeasures.net/score-and-interpret/interpret-scores/neuro-qol/neuro-qol-score-cut-points [accessed 2020-10-09]
45. Miller DM, Bethoux F, Victorson D, Nowinski CJ, Buono S, Lai J, et al. Validating Neuro-QoL short forms and targeted scales with people who have multiple sclerosis. Mult Scler 2016 May;22(6):830-841 [FREE Full text] [doi: 10.1177/1352458515599450] [Medline: 26238464]
46. Validated neurocognitive testing: Digit Symbol Substitution. Iowa City, IA: BrainBaseline; 2020. URL: https://www.brainbaseline.com/product/cognitive-testing#symbolsub [accessed 2020-05-01]
47. Snyder P, Tummalacherla M, Perumal T, Omberg L. mhealthtools: A Modular R Package for Extracting Features from Mobile and Wearable Sensor Data. J Open Source Software 2020;5(47):2106 [FREE Full text] [doi: 10.21105/joss.02106]
48. ElevateMS Feature Definitions. Seattle, WA: Sage Bionetworks; 2020. URL: https://github.com/Sage-Bionetworks/elevateMS_data_release/blob/master/FeatureDefinitions.md [accessed 2020-10-09]
49. Bridge Platform. Seattle, WA: Sage Bionetworks; 2020. URL: https://sagebionetworks.org/bridge-platform/ [accessed 2020-10-09]
50. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using. J Stat Soft 2015;67(1). [doi: 10.18637/jss.v067.i01]
51. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. J Stat Soft 2017;82(13). [doi: 10.18637/jss.v082.i13]

XSL•FO
RenderX

52. The R Project for Statistical Computing. Vienna, Austria: The R Foundation; 2018. URL: https://www.R-project.org/ [accessed 2020-10-09]
53. Supplementary Data ElevateMS Manuscript. 2020. URL: http://doi.org/10.7303/syn22005477 [accessed 2020-10-10]
54. Synapse. URL: http://www.synapse.org/elevateMS [accessed 2020-10-10]
55. Synapse - Baseline Characteristics vs Active Test Performance (syn22006660.2). Seattle, WA: Sage Bionetworks; 2020. URL: http://doi.org/10.7303/syn22006660.2 [accessed 2020-05-01]
56. Synapse - Neuro-QoL vs Active Test Performance (syn22008560.3). Seattle, WA: Sage Bionetworks; 2020. URL: http://doi.org/10.7303/syn22008560.3 [accessed 2020-05-01]
57. Synapse - Weather vs Active Test Performance (syn22033185.1). Seattle, WA: Sage Bionetworks; 2020. URL: https://doi.org/10.7303/syn22033185.1 [accessed 2020-05-01]
58. Synapse - Weather vs Daily Check-ins (syn22033192.2). Seattle, WA: Sage Bionetworks; 2020. URL: https://doi.org/10.7303/syn22033192.2 [accessed 2020-05-01]
59. Synapse - Weather vs Daily Check-ins + Sensitivity Analysis (syn22033187.2). Seattle, WA: Sage Bionetworks; 2020. URL: https://doi.org/10.7303/syn22033187.2 [accessed 2020-05-01]
60. Dilokthornsakul P, Valuck RJ, Nair KV, Corboy JR, Allen RR, Campbell JD. Multiple sclerosis prevalence in the United States commercially insured population. Neurology 2016 Mar 15;86(11):1014-1021 [FREE Full text] [doi: 10.1212/WNL.0000000000002469] [Medline: 26888980]
61. Harbo HF, Gold R, Tintoré M. Sex and gender issues in multiple sclerosis. Ther Adv Neurol Disord 2013 Jul;6(4):237-248 [FREE Full text] [doi: 10.1177/1756285613488434] [Medline: 23858327]
62. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
63. Hirpara DH, Gupta V, Brown L, Kidane B. Patient-reported outcomes in lung and esophageal cancer. J Thorac Dis 2019 Mar;11(Suppl 4):S509-S514 [FREE Full text] [doi: 10.21037/jtd.2019.01.02] [Medline: 31032069]
64. Kalluri M, Luppi F, Ferrara G. What Patients With Idiopathic Pulmonary Fibrosis and Caregivers Want: Filling the Gaps With Patient Reported Outcomes and Experience Measures. Am J Med 2020 Mar;133(3):281-289 [FREE Full text] [doi: 10.1016/j.amjmed.2019.08.032] [Medline: 31520625]
65. Nair D, Wilson FP. Patient-Reported Outcome Measures for Adults With Kidney Disease: Current Measures, Ongoing Initiatives, and Future Opportunities for Incorporation Into Patient-Centered Kidney Care. Am J Kidney Dis 2019 Dec;74(6):791-802. [doi: 10.1053/j.ajkd.2019.05.025] [Medline: 31492487]
66. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. BMJ 2015 Feb 10;350:g7818. [doi: 10.1136/bmj.g7818] [Medline: 25670183]
67. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan A, King MT, the SPIRIT-PRO Group, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. JAMA 2018 Feb 06;319(5):483-494. [doi: 10.1001/jama.2017.21903] [Medline: 29411037]
68. Santana M, Feeny D. Framework to assess the effects of using patient-reported outcome measures in chronic care management. Qual Life Res 2014 Jun;23(5):1505-1513. [doi: 10.1007/s11136-013-0596-1] [Medline: 24318085]
69. Appendix 2 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man. Amsterdam, The Netherlands: European Medicines Agency; 2016. URL: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf [accessed 2020-10-09]
70. Value and Use of Patient-Reported Outcomes (PROs) in Assessing Effects of Medical Devices. Silver Spring, MD: US Food and Drug Administration; 2019. URL: https://www.fda.gov/media/109626/download [accessed 2020-10-09]
71. PRO Report Appendix: Patient-Reported Outcome Measure (PRO) Case Studies. Silver Spring, MD: US Food and Drug Administration; 2019. URL: https://www.fda.gov/media/125193/download [accessed 2020-10-09]
72. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Silver Spring, MD: US Food and Drug Administration; 2009. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims [accessed 2020-10-09]
73. Haase R, Schultheiss T, Kempcke R, Thomas K, Ziemssen T. Use and acceptance of electronic communication by patients with multiple sclerosis: a multicenter questionnaire study. J Med Internet Res 2012;14(5):e135 [FREE Full text] [doi: 10.2196/jmir.2133] [Medline: 23069209]
74. Kern R, Haase R, Eisele JC, Thomas K, Ziemssen T. Designing an Electronic Patient Management System for Multiple Sclerosis: Building a Next Generation Multiple Sclerosis Documentation System. Interact J Med Res 2016 Jan 08;5(1):e2 [FREE Full text] [doi: 10.2196/ijmr.4549] [Medline: 26746977]
75. Kremenchutzky M, Walt L. Perceptions of health status in multiple sclerosis patients and their doctors. Can J Neurol Sci 2013 Mar;40(2):210-218. [doi: 10.1017/s0317167100013755] [Medline: 23419570]
76. Stuchiner T, Chen C, Baraban E, Cohan S. The Pacific Northwest MS Registry: Year 4 Update and Diagnosis Validation (Poster DX82). 2012 Presented at: Consortium of the Multiple Sclerosis Centers (CMSC) Annual Meeting; May 31–June 4, 2012; San Diego, CA, USA.

XSL•FO
RenderX

77. Verdier-Taillefer MH, Roullet E, Cesaro P, Alpérovitch A. Validation of self-reported neurological disability in multiple sclerosis. Int J Epidemiol 1994 Feb;23(1):148-154. [doi: 10.1093/ije/23.1.148] [Medline: 8194911]
78. Frahm N, Hecker M, Zettl UK. Multi-drug use among patients with multiple sclerosis: A cross-sectional study of associations to clinicodemographic factors. Sci Rep 2019 Mar 06;9(1):3743 [FREE Full text] [doi: 10.1038/s41598-019-40283-5] [Medline: 30842515]
79. Jelinek GA, Weiland TJ, Hadgkiss EJ, Marck CH, Pereira N, van DMDM. Medication use in a large international sample of people with multiple sclerosis: associations with quality of life, relapse rate and disability. Neurol Res 2015 Aug;37(8):662-673 [FREE Full text] [doi: 10.1179/1743132815Y.0000000036] [Medline: 25905471]
80. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource. Silver Spring, MD: US Food and Drug Administration; 2016.
81. Hunter SF. Overview and diagnosis of multiple sclerosis. Am J Manag Care 2016 Jun;22(6 Suppl):s141-s150 [FREE Full text] [Medline: 27356023]

Abbreviations

DMT: disease-modifying therapy
DSST: Digit Symbol Substitution Test
LME: linear mixed effects
MS: multiple sclerosis
Neuro-QoL: Quality of Life in Neurological Disorders
PDDS: Patient-Determined Disease Steps
PRO: patient-reported outcome
QoL: quality of life

Edited by G Eysenbach; submitted 03.07.20; peer-reviewed by L Garcia-Gancedo; comments to author 27.07.20; revised version received 21.09.20; accepted 21.09.20; published 27.10.20.

Please cite as:
Pratap A, Grant D, Vegesna A, Tummalacherla M, Cohan S, Deshpande C, Mangravite L, Omberg L
Evaluating the Utility of Smartphone-Based Sensor Assessments in Persons With Multiple Sclerosis in the Real-World Using an App (elevateMS): Observational, Prospective Pilot Digital Health Study
JMIR Mhealth Uhealth 2020;8(10):e22108
URL: http://mhealth.jmir.org/2020/10/e22108/
doi:10.2196/22108
PMID:33107827

©Abhishek Pratap, Daniel Grant, Ashok Vegesna, Meghasyam Tummalacherla, Stanley Cohan, Chinmay Deshpande, Lara Mangravite, Larsson Omberg. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 27.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Fitness-Tracker Assisted Frailty-Assessment Before Transcatheter Aortic Valve Implantation: Proof-of-Concept Study

Markus Mach[1,2], MD; Victoria Watzal[2]; Waseem Hasan[3], BSc; Martin Andreas[1], MD, MBA, PhD; Bernhard Winkler[2], MD; Gabriel Weiss[2], MD; Andreas Strouhal[4], MD; Christopher Adlbrecht[4], MD, MBA, FHFA; Georg Delle Karth[4], MD; Martin Grabenwöger[2,5], MD

[1]Division of Cardiac Surgery, Department of Surgery, Medical University of Vienna, Vienna, Austria

[2]Department of Cardio-Vascular Surgery, Hospital Hietzing and Karl Landsteiner Institute for Cardio-Vascular Research, Vienna, Austria

[3]Faculty of Medicine, Imperial College London, London, United Kingdom

[4]Department of Cardiology, Vienna North Hospital - Clinic Floridsdorf and Karl Landsteiner Institute for Cardiovascular and Critical Care Research, Vienna, Austria

[5]Medical Faculty, Sigmund Freud University, Vienna, Austria

Corresponding Author:
Markus Mach, MD
Division of Cardiac Surgery
Department of Surgery
Medical University of Vienna
Ebene 20A
Waehringer Guertel 18-20
Vienna, 1090
Austria
Phone: 43 6645158360
Email: markus.mach@meduniwien.ac.at

Abstract

Background: While transcatheter aortic valve replacement (TAVR) has revolutionized the treatment of aortic valve stenosis, wearable health-monitoring devices are gradually transforming digital patient care.

Objective: The aim of this study was to develop a simple, efficient, and economical method for preprocedural frailty assessment based on parameters measured by a wearable health-monitoring device.

Methods: In this prospective study, we analyzed data of 50 consecutive patients with mean (SD) age of 77.5 (5.1) years and a median (IQR) European system for cardiac operative risk evaluation (EuroSCORE) II of 3.3 (4.1) undergoing either transfemoral or transapical TAVR between 2017 and 2018. Every patient was fitted with a wrist-worn health-monitoring device (Garmin Vivosmart 3) for 1 week prior to the procedure. Twenty different parameters were measured, and threshold levels for the 3 most predictive categories (ie, step count, heart rate, and preprocedural stress) were calculated. Patients were assigned 1 point per category for exceeding the cut-off value and were then classified into 4 stages (no, borderline, moderate, and severe frailty). Furthermore, the FItness-tracker assisted Frailty-Assessment Score (FIFA score) was compared with the scores of the preprocedural gait speed category derived from the 6-minute walk test (GSC-6MWT) and the Edmonton Frail Scale classification (EFS-C). The primary study endpoint was hospital mortality.

Results: The overall preprocedural stress level (P=.02), minutes of high stress per day (P=.02), minutes of rest per day (P=.045), and daily heart rate maximum (P=.048) as single parameters were the strongest predictors of hospital mortality. When comparing the different frailty scores, the FIFA score demonstrated the greatest predictive power for hospital mortality (FIFA area under the curve [AUC] 0.844, CI 0.656-1.000; P=.048; GSC-6MWT AUC 0.671, CI 0.487-0.855; P=.42; EFS-C AUC 0.636, CI 0.254-1.000; P=.44).

Conclusions: This proof-of-concept study demonstrates the strong predictive performance of the FIFA score compared to that of the conventional frailty assessments.

(JMIR Mhealth Uhealth 2020;8(10):e19227) doi:10.2196/19227

XSL•FO
RenderX

KEYWORDS

frailty; activity; fitness; tracker; transcatheter aortic valve implantation; transcatheter aortic valve repair

Introduction

Since the first transcatheter aortic valve replacement (TAVR) performed in 2002 by Alain Cribier, little development has taken place in the field of frailty assessment [1]. Despite being referred to as "the most important patient characteristic not included in current risk models" in the Valve Academic Research Consortium-2 (VARC-2) consensus document, the frailty assessment tool that is the most appropriate for clinical practice remains debated [2,3]. While existing objective performance measures such as gait speed or grip strength lack the specificity to distinguish between frail and nonfrail patients when used on their own, the development of multimodal and TAVR-specific frailty scores has yielded promising approaches that outperform the most widely used risk scores [4-9]. However, as most of these conventional frailty assessment tools are relatively time-consuming and resource-consuming, frailty is still often assessed subjectively based upon either an "eyeball test," which is an "end-of-the-bed" assessment, or not measured at all in clinical practice [5].

Wearable health-monitoring devices have already been used for patient-related basic measurements in different specialties. Several studies have proven that such devices can reliably predict flares in patients with rheumatic diseases or reverse remodeling in patients with heart failure treated with resynchronization therapy [2-4]. These mediums, therefore, represent a potential platform where measurements relevant to the patients' health status and subsequent intervention can be collected without significantly interfering with the activities of both patients and clinical staff. The aim of this study was to develop a preprocedural frailty assessment based on data measured by a wearable health-monitoring device.

Methods

Patient Population and Study Design

The present open-label, nonrandomized proof-of-concept study features permanent preprocedural activity monitoring with a wearable health-monitoring device (Vivosmart3, software version 2.9-5.10, Sensor Hub software version 6.3., Garmin) during the week prior to the intervention. All patients treated between March and December 2018 at the Heart Center Hietzing (Vienna, Austria) were invited to participate and they provided written informed consent. Baseline patient characteristics, as well as procedural and outcome data, were obtained from the VIennaCardioThOracic Aortic Valve RegistrY. The Ethics Committee of the City of Vienna approved the study (protocol EK-048-0318).

The indication for TAVR was assessed by a multidisciplinary Heart Team, consisting of cardiologists, cardiac surgeons, anesthetists, radiologists, and geriatricians. Intermediate-risk or high-risk patients with severe aortic stenosis were subject to evaluation and were included, conforming to the current guidelines of the European Association of Cardio-Thoracic Surgery [5]. Hence, inclusion and exclusion criteria for the trial were the same as those for the procedure itself. No patient had physical limitations preventing him/her from wearing an activity tracker. Surgical risk stratification was based on the following risk algorithms: the logistic European system for cardiac operative risk evaluation (EuroSCORE), EuroSCORE II, and the Society of Thoracic Surgeons (STS) score [6-8]. To determine the prevalence of frailty, the Edmonton Frail Scale (EFS) was used and the 6-minute walk test (6MWT) was performed prior to the intervention [9,10]. The primary endpoint of the study was in-hospital all-cause mortality. The evaluation and documentation of the postprocedural outcome data followed the VARC-2 criteria [11].

Frailty Assessment

The EFS includes an objective evaluation of the patient's health status as well as the quality of life and clinical frailty. It comprises 5 categories: 0-5 points, not frail; 6-7 points, vulnerable; 8-9 points, mild frailty; 10-11 points, moderate frailty; and >12 points, severe frailty [9]. As an additional frailty assessment tool, the 6MWT was performed in the same visit in the outpatient department. We distinguished between very slow walkers (<0.5 m/s), slow walkers (≥0.5 m/s), normal walkers (>0.83 m/s), and patients, in whom the 6MWT was not feasible because of general weakness, being bedridden, being wheelchair-bound, or pain [12].

Wearable Health-Monitoring Device

The Vivosmart 3 continuously measures daily physical activity. It assesses the daily step count, distance covered (in kilometers), calories burned, the time spent in different stress levels, the hours and depth of sleep, the minimum and maximum heart rate, and the number of flights of stairs climbed. Several reports have validated the accuracy of the heart rate measurement by Garmin Vivosmart devices with simultaneous electrocardiogram readings [13,14], and the device has shown excellent test-retest reliability as well as optimal step count accuracy at low and moderate walking speeds [15]. The stress level measurement calculated by the device is based on the analysis of the heart rate variability and it is a good reflector of autonomic activity. Real-life heart rate measurements using photoplethysmography are combined with mathematical modeling and algorithms to infer the present stress level of the patient wearing the monitoring device [16].

The waterproof device was worn for at least one week prior to the procedure around the wrist of the nondominant hand to avoid additional step count through repetitive gestures during daily activities. Patients were encouraged to follow their regular routines and activities while continuously wearing the device and they were instructed not to remove it temporarily for bathing, swimming, or sleeping.

On the day of admission, the device was removed, and the activity data were uploaded via password-encrypted Bluetooth transfer to anonymous accounts in the Garmin Connect app

XSL•FO
RenderX

(version 3.22.0.1-4.20). The device had no GPS-tracking function, thereby ensuring maximum data security.

Fitness-Tracker Assisted Frailty-Assessment

The objective of this study was to develop a modern, easy-to-use, time-saving, and resource-saving frailty assessment method that allows accurate prediction of adverse outcomes after TAVR. The daily output was used to calculate the weekly average values, excluding the incomplete activity data available from the first and last day of monitoring. Threshold levels in 3 predefined categories (ie, heart rate, preprocedural stress, and walking) were calculated. The patients were assigned 1 point per category when exceeding (in categories with positive correlation) or subceeding (in categories with negative correlation) the threshold levels and then grouped into 4 categories (0, no frailty; 1, mild frailty; 2, moderate frailty; 3, severe frailty). Sleep pattern parameters were assessed for their discriminatory power but not included in the FItness-tracker assisted Frailty-Assessment (FIFA) score, as currently, no evidence exists that such parameters are correlated with adverse postoperative outcomes.

Statistical Analysis

Continuous data were expressed as mean (SD) or median (IQR) depending on the normal distribution. The comparison of continuous data between groups was performed using the Kruskal-Wallis test (H test) or univariate analysis when appropriate. A chi-square or Fisher exact test was performed to compare categorical data. Posthoc testing was performed using the Bonferroni corrected z test, Duncan, or Scheffé test depending on variance homogeneity and sample size.

For the development of the FIFA score, different parameters identifying frail patients and positively predicting hospital mortality were determined by receiver operating characteristic (ROC) analysis. The discriminatory ability was assessed via the area under the curve (AUC). Threshold values were calculated retrospectively with the Youden Index *(J = sensitivity + specificity –1)*. The different frailty and risk assessment tools were then again compared for their predictive power of hospital mortality with ROC and AUC analysis. The alpha level was set at less than .05. All tests were two-tailed. Statistical calculations were performed using the SPSS statistical software version 24.0 (IBM Corp).

Results

Baseline Patient Characteristics

The baseline characteristics of the cohort are outlined in Table 1. In total, 50 patients (22 women, 44%) with a mean (SD) age of 77.5 (5.1) years were included. The surgical risk profile of the cohort was as follows: mean (SD) EuroSCORE II, 3.3 (4.1) and median (IQR) STS score, 2.9 (2.3). The patients were stratified according to their FIFA score classification (no, mild, moderate, or severe frailty). There was a strong correlation between the FIFA score and the baseline serum albumin level, as more frail patients had significantly lower albumin levels (P=.005). Apart from the serum albumin levels, the baseline characteristics, including echocardiographic parameters, were similar among the groups.

XSL•FO
RenderX

Table 1. Baseline characteristics of the cohort.

Characteristics	Overall, N=50	No frailty, n=11	Mild frailty, n=10	Moderate frailty, n=21	Severe frailty, n=8	*P* value[a]
Demographics						
Age (years), mean (SD)	77.5 (5.1)	78.0 (3.3)	76.4 (4.8)	78.7 (4.9)	74.9 (7.2)	.28
Female, n (%)	22 (44)	5 (46)	6 (60)	8 (38)	3 (38)	.69
Body mass index (kg/m^2), mean (SD)	27.2 (5.3)	28.5 (3.9)	27.3 (7.5)	26.1 (4.3)	28.0 (6.7)	.63
Serum albumin (g/dl), mean (SD)	40.3 (3.3)	42.7 (2.0)	38.6 (3.9)	40.4 (2.2)	38.1 (3.9)	*.005*
Risk scores						
EuroSCORE II[b], median (IQR)	3.3 (4.1)	4.9 (6.4)	2.6 (1.7)	3.3 (5.2)	3.4 (2.6)	.32
Logistic EuroSCORE, mean (SD)	12.0 (8.1)	16.3 (11.0)	11.3 (6.4)	11.6 (7.4)	8.2 (5.4)	.17
STS[c] score, median (IQR)	2.9 (2.3)	2.7 (1.7)	2.4 (2.9)	3.6 (3.6)	2.8 (1.6)	.34
Comorbidities						
Hypertension, n (%)	47 (94)	11 (100)	9 (90)	19 (91)	8 (100)	.58
Diabetes mellitus (IDDM)[d], n (%)	19 (38)	2 (18)	5 (50)	9 (43)	3 (38)	.59
Atrial fibrillation, n (%)	26 (52)	8 (73)	4 (40)	10 (48)	4 (40)	.45
Peripheral vascular disease, n (%)	23 (46)	5 (46)	3 (30)	11 (52)	4 (50)	.69
COPD[e], n (%)	15 (30)	2 (18)	5 (50)	5 (24)	3 (38)	.09
Cerebrovascular accident, n (%)	10 (20)	2 (18)	1 (10)	6 (29)	1 (13)	.76
Creatinine (mg/dL), median (IQR)	1.1 (0.7)	1.1 (0.6)	0.9 (0.5)	1.2 (1.0)	1.5 (0.8)	.16
Prior myocardial infarction, n (%)	6 (12)	2 (18)	1 (10)	3 (14)	0 (0)	.65
Prior PCI[f], n (%)	15 (30)	6 (55)	0 (0)	5 (24)	4 (50)	*.03*
Previous CABG[g], n (%)	6 (12)	3 (27)	1 (10)	2 (10)	0 (0)	.30
Previous valve surgery, n (%)	4 (8)	2 (18)	0 (0)	2 (10)	0 (0)	.37
sPAP[h], mean (SD)	47.1 (15.8)	48.6 (20.9)	36.3 (3.2)	47.0 (14.5)	51.8 (14.0)	.49
LVEF%[i], mean (SD)	50.0 (15.5)	54.3 (13.1)	63.5 (12.1)	41.7 (16.0)	50.8 (12.2)	.10

[a]Statistically significant *P* values are italicized in the table.

[b]EUROScore: European system for cardiac operative risk evaluation.

[c]STS: Society of Thoracic Surgeons.

[d]IDDM: insulin-dependent diabetes mellitus.

[e]COPD: chronic obstructive pulmonary disease.

[f]PCI: percutaneous coronary intervention.

[g]CABG: coronary artery bypass graft.

[h]sPAP: systolic pulmonary artery pressure.

[i]LVEF: left ventricular ejection fraction.

Procedural Outcomes and Adverse Events

The procedural outcomes are shown in Table 2. The procedural time, total intensive care unit hours, and hospital stay did not differ between the frailty stages. The same applied to adverse events, as seen in Table 3.

XSL•FO
RenderX

Table 2. Procedural characteristics and outcomes in the different groups of the cohort.

Procedures	Overall, N=50	No frailty, n=11	Mild frailty, n=10	Moderate frailty, n=21	Severe frailty, n=8	*P* value
Prosthesis used, n (%)						.73
Sapien XT	4 (8)	2 (18)	0 (0)	2 (10)	0 (0)	
Sapien 3	27 (54)	6 (55)	6 (60)	10 (48)	5 (63)	
Symetis Acurate	15 (30)	3 (27)	3 (30)	6 (29)	3 (38)	
Core Valve Evolut	4 (8)	0 (0)	1 (10)	3 (14)	0 (0)	
Total ICU[a] hours, median (IQR)	22.0 (29.0)	32.0 (48.0)	23.0 (33.0)	21.0 (4.0)	18.5 (281.0)	.21
Length of stay in days, median (IQR)	13.0 (10.0)	15.0 (13.0)	11.0 (12.0)	13.0 (8.0)	17.0 (20.0)	.42

[a]ICU: intensive care unit.

Table 3. Adverse events in the different groups of the cohort.

Types of adverse events	Overall, N=50	No frailty, n=11	Mild frailty, n=10	Moderate frailty, n=21	Severe frailty, n=8	*P* value
Cerebrovascular events, n (%)						.70
Transient ischemic attack	2 (4)	0 (0)	0 (0)	1 (5)	1 (13)	
Major stroke	1 (2)	0 (0)	0 (0)	1 (5)	0 (0)	
Access site complication, n (%)						.52
Minor access complication	3 (6)	1 (9)	1 (10)	1 (5)	0 (0)	
Major access complication	1 (2)	0 (0)	1 (10)	0 (0)	0 (0)	
Myocardial infarction	2 (4)	0 (0)	0 (0)	1 (5)	1 (13)	.49
Postoperative pacemaker	7 (14)	2 (18)	2 (20)	2 (10)	1 (13)	.84
Postoperative acute kidney injury, n (%)						.43
Stage I	6 (12)	0 (0)	1 (10)	3 (14)	2 (25)	
Stage II	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
Stage III	2 (4)	0 (0)	0 (0)	2 (10)	0 (0)	
Postoperative bleeding, n (%)						.43
Minor bleeding	7 (14)	3 (27)	1 (10)	2 (10)	1 (13)	
Major bleeding	2 (4)	1 (9)	0 (0)	0 (0)	1 (13)	
Hospital mortality	3 (6)	0 (0)	0 (0)	1 (5)	2 (25)	.09

Discriminatory Power of Health-Monitoring Device–Generated Measurements as Predictors of Hospital Mortality

The overall preprocedural stress level (AUC 0.915, CI 0.795-1.000; *P*=.02), the minutes spent in high stress (AUC 0.908, CI 0.801-1.000; *P*=.02), the minutes at rest (AUC 0.848, CI 0.711-0.984; *P*=.045), and the heart rate maximum (AUC 0.844, CI 0.717-0.97; *P*=.048) as individual parameters were the strongest predictors of hospital mortality in the ROC analysis. High preprocedural stress levels, long periods spent in high stress, and an elevated maximum heart rate correlated positively with hospital mortality, whereas the amount of time at rest had an inverse correlation with hospital mortality (Table 4).

XSL•FO
RenderX

Table 4. Receiver operating characteristic analysis and hospital mortality.

Parameters	Area under the curve	Confidence interval		*P* value[a]
		Lower	Upper	
Tracker parameters				
Heart rate				
Heart rate minimum[b]	0.688	0.435	0.941	.28
Heart rate maximum[b]	0.844	0.717	0.971	*.048*
Preprocedural stress				
Stress level at rest[c]	0.848	0.711	0.984	*.045*
Stress level, high[b]	0.908	0.801	1.000	*.02*
Overall stress level[b]	0.915	0.795	1.000	*.02*
Walking				
Step count[c]	0.518	0.069	0.966	.92
Walking distance[c]	0.525	0.076	0.974	.89
Sleeping				
Sleep state, deep[c]	0.560	0.085	1.000	.95
Sleep, light[b]	0.457	0.059	0.856	.81
Sleep, awake[b]	0.532	0.113	0.951	.85
Sleep, in total[c]	0.468	0.071	0.865	.95
Risk scores				
EuroSCORE[d] II[b]	0.624	0.479	0.769	.47
Logistic EuroSCORE[b]	0.461	0.241	0.681	.82
STS[e] score[b]	0.376	0.198	0.554	.47
Frailty scores				
Gait speed classification (6MWT[f])[b]	0.671	0.487	0.855	.42
EFS[g] classification[b]	0.636	0.254	1.000	.44
FIFA[h] score	0.844	0.656	1.000	*.048*

[a]Statistically significant *P* values are italicized in the table.

[b]Positive correlation with hospital mortality.

[c]Negative correlation with hospital mortality.

[d]EUROScore: European system for cardiac operative risk evaluation.

[e]STS: Society of Thoracic Surgeons.

[f]6MWT: 6-minute walk test.

[g]EFS: Edmonton Frail Scale.

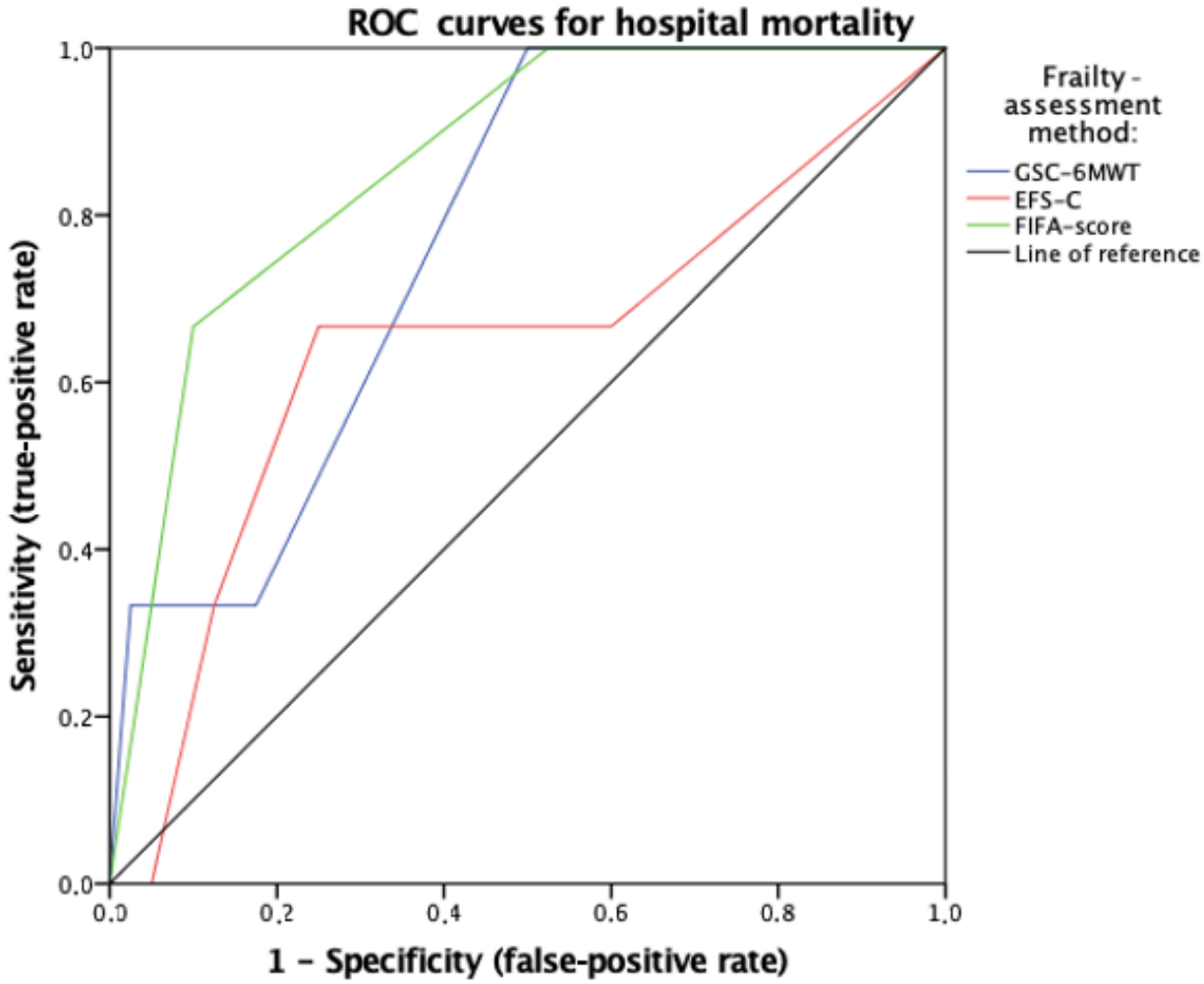
[h]FIFA: FItness-tracker assisted Frailty-Assessment.

Comparison of the Incremental Predictive Value of Frailty Scores

Depending on the frailty-assessment method used, the prevalence of frailty ranged from 55.5% (FIFA) and 60.6% (EFS-C) to 62.5% (gait speed category). Therefore, frailty was defined as either “moderate frailty” or “severe frailty” in the FIFA score classification, as “mild frailty,” “moderate frailty,” and “severe frailty” in the EFS-C or as unfeasible 6MWT or gait speed less than 0.83 m/s. The FIFA score demonstrated the highest predictability of hospital mortality (AUC 0.844, CI 0.656-1.000; *P*=.048) when compared to 6MWT gait speed classification (AUC 0.671, CI 0.487-0.855; *P*=.42) and EFS-C (AUC 0.636, CI 0.254-1.000; *P*=.44) (Table 4, Figure 1).

XSL•FO
RenderX

Figure 1. Receiver operating characteristic curves predicting hospital mortality. ROC: receiver operating characteristic; GSC-6MWT: gait speed category derived from the 6-minute walk test; EFS-C: Edmonton Frail Scale classification; FIFA: FItness-tracker assisted Frailty-Assessment.



Discussion

Principal Results

This study demonstrates the strong predictive performance of a modern health-monitoring device–based frailty assessment, and, to the best of our knowledge, this is the first trial to investigate the potential of a wearable activity-monitoring device in this specific clinical setting. The main findings of our study were as follows. The FIFA score correctly identifies frail patients, as demonstrated by the strong correlation with baseline serum albumin levels—a well-established biomarker for frailty. Device-recorded preprocedural stress and heart rates in patients with TAVR are independent predictors with increased postprocedural hospital mortality. The FIFA score outperformed conventional frailty assessment methods by correctly identifying patients at higher risk for procedure-related mortality.

The FIFA trial has developed a modern, intuitive, and efficient frailty assessment tool that adds substantially to the growing body of evidence on the relevance of frailty in the selection of patients with TAVR. Its potential in clinical practice is substantiated by its relationship with established measures of frailty such as serum albumin levels, which play a prominent role in the estimation of frailty by reflecting the extent of muscle weakness and malnutrition. Preoperative serum albumin levels <3.5 g/dl have been proven to be an independent predictor of mortality and the inverse and proportional correlation between baseline serum albumin levels and the FIFA score is thus indicative of the assessment tool's ability to correctly identify frail patients [9].

Furthermore, it has also been shown for the first time that several preprocedural stress parameters (the overall preprocedural stress level, net time in high stress/at rest) are independent predictors of postprocedural mortality. Establishing the link between stress measurements and postprocedural outcomes was enabled by wearable devices and represents a novel finding. The physiological basis for this relationship may be underpinned by the pivotal role that heart rate variability has in deriving the stress parameters. Several studies have verified the association between deviations in heart rate variability scores and postoperative outcomes. The standard deviation of normal to normal intervals has been established as an independent predictor of increased postprocedural mortality in various clinical populations [17-19]. Varadhan et al [20] demonstrated that impaired cardiac autonomic function reflected by a lower heart rate variability is a predictor of frailty and mortality in

XSL•FO
RenderX

older women. It is assumed that the underlying pathophysiological process may be the degradation of autonomic control mechanisms, which leads to an overactivated sympathetic system and subsequent neuroendocrine responses in resting stages, thereby hampering physiological recovery. Accordingly, the FIFA trial has also demonstrated that an elevated preprocedural overall maximum heart rate correlates positively with increased hospital mortality. We, therefore, postulate that stress parameters represent a directly measurable surrogate marker of impaired hemodynamic regulation in those patients. Consequently, stress parameters as measured with activity-monitoring devices may help detect high-risk patients and reduce therapeutic futility. Importantly, this information could be combined with other established parameters (eg, albumin levels) in improving the specificity of preoperative risk assessments.

Conventional frailty scores (the EFS-C and the 6MWT gait speed classification) showed limited association with hospital mortality compared to the FIFA score, but given the fact that these criteria were designed to predict long-term outcomes, correlations may improve, as the study progresses over time. In contrast to our findings, gait speed as a single measure was shown to be independently associated with mortality after TAVR procedures. Alfredsson et al [21] reported that the slowest walkers in their study had a 35% higher 30-day mortality than normal walkers, and each 0.2-m/s decrease in gait speed corresponded to an 11% mortality increase. Although the EFS applied in this study has not been subject to validation in a population with TAVR, it demonstrated good predictability of 30-day mortality in older patients after cardiac surgery [22]. The differences in the findings may be attributed to the smaller population examined in this study; however, the ability of the FIFA score to predict short-term mortality within our initial study population is promising. With an AUC of 0.844, it demonstrates excellent discriminatory power [23]. Moreover, not only does it correctly identify frail patients and potentially exceed conventional frailty assessment methods in their predictive ability of postprocedural mortality, but its ease of use and economical approach saving both time and valuable human resources make it an especially attractive tool to use in clinical practice. A further benefit is that while patients may perform differently in a clinical test setting, the gathering of real-life data allows more objective interpretation of patients' frailty with high interobserver reliability.

Limitations

There are several limitations to this study. As this is a proof-of-concept study, benchmark data had to be established for comparison with conventional frailty assessment methods; hence, the open-label, nonrandomized, prospective trial design. Owing to the lack of existing data, the concept was established and tested in a relatively small number of patients. Even though the primary endpoint is sufficiently powered, the concept needs to be validated in a larger patient population. Furthermore, the short postprocedural observation period may lead to underestimation of the predictive value of the FIFA score as most frailty-related studies assess long-term outcomes, including 1-year mortality.

Conclusion

For the first time, the strong predictive performance of wearable health-monitoring device–related assessment compared to that of conventional frailty methods has been shown. The ease of use, objectivity, and high predictive performance may not only save valuable clinical resources but ultimately improve patient selection and safety. The promising initial results warrant further evaluation of the FIFA score classification as a predictor of short-term and long-term mortality after structural heart interventions or conventional surgery.

Acknowledgments

This work was supported by the Medical Scientific Fund of the Mayor of the City of Vienna; grant number: 18116.

Conflicts of Interest

MM has received a research grant from Edwards Lifesciences, JenaValve, and Symetis. MA is proctor (Edwards, Abbott) and advisor (Medtronic). All other authors have reported that they have no relationships relevant to the content.

References

1. Cribier A, Eltchaninoff H, Bash A, Borenstein N, Tron C, Bauer F, et al. Percutaneous transcatheter implantation of an aortic valve prosthesis for calcific aortic stenosis: first human case description. Circulation 2002 Dec 10;106(24):3006-3008. [doi: 10.1161/01.cir.0000047200.36165.b8] [Medline: 12473543]
2. Stocker TJ, Scheck F, Orban M, Braun D, Hertell H, Lackermair K, et al. Physical activity tracking in correlation to conventional heart failure monitoring assessing improvements after transcatheter mitral and tricuspid valve repair. Eur J Heart Fail 2019 Jul;21(7):943-945 [FREE Full text] [doi: 10.1002/ejhf.1418] [Medline: 30690826]
3. Jacquemin C, Molto A, Servy H, Sellam J, Foltz V, Gandjbakhch F, et al. Flares assessed weekly in patients with rheumatoid arthritis or axial spondyloarthritis and relationship with physical activity measured using a connected activity tracker: a 3-month study. RMD Open 2017;3(1):e000434. [doi: 10.1136/rmdopen-2017-000434] [Medline: 28879046]
4. Vegh EM, Kandala J, Orencole M, Upadhyay GA, Sharma A, Miller A, et al. Device-measured physical activity versus six-minute walk test as a predictor of reverse remodeling and outcome after cardiac resynchronization therapy for heart failure. Am J Cardiol 2014 May 01;113(9):1523-1528. [doi: 10.1016/j.amjcard.2014.01.430] [Medline: 24641966]

XSL•FO
RenderX

5. Baumgartner H, Falk V, Bax JJ, De Bonis M, Hamm C, Holm PJ, et al. 2017 ESC/EACTS Guidelines for the Management of Valvular Heart Disease. Rev Esp Cardiol (Engl Ed) 2018 Feb;71(2):110. [doi: 10.1016/j.rec.2017.12.013] [Medline: 29425605]
6. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999 Jul;16(1):9-13. [doi: 10.1016/s1010-7940(99)00134-7] [Medline: 10456395]
7. Nashef SAM, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. Eur J Cardiothorac Surg 2012 Apr;41(4):734-744. [doi: 10.1093/ejcts/ezs043] [Medline: 22378855]
8. O'Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Society of Thoracic Surgeons Quality Measurement Task Force. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2 isolated valve surgery. Ann Thorac Surg 2009 Jul;88(1 Suppl):S23-S42. [doi: 10.1016/j.athoracsur.2009.05.056] [Medline: 19559823]
9. Rolfson DB, Majumdar SR, Tsuyuki RT, Tahir A, Rockwood K. Validity and reliability of the Edmonton Frail Scale. Age Ageing 2006 Sep;35(5):526-529 [FREE Full text] [doi: 10.1093/ageing/afl041] [Medline: 16757522]
10. Green P, Cohen DJ, Généreux P, McAndrew T, Arnold SV, Alu M, et al. Relation between six-minute walk test performance and outcomes after transcatheter aortic valve implantation (from the PARTNER trial). Am J Cardiol 2013 Sep 01;112(5):700-706 [FREE Full text] [doi: 10.1016/j.amjcard.2013.04.046] [Medline: 23725996]
11. Kappetein AP, Head SJ, Généreux P, Piazza N, van Mieghem NM, Blackstone EH, et al. Updated standardized endpoint definitions for transcatheter aortic valve implantation: the Valve Academic Research Consortium-2 consensus document. Eur Heart J 2012 Oct;33(19):2403-2418. [doi: 10.1093/eurheartj/ehs255] [Medline: 23026477]
12. Kano S, Yamamoto M, Shimura T, Kagase A, Tsuzuki M, Kodama A, et al. Gait Speed Can Predict Advanced Clinical Outcomes in Patients Who Undergo Transcatheter Aortic Valve Replacement: Insights From a Japanese Multicenter Registry. Circ Cardiovasc Interv 2017 Sep;10(9). [doi: 10.1161/CIRCINTERVENTIONS.117.005088] [Medline: 28916601]
13. Boudreaux BD, Hebert EP, Hollander DB, Williams BM, Cormier CL, Naquin MR, et al. Validity of Wearable Activity Monitors during Cycling and Resistance Exercise. Med Sci Sports Exerc 2018 Mar;50(3):624-633. [doi: 10.1249/MSS.0000000000001471] [Medline: 29189666]
14. Reddy RK, Pooni R, Zaharieva DP, Senf B, El Youssef J, Dassau E, et al. Accuracy of Wrist-Worn Activity Monitors During Common Daily Physical Activities and Types of Structured Exercise: Evaluation Study. JMIR Mhealth Uhealth 2018 Dec 10;6(12):e10338 [FREE Full text] [doi: 10.2196/10338] [Medline: 30530451]
15. Fokkema T, Kooiman TJM, 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]
16. Kettunen JS, Saalasti S. Procedure for detection of stress by segmentation and analyzing a heart beat signal. 2005. URL: https://patentimages.storage.googleapis.com/c0/da/86/0013898102e400/US20050256414A1.pdf [accessed 2020-09-20]
17. Ernst G, Watne LO, Frihagen F, Wyller TB, Dominik A, Rostrup M. Decreases in heart rate variability are associated with postoperative complications in hip fracture patients. PLoS One 2017;12(7):e0180423 [FREE Full text] [doi: 10.1371/journal.pone.0180423] [Medline: 28742855]
18. Dekker JM, Crow RS, Folsom AR, Hannan PJ, Liao D, Swenne CA, et al. Low heart rate variability in a 2-minute rhythm strip predicts risk of coronary heart disease and mortality from several causes: the ARIC Study. Atherosclerosis Risk In Communities. Circulation 2000 Sep 12;102(11):1239-1244. [doi: 10.1161/01.cir.102.11.1239] [Medline: 10982537]
19. Karmali SN, Sciusco A, May SM, Ackland GL. Heart rate variability in critical care medicine: a systematic review. Intensive Care Med Exp 2017 Dec;5(1):33 [FREE Full text] [doi: 10.1186/s40635-017-0146-1] [Medline: 28702940]
20. Varadhan R, Chaves PHM, Lipsitz LA, Stein PK, Tian J, Windham BG, et al. Frailty and impaired cardiac autonomic control: new insights from principal components aggregation of traditional heart rate variability indices. J Gerontol A Biol Sci Med Sci 2009 Jun;64(6):682-687 [FREE Full text] [doi: 10.1093/gerona/glp013] [Medline: 19223607]
21. Alfredsson J, Stebbins A, Brennan JM, Matsouaka R, Afilalo J, Peterson ED, et al. Gait Speed Predicts 30-Day Mortality After Transcatheter Aortic Valve Replacement: Results From the Society of Thoracic Surgeons/American College of Cardiology Transcatheter Valve Therapy Registry. Circulation 2016 Apr 05;133(14):1351-1359. [doi: 10.1161/CIRCULATIONAHA.115.020279] [Medline: 26920495]
22. Amabili P, Wozolek A, Noirot I, Roediger L, Senard M, Donneau A, et al. The Edmonton Frail Scale Improves the Prediction of 30-Day Mortality in Elderly Patients Undergoing Cardiac Surgery: A Prospective Observational Study. J Cardiothorac Vasc Anesth 2019 Apr;33(4):945-952. [doi: 10.1053/j.jvca.2018.05.038] [Medline: 30049520]
23. Hosmer JD, Lemeshow S, Sturdivant R. Applied logistic regression. Hoboken, New Jersey, United States: Wiley; 2013.

Abbreviations

6MWT: 6-minute walk test
AUC: area under the curve
EFS-C: Edmonton Frail Scale classification
EuroSCORE: European system for cardiac operative risk evaluation

XSL•FO
RenderX

FIFA: FItness-tracker assisted Frailty-Assessment
ROC: receiver operating characteristic
STS: Society of Thoracic Surgeons
TAVR: transcatheter aortic valve replacement
VARC: Valve Academic Research Consortium

Edited by G Eysenbach; submitted 09.04.20; peer-reviewed by P Neves, A Colli, R Pooni; comments to author 12.06.20; revised version received 18.06.20; accepted 25.06.20; published 15.10.20.

Please cite as:
Mach M, Watzal V, Hasan W, Andreas M, Winkler B, Weiss G, Strouhal A, Adlbrecht C, Delle Karth G, Grabenwöger M
Fitness-Tracker Assisted Frailty-Assessment Before Transcatheter Aortic Valve Implantation: Proof-of-Concept Study
JMIR Mhealth Uhealth 2020;8(10):e19227
URL: https://mhealth.jmir.org/2020/10/e19227
doi:10.2196/19227
PMID:33055057

©Markus Mach, Victoria Watzal, Waseem Hasan, Martin Andreas, Bernhard Winkler, Gabriel Weiss, Andreas Strouhal, Christopher Adlbrecht, Georg Delle Karth, Martin Grabenwöger. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 15.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Use of mHealth Devices to Screen for Atrial Fibrillation: Cost-Effectiveness Analysis

Godwin D Giebel[1], MSc

E-Government/E-Health, Department of Business Information Systems, Baden-Wuerttemberg Cooperative State University Mannheim, Mannheim, Germany

Corresponding Author:
Godwin D Giebel, MSc
E-Government/E-Health, Department of Business Information Systems
Baden-Wuerttemberg Cooperative State University Mannheim
Coblitzallee 1-9
Mannheim, 68163
Germany
Phone: 49 157 57953048
Email: godwin.giebel@dhbw-mannheim.de

Abstract

Background: With an estimated prevalence of around 3% and an about 2.5-fold increased risk of stroke, atrial fibrillation (AF) is a serious threat for patients and a high economic burden for health care systems all over the world. Patients with AF could benefit from screening through mobile health (mHealth) devices. Thus, an early diagnosis is possible with mHealth devices, and the risk for stroke can be markedly reduced by using anticoagulation therapy.

Objective: The aim of this work was to assess the cost-effectiveness of algorithm-based screening for AF with the aid of photoplethysmography wrist-worn mHealth devices. Even if prevented strokes and prevented deaths from stroke are the most relevant patient outcomes, direct costs were defined as the primary outcome.

Methods: A Monte Carlo simulation was conducted based on a developed state-transition model; 30,000 patients for each CHA_2DS_2-VASc (Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category [female]) score from 1 to 9 were simulated. The first simulation served to estimate the economic burden of AF without the use of mHealth devices. The second simulation served to simulate the economic burden of AF with the use of mHealth devices. Afterwards, the groups were compared in terms of costs, prevented strokes, and deaths from strokes.

Results: The CHA_2DS_2-VASc score as well as the electrocardiography (ECG) confirmation rate had the biggest impact on costs as well as number of strokes. The higher the risk score, the lower were the costs per prevented stroke. Higher ECG confirmation rates intensified this effect. The effect was not seen in groups with lower risk scores. Over 10 years, the use of mHealth (assuming a 75% ECG confirmation rate) resulted in additional costs (€1=US $1.12) of €441, €567, €536, €520, €606, €625, €623, €692, and €847 per patient for a CHA_2DS_2-VASc score of 1 to 9, respectively. The number of prevented strokes tended to be higher in groups with high risk for stroke. Higher ECG confirmation rates led to higher numbers of prevented strokes. The use of mHealth (assuming a 75% ECG confirmation rate) resulted in 25 (7), –68 (–54), 98 (–5), 266 (182), 346 (271), 642 (440), 722 (599), 1111 (815), and 1116 (928) prevented strokes (fatal) for CHA_2DS_2-VASc score of 1 to 9, respectively. Higher device accuracy in terms of sensitivity led to even more prevented fatal strokes.

Conclusions: The use of mHealth devices to screen for AF leads to increased costs but also a reduction in the incidence of stroke. In particular, in patients with high CHA_2DS_2-VASc scores, the risk for stroke and death from stroke can be markedly reduced.

(JMIR Mhealth Uhealth 2020;8(10):e20496) doi:10.2196/20496

KEYWORDS

mHealth; atrial fibrillation; screening devices; strokes; cost-effectiveness; photoplethysmography

XSL•FO
RenderX

Introduction

With an estimated prevalence of about 3%, atrial fibrillation (AF) is one of the most common cardiac arrhythmias [1]. On the one hand, AF can be considered as an independent disease; on the other hand, AF can be considered as a risk factor for secondary diseases. AF is associated with an increased risk of all-cause mortality, as well as cardiovascular mortality and stroke [2,3].

An established way to estimate the risk for stroke in patients with AF is the CHA_2DS_2-VASc score (Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category [female]) [4]. To reduce the risk of stroke, it is recommended to consider anticoagulation therapy after the diagnosis of AF in male patients with a CHA_2DS_2-VASc score of 1 and in female patients with a score of 2 [1].

AF can occur in 5 different forms (first diagnosed, paroxysmal, persistent, long-standing persistent, and permanent), which can be either symptomatic or asymptomatic. The European Society of Cardiology recommends opportunistic screening by pulse taking or electrocardiogram rhythm strip in patients older than 65 years because undiagnosed AF remains a common problem [1].

While screening during visits to the doctor often misses irregular forms of AF, screening with the aid of implantable cardioverter-defibrillators, pacemakers, and implantable loop recorders is, at the same time, only eligible for a minority of patients with previous cardiac illnesses. An innovative and accurate approach to detect AF might be the application of mobile health (mHealth) in combination with algorithms. Nevertheless, the diagnosis should always be confirmed by electrocardiography (ECG) as the gold standard [1].

The aim of this work was to evaluate the fictitious use of photoplethysmography (PPG) in combination with algorithms integrated in wrist-worn mHealth devices over a period of 10 years to support the diagnosis of AF as an add-on to the existing health care system in Germany. The focus of this study was on the different outcomes. The primary outcome was AF-related costs. The secondary outcomes were the number of prevented strokes and prevented deaths from stroke.

Methods

Model Description

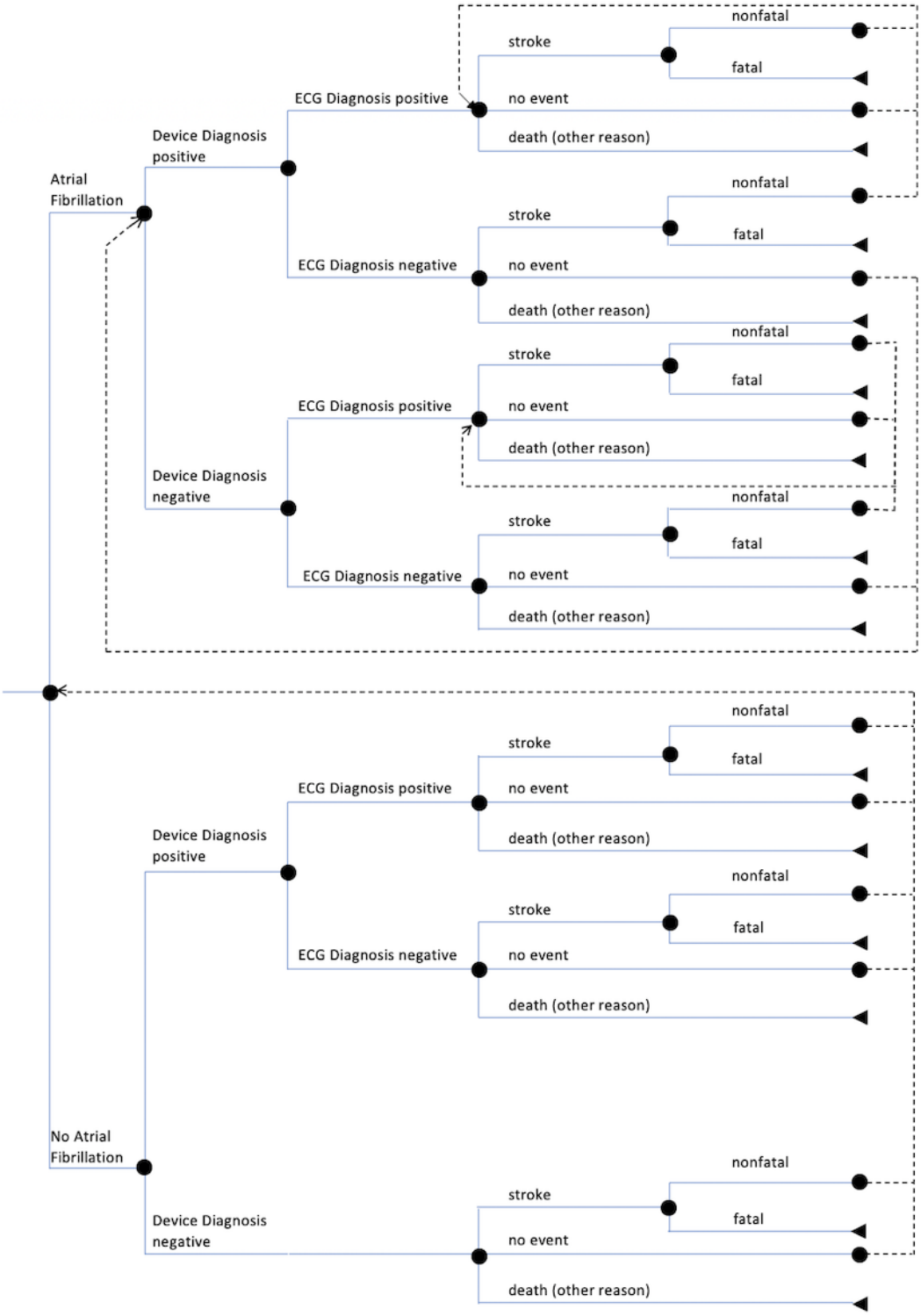
A Markov Model, a practical tool for medical decision making [5], was developed to assess the health economic impact of wrist-worn PPG mHealth devices in the diagnosis of AF. A model previously published by Reinhold et al who compared implantable cardioverter-defibrillators was adapted [6]. A Monte Carlo simulation was conducted based on a developed state-transition model. Depending on the underlying patient group, either with or without devices, different states and transitions were restricted (Figure 1 and Figure 2). For both groups, simulations were based on a time horizon of 10 years. This was assumed since technological changes might probably lead to even more accurate devices. During this period, changings of state were calculated based on a 1-year cycle. Whether the health state of individuals changes or not, depends on the previous state as well as on defined probabilities of state transition as listed in Table 1.

The simulation ends for an individual in case of death or by reaching the time horizon of 10 years. In all other cases, the subject re-enters the simulation at a point defined by the previous state. The re-entering points are indicated in Figure 1 and Figure 2. The end point of a 1-year cycle is the starting point for the next cycle.

An individual enters the simulation either with AF or without AF. The initial health state is defined by the prevalence of AF. The following path is determined by the incidence of the alternatives at each decision node. Cardioversion through surgical interventions (eg, catheter ablation) to restore normal sinus rhythm was excluded. Thus, it was assumed that once an individual experiences AF, it cannot be cured. With AF, an individual cannot leave the upper branch (Figure 1 and Figure 2, “Atrial Fibrillation”) of the decision tree.

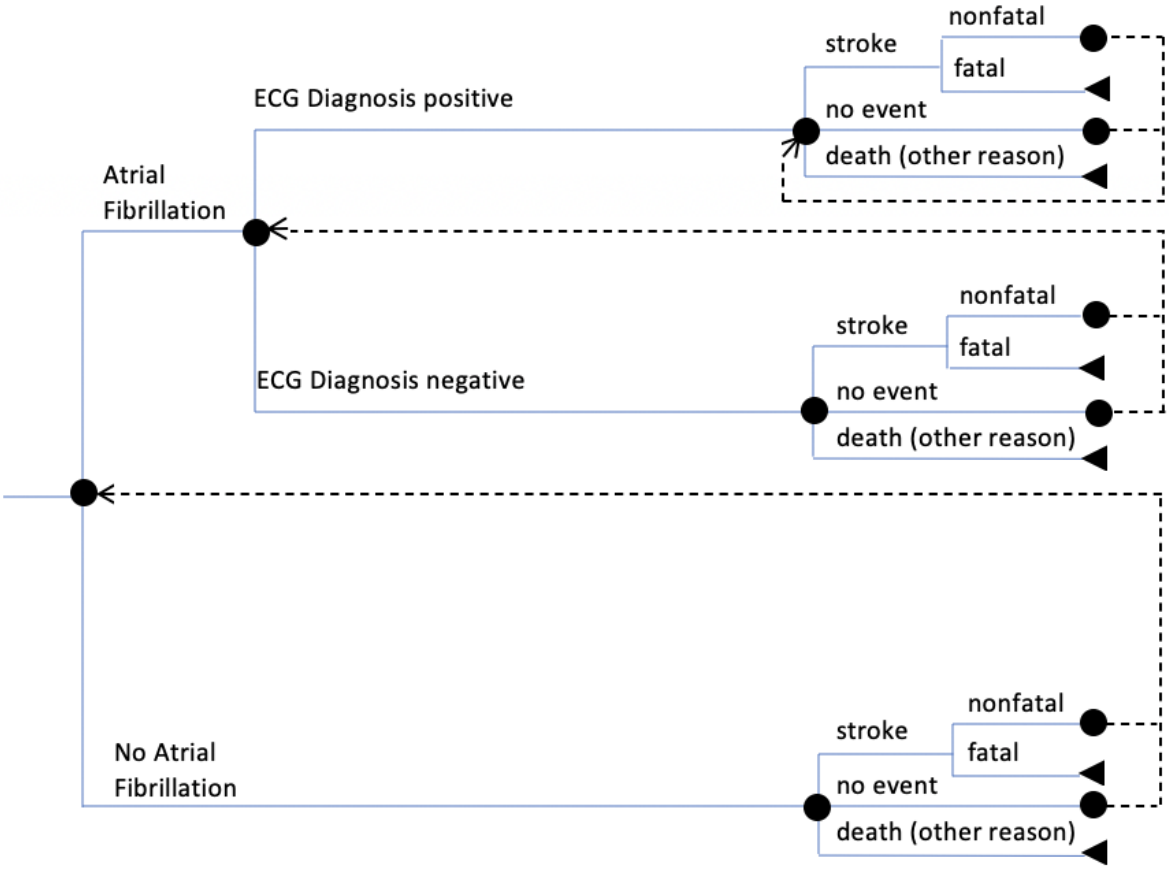
XSL•FO
RenderX

Figure 1. Model structure of the group with mobile health devices (each end point is a different scenario). Additional bleeding events can occur in each end point. ECG: electrocardiography.



XSL•FO
RenderX

Figure 2. Model structure of the group without mobile health devices (each end point is a different scenario). Additional bleeding events can occur in each end point. ECG: electrocardiography.



XSL•FO
RenderX

Table 1. Probabilities of annual state transition as well as underlying assumptions and sources.

Serial number	Model item	Assumptions	Sources and description
1	Prevalence of AF[a] at baseline	Based on the CHA_2DS_2-VASc[b] score: 0.01, 0.015, 0.034, 0.067, 0.118, 0.182, 0.255, 0.302, 0.403, 0.492	Derived from the study of Saliba et al [7]. The prevalence was used to simulate the initial proportion of patients with AF.
2	Incidence of AF in the general population (per 100 person-years)	Based on the CHA_2DS_2-VASc score: 0.17, 0.21, 0.49, 0.94, 1.65, 2.31, 2.75, 3.39, 4.09, 6.71	Derived from the study of Saliba et al [7]. The incidence was used to estimate the number of new cases of AF each year.
3	Sensitivity of mHealth[c] devices	93%[d]	Derived from the study of Bonomi et al [8]. Sensitivity of PPG compared to 24/48-hour Holter electrocardiogram readings in outpatient settings; 93 out of 100 patients with AF receive a true-positive diagnosis.
4	False-positive AF detection rate (mHealth device)	0.2%[d]	Bonomi et al [8] described the false-positive detection rate as lower than 0.2%; 0.2% of subjects without AF receive a false-positive diagnosis.
5	Confirmation of the mHealth diagnosis (by a physician using ECG[e])	100%, 75%, and 50%	Because of the nonpersistent forms of AF, the disease cannot always be confirmed through ECG follow-up. Nevertheless, in the first step, the assumption was made that a true-positive mHealth diagnosis of AF can always be confirmed by a physician. In subsequent simulations, the proportion was altered.
6	Clarification of a wrong mHealth diagnosis (by a physician using ECG)	100%	Assumption that in patients with no AF, the attending physician will not find artefacts of arrhythmia in the electrocardiogram.
7	Proportion of AF detected without a device	36.09%	Steinhubl et al [9] investigated the detection rate of AF in active home-based monitored individuals. They found newly diagnosed AF in 6.7 per 100 person-years in the monitored individuals and 2.6 per 100 person-years in unmonitored individuals. The proportion of AF detected with the aid of wearables was multiplied with the AF ratio between unmonitored and monitored individuals. Yearly, 36.09% of AF cases can be detected without the use of mHealth devices.
8	Stroke incidence in untreated patients with AF (per 100 person-years)	Based on the CHA_2DS_2-VASc score: 0.2, 0.6, 2.5, 3.7, 5.5, 8.4, 11.4, 13.1, 12.6, 14.44	Derived from the study of Friberg et al [10]. The stroke incidence yields the probability of experiencing a stroke.
9	Stroke incidence in patients with no AF (per 100 person-years)	Based on the CHA_2DS_2-VASc score: 0.0826, 0.2479, 1.0331, 1.5289, 2.2727, 3.4711, 4.7107, 5.4132, 5.2066, 5.9669	According to Odutayo et al [2], patients with AF have a 2.42-fold increased risk for stroke compared to patients with no AF. The stroke incidence in untreated patients with AF was divided by 2.42.
10	Stroke incidence in patients with AF receiving NOAC[f] (per 100 person-years)	Based on the CHA_2DS_2-VASc score: 0.068, 0.204, 0.85, 1.258, 1.87, 2.856, 3.876, 4.454, 4.284, 4.9096	VKA[g] reduces the risk of stroke by two-third (66%) [1]. Rivaroxaban is noninferior to warfarin [11]. Thus, the risk reduction through NOAC should be at least as high as the one from VKA.
11	Stroke mortality in patients with no AF	34%	Derived from the study of Reinhold et al [6]. If a patient does not have AF but experiences a stroke, there is a 34% probability that the stroke is fatal.
12	Stroke mortality in untreated patients with AF	63%	Derived from the study of Reinhold et al [6]. If a patient has AF and does not receive medication, there is a 63% probability that the stroke is fatal.
13	Stroke mortality in patients with AF receiving NOAC	42%	Derived from the study of Reinhold et al [6]. If a patient has AF and receives medication, the probability that an occurring stroke is fatal is 42%.
14	Mortality in patients with no AF, no stroke	6%	Derived from the study of Reinhold et al [6]. Probability that an individual who does not have AF dies due to reasons other than stroke.
15	Mortality in untreated patients with AF, no stroke	11.1%	Derived from the study of Reinhold et al [6]. The probability that an untreated patient with AF dies due to reasons other than stroke.

XSL•FO
RenderX

[a]AF: atrial fibrillation.

[b]CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female).

[c]mHealth: mobile health.

[d]values were changed in sensitivity analysis.

[e]ECG: electrocardiography.

[f]NOAC: non–vitamin K antagonist.

[g]VKA: vitamin K antagonist.

Once the individual health state is set and the underlying individual is part of the group with mHealth devices, there is a given probability of a device-based diagnosis (either true-positive diagnosis or false-positive diagnosis). If the device-based diagnosis is positive, the patient visits a doctor and an ECG is recorded. If the mHealth diagnosis was false positive, the doctor will clear up the misdiagnosis and the individual is considered as healthy and remains in the group without AF. If the individual truly has AF and the mHealth device–based diagnosis is positive, the diagnosis might be confirmed by the doctor. Either way, if the diagnosis is confirmed or not, the patient remains in the AF group (Figure 1). Therefore, different probabilities were assumed (Table 1, Serial number 5).

In case the device misses a diagnosis of AF or the individual is in the group without mHealth devices, there is a chance that AF is diagnosed during a visit to the physician in terms of standard care (Table 1, Serial number 7). Furthermore, it is supposed that a stroke in patients with previously undetected AF leads to an AF diagnosis and therapy as well.

Once an individual receives an ECG-driven diagnosis of AF, it is valid for the rest of the simulation and the possible states are restricted according to the state-transition model. Based on the diagnosis, it is assumed that the patient receives anticoagulation therapy in the form of non–vitamin K antagonists (NOAC).

The possible end points at the end of each cycle are identical, irrespective of the preceding arms of the decision tree. The first possible end point could be experiencing a stroke, which can be either fatal or nonfatal. The second possible end point could be that the individual does not face any event influencing the simulation. The third end point could be that the patient can die due to reasons other than stroke. In all the end points, additional bleeding events can occur.

State Transition Probabilities

The underlying probabilities for state transition are depicted in Table 1. The transition possibilities differ for the implemented CHA_2DS_2-VASc score. Increasing scores correlate with higher prevalence and incidence of AF as well as higher risk for stroke. The initiation of NOAC reduces the risk of stroke and mortality in patients with AF; however, it increases the risk for major bleeding. To assess the accuracy of mHealth devices in screening for AF, a study focusing on the use of PPG was used [8]. PPG is one of the most widespread technologies in mHealth devices to screen for AF.

Costs

AF-related direct costs were considered from the view of the German statutory health insurance. Device costs, costs incurred during a visit to the doctor, costs incurred in diagnostics, costs incurred in the therapy in form of NOAC, as well as costs related to stroke and major bleeding were integrated. Device costs were derived from the most popular mHealth AF screening device, the Apple Watch 5 (€437.65, €1=US $1.12) [12].

To confirm the mHealth device–based diagnosis by a physician, the costs were represented by adding single cost factors incurred during the physician visit (ordination, consultation, urgent care, telephone advice, telemedical care) (€35.62) with cost factors resulting from diagnostics (long-term ECG, 12-lead ECG, stress ECG) (€31.61) [13,14]. The cost components were derived from [13] but the costs were adapted to the year of the study. As medication costs for oral anticoagulation, the use of rivaroxaban as the most prescribed NOAC in Germany was assumed. Thus, the costs for pharmaceuticals resulted in €1226 per year [15]. Costs for individuals with stroke, either fatal or not, were derived from the study of Kolominsky-Rabas et al [16]. An interpolation and an extrapolation were made to receive period-specific costs (Figure 3 and Table 2). The costs for major bleeding (€1995) were directly derived from the study of Reinhold et al [6]. The present value was calculated using a discount rate of 3% per year.

XSL•FO
RenderX

Figure 3. Interpolation and extrapolation of costs determined by using least squares adjustment. Values for year 1, year 5, and year 10 derived from Kolominsky-Rabas et al [16]. €1=US $1.12.

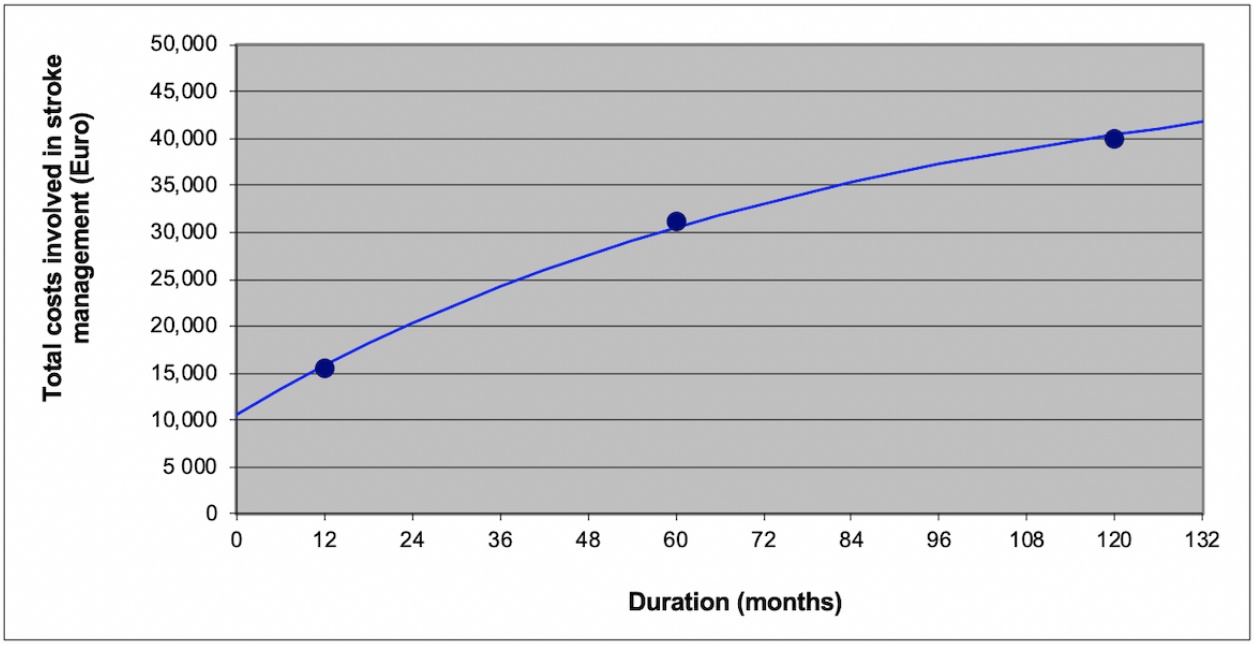


Table 2. Relevant cost factors as well as sources and descriptions.

Cost factor	Costs[a]	Reasons and description
Device costs	€437.65	The price was derived from the most popular PPG[b] AF[c] screening device, the Apple Watch Series 5 [12]. An integrated algorithm diagnoses AF automatically. Trained personnel for interpretation is not needed.
Visit to the doctor and diagnostics	€67.23	Physician visit: ordination and consultation, €13.20; urgent care, €12.90; and telemedical care, €9.52. Diagnostics: long-term ECG and 12-lead ECG, €9.96; stress ECG, €21.65; derived from the study of McBride et al [13] and adapted to current conditions [14].
Medication costs for oral anticoagulation (NOAC[d])	€1226	The use of rivaroxaban was assumed because it is the most prescribed NOAC in Germany [15].[e]
Per year costs incurred after surviving a stroke	€15,753 (year 1), €4480 (year 2)… €1481 (year 10)	Interpolation and extrapolation of costs derived from the study of Kolominsky-Rabas et al [16] (Figure 3).[e]
Costs for major bleeding	€1995	Directly derived from the study of Reinhold et al [6].
Annual discounting rate	3%	Own assumption.

[a]€1=US $1.12.

[b]PPG: photoplethysmography.

[c]AF: atrial fibrillation.

[d]NOAC: non–vitamin K antagonist.

[e]The program was realized using unrounded amounts in Euro.

Implementation

As relevant outcomes costs, prevented strokes and prevented deaths from stroke were defined. To receive these outcomes, an implementation of the simulation was conducted in Excel (Microsoft Corp) by using Visual Basic for Applications.

Four different scenarios were simulated for each CHA_2DS_2-VASc score from 1 to 9: 3 scenarios with mHealth devices but different ECG confirmation rates (100%, 75%. and 50%) (Table 1, Serial number 5) and 1 scenario for patients without mHealth devices. Each simulation included 30,000 fictitious patients. Subsequently, a sensitivity analysis for device sensitivity and false-positive AF detection rate was conducted. According to the European Society of Cardiology Guidelines for the management of AF, it was assumed that anticoagulation therapy was initiated in male patients with a CHA_2DS_2-VASc score of 1 and in female patients with a score of 2 [1]. Therefore, a comparison in patients with a risk score of 0 was deemed as dispensable. To estimate the difference in the patients with a risk score of 1, it was assumed that half of the individuals were females. This is in accordance with the distributions of the sexes in the publications used to determine the prevalence and

XSL•FO
RenderX

incidence of AF [7] as well as the stroke incidence [10] used in the simulation.

Results

Costs

The economic effect of mHealth intervention was assessed in 2 steps. First, the focus was on costs per patient. Secondly, costs were assessed in relation to prevented strokes and fatal strokes. As seen in Table 3 and Table 4, an increasing risk score has a major impact on costs per patient in all the groups. The higher the CHA_2DS_2-VASc score, the higher are the costs per patient on average. While device ECG confirmation rate has little impact on costs per patient, the use of mHealth devices increases the costs per patient clearly (Figure 4).

To assess the costs per prevented stroke, the groups with and without mHealth devices were compared. The difference in the sum of the costs for all the patients in each group as well as the difference in the number of strokes were determined for each risk score. The ratio between the difference of the sum of costs and the difference in number of strokes resulted in costs per prevented stroke (Table 5).

Although costs per patient increase with increasing CHA_2DS_2-VASc scores, the costs per stroke tend to decrease in general. This effect is intensified by an increasing ECG confirmation rate. The effect is not seen in groups with lower risk scores. In these groups, the underlying basic risk for stroke is low. Thus, the risk reduction by use of mHealth devices is low as well. Findings for costs per fatal stroke fluctuated more than costs per patient and the number of fatal strokes. This can be explained by a small denominator (number of prevented [fatal] strokes) in relation to a large numerator (cost difference for all patients). Thus, small changes in the number of prevented (fatal) strokes have a big impact on costs per prevented (fatal) stroke.

With increasing ECG confirmation rates, the effect of mHealth use becomes more evident. Low ECG confirmation rates lead to results mainly driven by chance. In particular, regarding the costs per prevented fatal stroke, the impact of higher risk scores as well as ECG confirmation rates is even more pronounced.

Table 3. Summarized results of the simulations. Costs, strokes, and fatal strokes classified on the basis of the CHA_2DS_2-VASc score as well as the investigated group (N=30,000 patients per group per score).

	Study arm without device				Study arm with device (50% ECG[a] confirmation)			
CHA_2DS_2-VASc score[b]	Average costs per patient (in €[c], whole simulation duration)	Total number of strokes[d]	Number of nonfatal strokes	Number of fatal strokes	Average costs per patient (in €, whole simulation duration)	Total number of strokes[d]	Number of nonfatal strokes	Number of fatal strokes
1	873	581	379	202	1330	599	402	197
2	2280	2338	1571	767	2788	2351	1513	838
3	3351	3493	2283	1210	3815	3460	2232	1228
4	4860	5260	3288	1972	5239	4903	3100	1803
5	6877	7808	4844	2964	7233	7437	4569	2858
6	8802	10,397	6286	4111	9375	10,163	6228	3935
7	10,023	11,804	7024	4780	10,414	11,237	6857	4380
8	10,154	11,485	6591	4894	10,761	11,039	6469	4570
9	11,299	12,565	6944	5621	12,086	12,201	6964	5237
mean	6502	7303	4357	2947	7005	7043	4259	2784

[a]ECG: electrocardiography.

[b]CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female).

[c]€1=US $1.12.

[d]Total number of strokes includes nonfatal and fatal strokes.

XSL•FO
RenderX

Table 4. Summarized results of the simulations. Costs, strokes, and fatal strokes classified on the basis of the CHA_2DS_2-VASc score as well as the investigated group (N=30,000 patients per group per score).

	Study arm with device (75% ECG[a] confirmation)				Study arm with device (100% ECG confirmation)			
CHA_2DS_2-VASc score[b]	Average costs per patient (in €[c], whole simulation duration)	Total number of strokes[d]	Number of nonfatal strokes	Number of fatal strokes	Average costs per patient (in €, whole simulation duration)	Total number of strokes[d]	Number of nonfatal strokes	Number of fatal strokes
1	1314	556	361	195	1290	528	331	197
2	2847	2406	1585	821	2876	2364	1550	814
3	3887	3395	2180	1215	3876	3339	2180	1159
4	5380	4994	3204	1790	5421	4894	3154	1740
5	7483	7444	4751	2693	7543	7263	4700	2563
6	9427	9755	6084	3671	9508	9549	6107	3442
7	10,646	11,082	6901	4181	10,627	10,703	6771	3932
8	10,846	10,374	6295	4079	10,937	10,122	6301	3821
9	12,146	11,449	6756	4693	12,463	11,210	6897	4313
mean	7108	6828	4235	2593	7171	6664	4221	2442

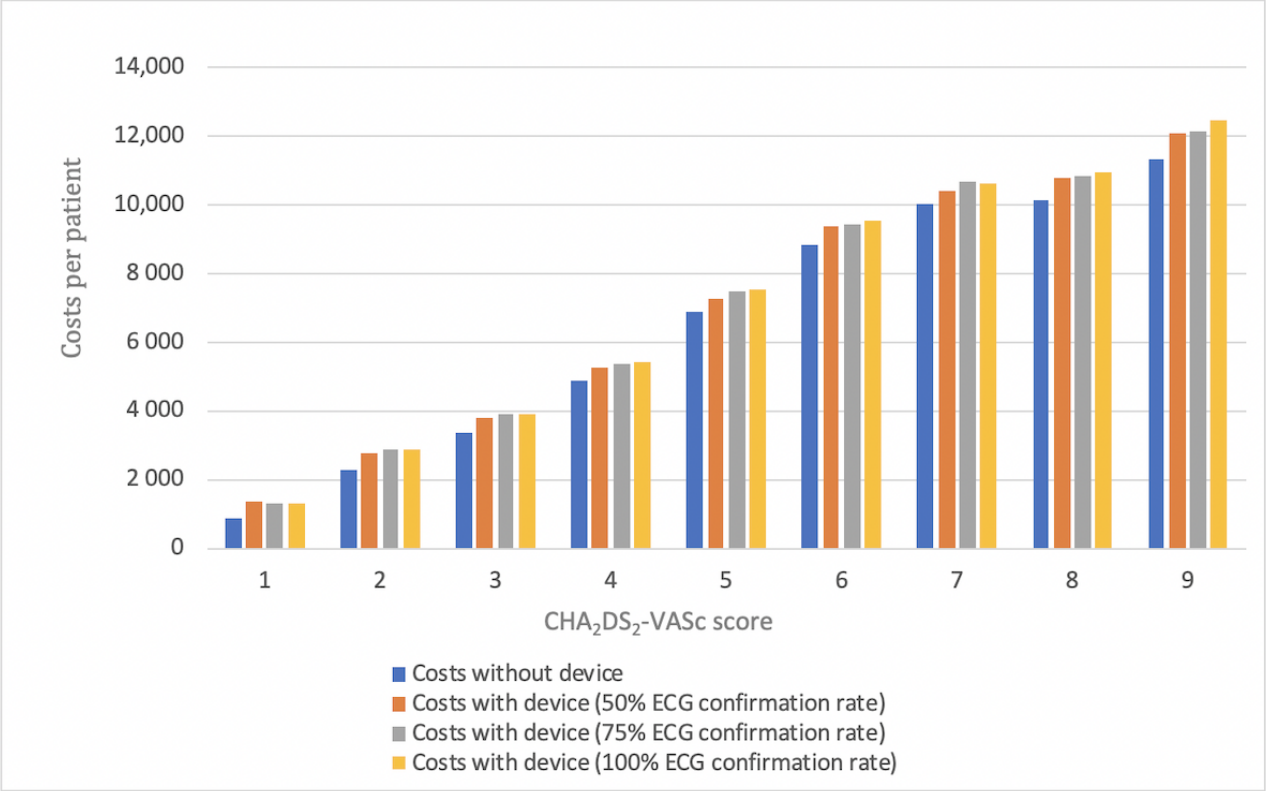
[a]ECG: electrocardiography.

[b]CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female).

[c]€1=US $1.12.

[d]Total number of strokes includes nonfatal and fatal strokes.

Figure 4. Costs per patient classified on the basis of the CHA_2DS_2-VASc score as well as the investigated group (with or without device and ECG confirmation rate). ECG: electrocardiography; CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female). €1=US $1.12.



XSL•FO
RenderX

Table 5. Number of prevented strokes and costs per prevented stroke in each intervention group.

Study arm, CHA_2DS_2-VASc score[a]	Cost difference for all patients[b] (in €[c])	Prevented strokes	Costs per prevented stroke (in €)	Prevented fatal strokes	Costs per prevented fatal stroke (in €)
Study arm with device (100% ECG[d] confirmation)					
1	12,519,300	53	236,213	5	2,503,860
2	17,893,200	–26	–688,200	–47	–380,706
3	15,759,300	154	102,333	51	309,006
4	16,852,500	366	46,045	232	72,640
5	19,992,600	545	36,684	401	49,857
6	21,174,300	848	24,970	669	31,651
7	18,103,800	1101	16,443	848	21,349
8	23,481,300	1363	17,228	1073	21,884
9	34,921,800	1355	25,773	1308	26,699
Study arm with device (75% ECG confirmation)					
1	13,228,200	25	529,128	7	1,889,743
2	17,028,300	–68	–250,416	–54	–315,339
3	16,074,000	98	164,020	–5	–3,214,800
4	15,609,900	266	58,684	182	85,769
5	18,181,800	364	49,950	271	67,092
6	18,732,600	642	29,179	440	42,574
7	18,676,800	722	25,868	599	31,180
8	20,762,700	1111	18,688	815	25,476
9	25,423,200	1116	22,781	928	27,396
Study arm with device (50% ECG confirmation)					
1	13,704,000	–18	–761,333	5	2,740,800
2	15,242,700	–13	–1,172,515	–71	–214,686
3	13,933,500	33	422,227	–18	–774,083
4	11,367,300	357	31,841	169	67,262
5	10,708,500	371	28,864	96	111,547
6	17,187,900	234	73,453	176	97,659
7	11,712,000	567	20,656	400	29,280
8	18,208,800	446	40,827	324	56,200
9	23,614,500	364	64,875	384	61,496

[a]CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female).

[b]Cost difference between group with devices and group without devices.

[c]€1=US $1.12.

[d]ECG: electrocardiography.

Prevented Strokes

With respect to patients, prevented strokes are considered as the most relevant outcome in this Monte Carlo simulation. Prevented strokes were analyzed as prevented strokes in total on the one hand and as prevented fatal strokes on the other hand. Both of them were calculated as the difference between the number of (fatal) strokes in the group without devices and the number of (fatal) strokes in each of the groups with devices (Table 5, Figure 5, and Figure 6).

XSL•FO
RenderX

Figure 5. Stroke analysis on the basis of the CHA_2DS_2-VASc score as well as the investigated group (with or without device and ECG confirmation rate). ECG: electrocardiography; CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female).

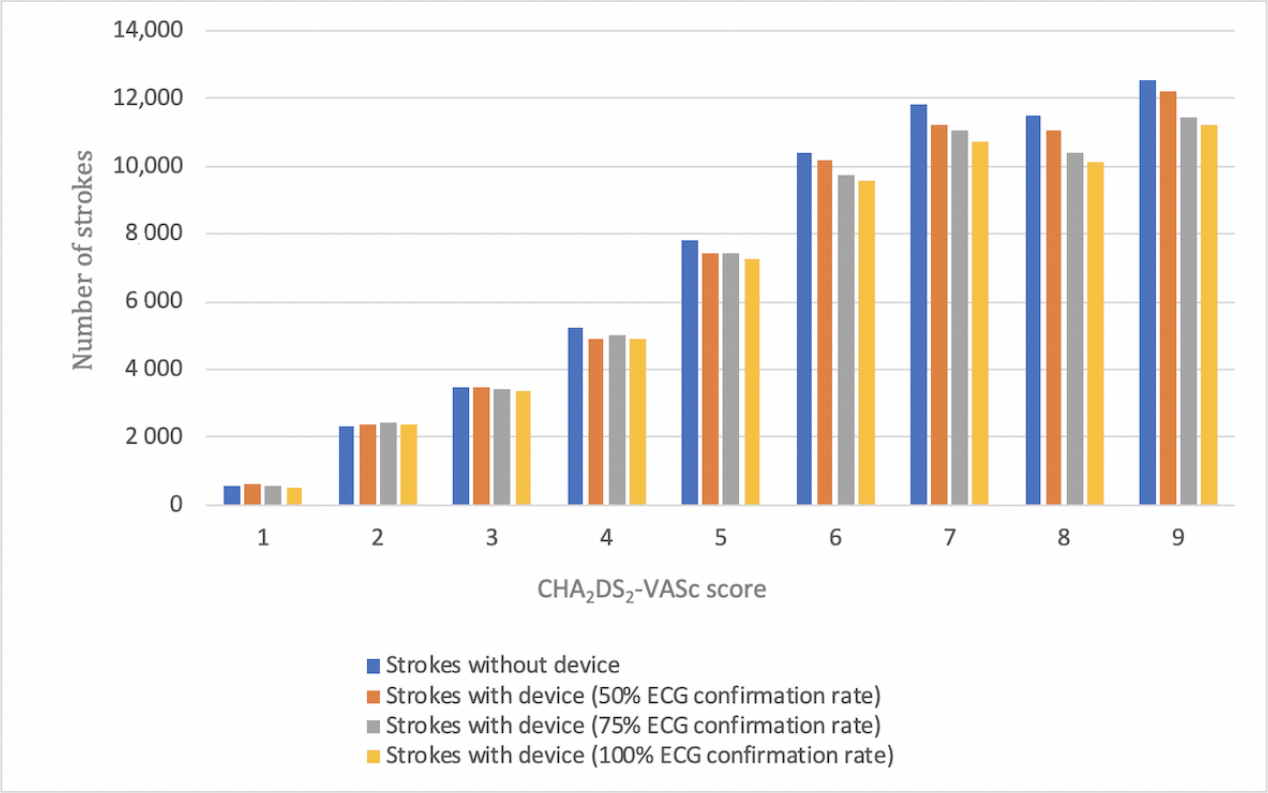
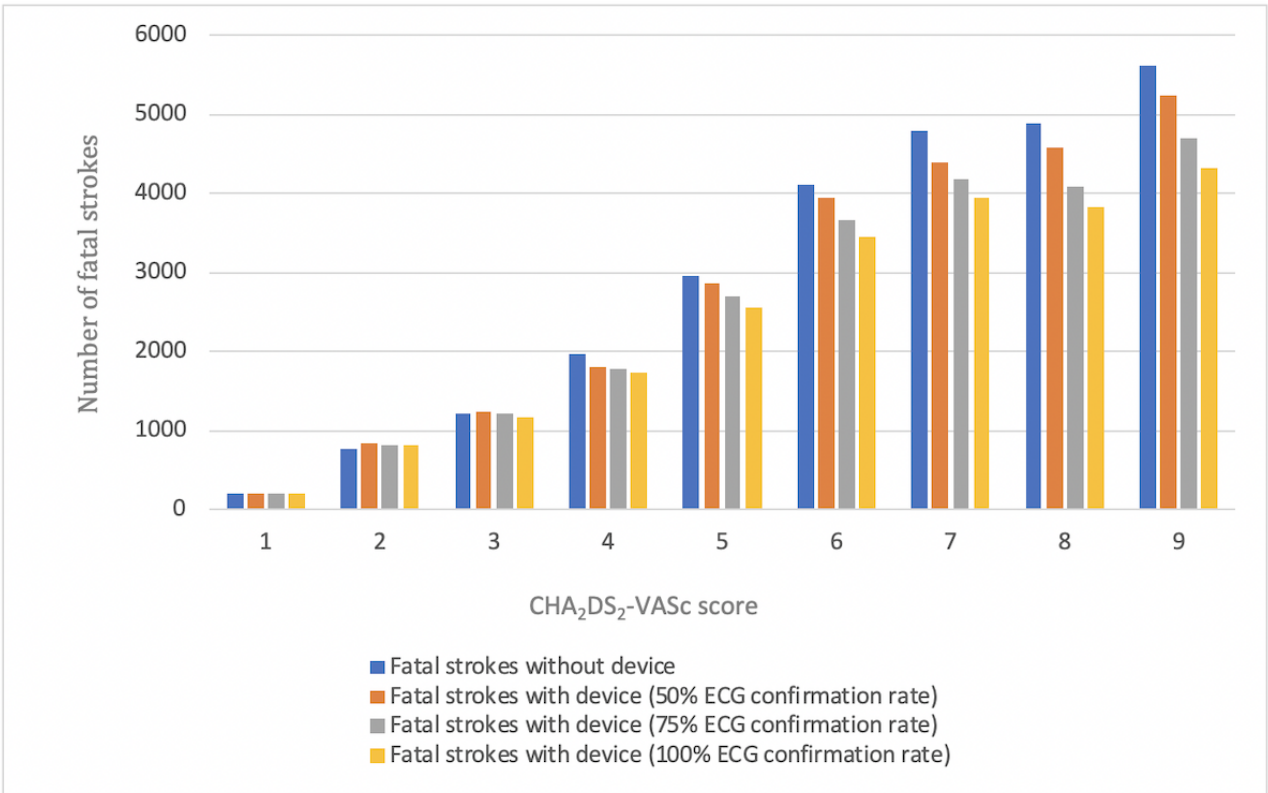


Figure 6. Fatal stroke analysis on the basis of the CHA_2DS_2-VASc score as well as the investigated group (with or without device and ECG confirmation rate). ECG: electrocardiography; CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female).



XSL•FO
RenderX

The chance to prevent strokes by the use of mHealth devices is mainly driven by 2 factors. First, as seen in Table 5, the incidence of prevented strokes tends to increase with increasing CHA_2DS_2-VASc scores. The higher the risk score, the higher is the incidence of AF. More patients with AF provide a higher chance to detect AF by using mHealth devices, and thus, initiated anticoagulation therapy will most likely reduce the number of strokes. However, here again, this effect is not seen in groups with a low risk for stroke. The second factor influencing the number of prevented strokes is the ECG confirmation rate. The higher the predictive value of the device is, the more number of cases of AF can be confirmed, and the more strokes might be prevented. If the device diagnosis is more reliable, more cases of AF can be detected and the risk for stroke can be reduced by subsequent therapy. The effects of higher risk scores and high device ECG confirmation rate are even higher in prevented fatal strokes. Nevertheless, there is also no clear effect in low-risk patient groups.

Sensitivity Analysis

Based on the simulation, a sensitivity analysis was conducted for values of device sensitivity (86%, 93%, and 100%) as well as device false-positive AF detection rate (0.2%, 1%, and 5%) (Table 6 and Table 7). For sensitivity analysis, the confirmation of the mHealth diagnosis was determined to be 75%.

Device accuracy in terms of device sensitivity and device false-positive rate had little impact on the costs per patient but it had big impact on the number of fatal strokes. A higher device sensitivity leads to a higher number of prevented fatal strokes. In terms of the device false-positive rate, a higher value had little impact on costs per patient and the number of strokes. Nevertheless, it should be considered that high false device–positive rates frighten patients and lead to more frequent physician-patient interactions, which are a burden for the health care system.

Table 6. Sensitivity analysis. The values were changed to 86% and 100%; 93% was the standard case.

	Device sensitivity								
	86%			93%[a]			100%		
CHA_2DS_2-VASc score[b]	Average costs per patient (in €)[c]	Total number of strokes	Number of fatal strokes	Average costs per patient (in €)	Total number of strokes	Number of fatal strokes	Average costs per patient (in €)	Total number of strokes	Number of fatal strokes
1	1275	515	175	1308	558	210	1326	586	209
2	2794	2362	868	2847	2406	821	2816	2312	792
3	3908	3458	1239	3887	3395	1215	3912	3446	1234
4	5445	5111	1861	5380	4994	1790	5456	4986	1728
5	7504	7450	2742	7483	7444	2693	7466	7329	2695
6	9498	9878	3704	9427	9755	3671	9542	9830	3600
7	10,430	10,902	4219	10,646	11,082	4181	10,537	10,814	4106
8	10,831	10,468	4081	10,846	10,374	4079	10,922	10,447	4109
9	12,279	11,696	4715	12,146	11,449	4693	12,129	11,351	4602

[a]Base value.

[b]CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female).

[c]€1=US $1.12.

XSL•FO
RenderX

Table 7. Sensitivity analysis. Values altered for device false-positive atrial fibrillation detection rates.

	Device false-positive rate								
	0.2%[a]			1%			5%		
CHA_2DS_2-VASc score[b]	Average costs per patient (in €)[c]	Total number of strokes	Number of fatal strokes	Average costs per patient (in €)	Total number of strokes	Number of fatal strokes	Average costs per patient (in €)	Total number of strokes	Number of fatal strokes
1	1308	558	210	1336	584	207	1342	579	207
2	2847	2406	821	2863	2395	820	2835	2352	789
3	3887	3395	1215	3858	3425	1187	3864	3405	1198
4	5380	4994	1790	5414	5019	1803	5365	4961	1767
5	7483	7444	2693	7526	7452	2735	7447	7254	2626
6	9427	9755	3671	9537	9851	3693	9561	9833	3676
7	10,646	11,082	4181	10,594	10,931	4073	10,650	11,127	4264
8	10,846	10,374	4079	10,923	10,557	4112	10,772	10,305	4026
9	12,146	11,449	4693	12,076	11,373	4691	12,254	11,599	4631

[a]Base value.

[b]CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female).

[c]€1=US $1.12.

Discussion

Besides wrist-worn devices, ECG patches, hand-held devices, and apps provide a helpful method to screen for AF [17]. Recent cost-effectiveness analyses of hand-held ECG recorders showed that these devices are likely to be cost-effective in older patient groups [18-20]. Jacobs et al [18] investigated the effect of AF screening with mHealth devices during seasonal influenza vaccination; they found the screening to be cost-effective. A second cost-effectiveness analysis conducted by Aronsson et al [19] showed that 2 weeks of intermittent screening for asymptomatic AF resulted in costs of €4313 per gained quality-adjusted life-year and €6583 per avoided stroke [19]. Levin et al found that screening for silent AF after ischemic stroke in 75-year-old patients leads to decreased costs, extended lives, and improved quality of life [20]. The cost-effectiveness of wrist-worn mHealth devices to detect AF is not yet clarified [17].

The present model is the first to estimate the cost-effectiveness of mHealth interventions by using wrist-worn devices over a long period and assessing the cost-effectiveness of mHealth devices in relation to the CHA_2DS_2-VASc score. To assess the health economic effect of mHealth devices, several assumptions and simplifications were integrated in the model. Some costs were excluded. First, in the underlying simulation, indirect costs associated with strokes were not considered. Indirect costs include costs for work loss. Work loss was not considered because no eligible current analysis about those specific costs could have been found. Furthermore, indirect costs incurred by work absences are presumed to be relatively low because strokes mainly occur in older patients who are not working anymore. Second, this simulation was limited to a time period of 10 years. Long-term costs of care and medication were restricted in accordance with the model.

Mean cost values for a visit to the doctor included ordination, consultation, urgent care, telemedical care as well as different types of ECG. Other possible interventions such as international normalized ratio blood test, ultrasound, and radiography [13] were not considered. There were no eligible data for long-term patient care. Thus, subsequent visits were not integrated.

It was implemented that patients with AF receive rivaroxaban because it is the most prescribed NOAC in Germany. Besides rivaroxaban, there are many other pharmaceutical products such as apixaban, dabigatran, warfarin, and phenprocoumon for the treatment of AF. Some patients are not eligible for treatment with NOACs and should take oral anticoagulants in form of vitamin K antagonists (VKAs). Exclusion criteria are, for example, use of mechanical heart valves or moderate as well as severe mitral stenosis [1]. Since the most prescribed VKA in Germany (phenprocoumon: €54.75 per year) is cheaper than rivaroxaban (€1226.40 per year) [15], the estimates in this study are even more conservative. In other studies, the costs for anticoagulation therapy were estimated to be lower. Jacobs et al [18] estimated the costs for NOAC to be €235 in the Netherlands. Aronsson et al [19] suggested the use of apixaban, which resulted in costs of €844 in Sweden.

This simulation is based on published data. However, this published data did not represent a consistent patient pool. Therefore, a special focus was put on the patient characteristics in the underlying studies. The proportion of male and female patients was always near 50%. Patient age as well as other relevant characteristics were represented consistently by the CHA_2DS_2-VASc score. A weakness of the simulation was that general mortality in healthy subjects was assumed to be 6%, irrespective of their age.

XSL•FO
RenderX

The stroke incidence in patients with no AF was determined by a division; the stroke incidence of untreated patients with AF was divided by their additional risks for stroke compared to patients with no AF. The most popular study on AF-related stroke risk, the Framingham Study, estimates that the additional risk for stroke in untreated patients with AF compared to that in patients with no AF is 4.8-fold [3]. In this study, this risk was determined to be 2.42-fold according to a meta-analysis by Odutayo et al [2].

The Apple Heart Study showed that only 57% of the patients went to the doctor after receiving an irregular pulse notification [21]. In this simulation, it was modelled that every individual who receives a notification visits the doctor. According to the results, fewer visits to the doctor are related to lower overall costs as well as fewer prevented strokes.

A further problem was to assess the accuracy of the mHealth devices. The assumed accuracy published by Bonomi et al [8] could be overestimated because physical activity, darker skin color, higher body mass index, or male gender may influence the accuracy [22]. With respect to newer devices such as the Apple Watch, more cases of AF can be diagnosed with the aid of ECG recordings in addition to PPG technology. To derive the ratio of AF detected between the groups with and without a device, the findings of a study by Steinhubl et al were used [9]. They investigated the effect of a home-based wearable intervention to detect AF by using ECG patches over a period of 4 weeks. Although Steinhubl et al [9] used ECG patches for a shorter period, their results were integrated in the simulation. Tischer et al [23] found that patients with high CHA_2DS_2-VASc scores experienced thromboembolic complications, irrespective of the presence of AF. In these patients, anticoagulation therapy may be initiated, regardless of AF. Thus, particularly in the group with devices, for higher scores, the costs of the prescribed NOACs could be overestimated because some patients would receive anticoagulation therapy, irrespective of AF.

In conclusion, the results of this simulation allow the assessment of the use of mHealth devices in different risk groups. From an economic point of view, the use of these devices in patients with high risk scores increases the costs per patient. With higher risk scores, costs per prevented stroke decrease. Higher device accuracy leads to more stable results. From a patient-oriented perspective, the use of mHealth devices results in reduced number of strokes. More strokes can be prevented if the underlying CHA_2DS_2-VASc score is higher. In addition, a high ECG confirmation rate and increased device accuracy lead to more prevented strokes.

This study shows that mHealth devices are a recommendable tool to screen for AF in patients with high CHA_2DS_2-VASc scores. The higher the risk for stroke in patients with AF, the more cost-effective are the devices.

Acknowledgments

I thank Prof. Dr. Dierk Brockmeier for his valuable advice regarding the set-up of the model as well as the interpolation and extrapolation of costs associated with stroke management. I thank Prof. Dr. Adriaan Dorresteijn for his support, which made the publication possible. The publication fee was funded by the Ministry of Science, Research and Art Baden-Württemberg.

Conflicts of Interest

None declared.

References

1. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, ESC Scientific Document Group. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 2016 Oct 07;37(38):2893-2962. [doi: 10.1093/eurheartj/ehw210] [Medline: 27567408]
2. Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. BMJ 2016 Sep 06;354:i4482 [FREE Full text] [doi: 10.1136/bmj.i4482] [Medline: 27599725]
3. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991 Aug;22(8):983-988. [doi: 10.1161/01.str.22.8.983] [Medline: 1866765]
4. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 2010 Feb;137(2):263-272. [doi: 10.1378/chest.09-1584] [Medline: 19762550]
5. Sonnenberg FA, Beck JR. Markov Models in Medical Decision Making. Med Decis Making 2016 Jul 02;13(4):322-338. [doi: 10.1177/0272989x9301300409]
6. Reinhold T, Belke R, Hauser T, Grebmer C, Lennerz C, Semmler V, et al. Cost Saving Potential of an Early Detection of Atrial Fibrillation in Patients after ICD Implantation. Biomed Res Int 2018;2018:3417643 [FREE Full text] [doi: 10.1155/2018/3417643] [Medline: 30186856]
7. Saliba W, Gronich N, Barnett-Griness O, Rennert G. Usefulness of CHADS2 and CHA2DS2-VASc Scores in the Prediction of New-Onset Atrial Fibrillation: A Population-Based Study. Am J Med 2016 Aug;129(8):843-849. [doi: 10.1016/j.amjmed.2016.02.029] [Medline: 27012854]

XSL•FO
RenderX

8. Bonomi AG, Schipper F, Eerikäinen LM, Margarito J, van Dinther R, Muesch G, et al. Atrial Fibrillation Detection Using a Novel Cardiac Ambulatory Monitor Based on Photo - Plethysmography at the Wrist. J Am Heart Assoc 2018 Aug 07;7(15). [doi: 10.1161/jaha.118.009351]
9. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a Home-Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mSToPS Randomized Clinical Trial. JAMA 2018 Jul 10;320(2):146-155 [FREE Full text] [doi: 10.1001/jama.2018.8102] [Medline: 29998336]
10. Friberg L, Rosenqvist M, Lip G. Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. Eur Heart J 2012 Jun;33(12):1500-1510. [doi: 10.1093/eurheartj/ehr488] [Medline: 22246443]
11. Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med 2011 Sep 08;365(10):883-891. [doi: 10.1056/NEJMoa1009638] [Medline: 21830957]
12. Apple. Apple Watch. 2020. URL: https://www.apple.com/de/watch/ [accessed 2020-08-06]
13. McBride D, Mattenklotz AM, Willich SN, Brüggenjürgen B. The costs of care in atrial fibrillation and the effect of treatment modalities in Germany. Value Health 2009;12(2):293-301 [FREE Full text] [doi: 10.1111/j.1524-4733.2008.00416.x] [Medline: 18657103]
14. KBV Kassenärztliche Bundesvereinigung. Einheitlicher Bewertungsmaßstab (EBM). 2019. URL: https://www.kbv.de/media/sp/EBM_Gesamt_-_Stand_4._Quartal_2019.pdf [accessed 2020-09-29]
15. Hein L, Wille H. Antithrombotika und Antihämorrhagika. In: Schwabe U, Paffrath D, Ludwig WD, Klauber J, editors. Arzneiverordnungs-Report 2018. Berlin, Heidelberg: Springer; 2018:401-421.
16. Kolominsky-Rabas PL, Heuschmann PU, Marschall D, Emmert M, Baltzer N, Neundörfer B, et al. Lifetime cost of ischemic stroke in Germany: results and national projections from a population-based stroke registry: the Erlangen Stroke Project. Stroke 2006 May;37(5):1179-1183. [doi: 10.1161/01.STR.0000217450.21310.90] [Medline: 16574918]
17. Giebel GD, Gissel C. Accuracy of mHealth Devices for Atrial Fibrillation Screening: Systematic Review. JMIR Mhealth Uhealth 2019 Jun 16;7(6):e13641 [FREE Full text] [doi: 10.2196/13641] [Medline: 31199337]
18. Jacobs MS, Kaasenbrood F, Postma MJ, van Hulst M, Tieleman RG. Cost-effectiveness of screening for atrial fibrillation in primary care with a handheld, single-lead electrocardiogram device in the Netherlands. Europace 2018 Jan 01;20(1):12-18. [doi: 10.1093/europace/euw285] [Medline: 27733465]
19. Aronsson M, Svennberg E, Rosenqvist M, Engdahl J, Al-Khalili F, Friberg L, et al. Cost-effectiveness of mass screening for untreated atrial fibrillation using intermittent ECG recording. Europace 2015 Jul;17(7):1023-1029. [doi: 10.1093/europace/euv083] [Medline: 25868469]
20. Levin L, Husberg M, Sobocinski PD, Kull VF, Friberg L, Rosenqvist M, et al. A cost-effectiveness analysis of screening for silent atrial fibrillation after ischaemic stroke. Europace 2015 Feb;17(2):207-214. [doi: 10.1093/europace/euu213] [Medline: 25349228]
21. Stanford Medicine. Apple Heart Study demonstrates ability of wearable technology to detect atrial fibrillation. 2019. URL: https://med.stanford.edu/news/all-news/2019/03/apple-heart-study-demonstrates-ability-of-wearable-technology.html [accessed 2020-09-29]
22. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. J Pers Med 2017 May 24;7(2) [FREE Full text] [doi: 10.3390/jpm7020003] [Medline: 28538708]
23. Tischer TS, Schneider R, Lauschke J, Nesselmann C, Klemm A, Diedrich D, et al. Prevalence of atrial fibrillation in patients with high CHADS2- and CHA2DS2VASc-scores: anticoagulate or monitor high-risk patients? Pacing Clin Electrophysiol 2014 Dec;37(12):1651-1657 [FREE Full text] [doi: 10.1111/pace.12470] [Medline: 25621351]

Abbreviations

AF: atrial fibrillation
CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65-74 years, Sex category (female)
ECG: electrocardiography
mHealth: mobile health
NOAC: non–vitamin K antagonist
PPG: photoplethysmography
VKA: vitamin K antagonist

XSL•FO
RenderX

Edited by G Eysenbach; submitted 20.05.20; peer-reviewed by D Bardey, H Heidbuchel; comments to author 23.07.20; revised version received 03.09.20; accepted 03.09.20; published 06.10.20.

Please cite as:
Giebel GD
Use of mHealth Devices to Screen for Atrial Fibrillation: Cost-Effectiveness Analysis
JMIR Mhealth Uhealth 2020;8(10):e20496
URL: http://mhealth.jmir.org/2020/10/e20496/
doi:10.2196/20496
PMID:33021489

©Godwin D Giebel. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 06.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Checklists for Complications During Systemic Cancer Treatment Shared by Patients, Friends, and Health Care Professionals: Prospective Interventional Cohort Study

Helen V Jones[1*], MD; Harry Smith[2], MD; Tim Cooksley[3], MD; Philippa Jones[4*]; Toby Woolley[1], MD; Derick Gwyn Murdoch[5*]; Dafydd Thomas[6]; Betty Foster[6]; Valerie Wakefield[6]; Pasquale Innominato[1,7,8*], MD, PhD; Anna Mullard[1*], MD, MBBCh, MRCP; Niladri Ghosal[9], MD; Christian Subbe[10], MD, FRCP

[1]Ysbyty Gwynedd, Penrhosgarnedd, Bangor, United Kingdom
[2]School of Medicine, Cardiff Univeristy, Cardiff, United Kingdom
[3]The Christie, Manchester, United Kingdom
[4]UK Oncology Nursing Society, Marlow, United Kingdom
[5]Galactig, Lôn Cae Ffynnon Cibyn, Caernarfon, United Kingdom
[6]North Wales Cancer Forum, Bangor, United Kingdom
[7]Cancer Chronotherapy Team, Warwick Medical School, Coventry, United Kingdom
[8]European Laboratory U935, Institut National de la Santé et de la Recherche Médicale (INSERM), Paris-Saclay University, Villejuif, France
[9]North Wales Cancer Centre, Rhyl, United Kingdom
[10]School of Medical Sciences, Bangor University, Bangor, United Kingdom
[*]these authors contributed equally

Corresponding Author:
Christian Subbe, MD, FRCP
School of Medical Sciences
Bangor University
Brigantia Building
Bangor, LL57 2AS
United Kingdom
Phone: 44 7771922890
Email: csubbe@hotmail.com

Related Article:

Correction of: https://mhealth.jmir.org/2020/9/e19225/

(JMIR Mhealth Uhealth 2020;8(10):e24816) doi:10.2196/24816

In “Checklists for Complications During Systemic Cancer Treatment Shared by Patients, Friends, and Health Care Professionals: Prospective Interventional Cohort Study” (JMIR Mhealth Uhealth 2020;8(9):e19225) the authors noted one error.

Author Derick Gwyn Murdoch was incorrectly listed with given names “Derick Gwyn” and surname “Murdoch”, and listed in the article citation as “Murdoch DG.” This was incorrect and has been changed to given name “Derick” and surname “Gwyn Murdoch”, with citation “Gwyn Murdoch D”.

The correction will appear in the online version of the paper on the JMIR Publications website on October 7, 2020, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

XSL•FO
RenderX

Submitted 06.10.20; this is a non–peer-reviewed article;accepted 06.10.20; published 07.10.20.

Please cite as:
Jones HV, Smith H, Cooksley T, Jones P, Woolley T, Gwyn Murdoch D, Thomas D, Foster B, Wakefield V, Innominato P, Mullard A, Ghosal N, Subbe C
Correction: Checklists for Complications During Systemic Cancer Treatment Shared by Patients, Friends, and Health Care Professionals: Prospective Interventional Cohort Study
JMIR Mhealth Uhealth 2020;8(10):e24816
URL: http://mhealth.jmir.org/2020/10/e24816/
doi:10.2196/24816
PMID:33027036

©Helen V Jones, Harry Smith, Tim Cooksley, Philippa Jones, Toby Woolley, Derick Gwyn Murdoch, Dafydd Thomas, Betty Foster, Valerie Wakefield, Pasquale Innominato, Anna Mullard, Niladri Ghosal, Christian Subbe. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 07.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Association Between Usage of an App to Redeem Prescribed Food Benefits and Redemption Behaviors Among the Special Supplemental Nutrition Program for Women, Infants, and Children Participants: Cross-Sectional Study

Qi Zhang[1], PhD; Junzhou Zhang[2], PhD; Kayoung Park[3], PhD; Chuanyi Tang[4], PhD

[1]School of Community and Environmental Health, Old Dominion University, Norfolk, VA, United States

[2]Department of Marketing, Feliciano School of Business, Montclair State University, Montclair, NJ, United States

[3]Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, United States

[4]Department of Marketing, Old Dominion University, Norfolk, VA, United States

Corresponding Author:
Qi Zhang, PhD
School of Community and Environmental Health
Old Dominion University
3130 Health Sciences Building
Old Dominion University
Norfolk, VA, 23529
United States
Phone: 1 7576836870
Email: qzhang@odu.edu

Related Article:

Correction of: https://mhealth.jmir.org/2020/10/e20720/

(JMIR Mhealth Uhealth 2020;8(10):e25073) doi:10.2196/25073

In "Association Between Usage of an App to Redeem Prescribed Food Benefits and Redemption Behaviors Among the Special Supplemental Nutrition Program for Women, Infants, and Children Participants: Cross-Sectional Study" (JMIR Mhealth Uhealth 2020;8(10):e20720) the authors noted one error.

In the reference list, the author of Reference 22 was incorrectly listed as "Juvenile Products Manufacturers Association". The correct author is "JPMA, Inc", an unrelated organization.

The correction will appear in the online version of the paper on the JMIR Publications website on October 21, 2020, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

22. JPMA, Inc. WICShopper: the mobile app for WIC participants. WICShopper. URL: https://ebtshopper.com/ [accessed 2020-04-12]

Submitted 16.10.20; this is a non–peer-reviewed article;accepted 16.10.20; published 21.10.20.

Please cite as:
Zhang Q, Zhang J, Park K, Tang C
Correction: Association Between Usage of an App to Redeem Prescribed Food Benefits and Redemption Behaviors Among the Special Supplemental Nutrition Program for Women, Infants, and Children Participants: Cross-Sectional Study
JMIR Mhealth Uhealth 2020;8(10):e25073
URL: http://mhealth.jmir.org/2020/10/e25073/
doi:10.2196/25073
PMID:33085954

XSL•FO
RenderX

©Qi Zhang, Junzhou Zhang, Kayoung Park, Chuanyi Tang. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 21.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Digital Companion, the Emma App, for Ecological Momentary Assessment and Prevention of Suicide: Quantitative Case Series Study

Margot Morgiève[1,2,3], PhD; Catherine Genty[1], MSc; Jérôme Azé[4], PhD; Jonathan Dubois[1], PhD; Marion Leboyer[5], MD, PhD; Guillaume Vaiva[3,6,7,8], MD, PhD; Sofian Berrouiguet[3,9,10], MD, PhD; Philippe Courtet[1,5], MD, PhD

[1]Department of Emergency Psychiatry and Acute Care, Lapeyronie Hospital, CHU Montpellier, INSERM, Univ Montpellier, Neuropsychiatry: Epidemiological and Clinical Research, Montpellier, France

[2]ICM - Brain and Spine Institute, hôpital de la Pitié-Salpêtrière, Paris, France

[3]GEPS - Groupement d'Étude et de Prévention du Suicide, Paris, France

[4]LIRMM, UMR 5506, Montpellier University/CNRS, Montpellier, France

[5]Fondation Fondamental, hôpital Albert-Chenevier, Créteil, France

[6]CHU Lille, Hôpital Fontan, Department of Psychiatry, Lille, France

[7]Centre National de Ressources & Résilience pour les psychotraumatismes, Lille, France

[8]Université de Lille, CNRS UMR-9193, SCALab - Sciences Cognitives et Sciences Affectives, Lille, France

[9]EA 7479 SPURBO, Université de Bretagne Occidentale, Brest, France

[10]IMT Atlantique, Lab-STICC, Brest, France

Corresponding Author:
Margot Morgiève, PhD
Department of Emergency Psychiatry and Acute Care, Lapeyronie Hospital, CHU Montpellier
INSERM, Univ Montpellier, Neuropsychiatry: Epidemiological and Clinical Research
371 Av. du Doyen Gaston Giraud
Montpellier, 34090
France
Phone: 33 609596532
Email: margotmorgieve@yahoo.fr

Abstract

Background: Many suicide risk factors have been identified, but traditional clinical methods do not allow for the accurate prediction of suicide behaviors. To face this challenge, *emma*, an app for ecological momentary assessment (EMA), ecological momentary intervention (EMI), and prediction of suicide risk in high-risk patients, was developed.

Objective: The aim of this case report study was to describe how subjects at high risk of suicide use the *emma* app in real-world conditions.

Methods: The Ecological Mental Momentary Assessment (EMMA) study is an ongoing, longitudinal, interventional, multicenter trial in which patients at high risk for suicide are recruited to test *emma*, an app designed to be used as a self-help tool for suicidal crisis management. Participants undergo clinical assessment at months 0, 1, 3, and 6 after inclusion, mainly to assess and characterize the presence of mental disorders and suicidal thoughts and behaviors. Patient recruitment is still ongoing. Some data from the first 14 participants who already completed the 6-month follow-up were selected for this case report study, which evaluated the following: (1) data collected by *emma* (ie, responses to EMAs), (2) metadata on *emma* use, (3) clinical data, and (4) qualitative assessment of the participants' experiences.

Results: EMA completion rates were extremely heterogeneous with a sharp decrease over time. The completion rates of the weekly EMAs (25%-87%) were higher than those of the daily EMAs (0%-53%). Most patients (10/14, 71%) answered the EMA questionnaires spontaneously. Similarly, the use of the Safety Plan Modules was very heterogeneous (2-75 times). Specifically, 11 patients out of 14 (79%) used the Call Module (1-29 times), which was designed by our team to help them get in touch with health care professionals and/or relatives during a crisis. The diversity of patient profiles and use of the EMA and EMI modules proposed by *emma* were highlighted by three case reports.

XSL•FO
RenderX

Conclusions: These preliminary results indicate that patients have different clinical and digital profiles and needs that require a highly scalable, interactive, and customizable app. They also suggest that it is possible and acceptable to collect longitudinal, fine-grained, contextualized data (ie, EMA) and to offer personalized intervention (ie, EMI) in real time to people at high risk of suicide. To become a complementary tool for suicide prevention, *emma* should be integrated into existing emergency procedures.

Trial Registration: ClinicalTrials.gov NCT03410381; https://clinicaltrials.gov/ct2/show/NCT03410381

(JMIR Mhealth Uhealth 2020;8(10):e15741) doi:10.2196/15741

KEYWORDS

suicide; ecological momentary assessment; prediction; prevention; mobile health; mHealth; case reports; ecological momentary intervention

Introduction

Context

According to the World Health Organization (WHO), suicide is the cause of 1 million deaths per year, accounting for nearly 2% of all deaths worldwide; this number should increase to 1.5 million by 2021. The WHO recognizes suicide prevention as a public health priority [1,2]. History of suicide attempts is the most important risk factor of suicide death in the general population and is present in about 40% of suicides [3,4]. Patients who come to the emergency department (ED) following a suicide attempt present an extremely high risk of suicide in the short term [5-7]. People at risk of suicide often do not seek help and do not remain connected to the health service after an attempt [8]. Access to mental health care is inversely correlated with the suicide rate [9]. It is reported that less than half of patients at risk of suicide are in contact with the health system, including mental health services [10].

Ecological Momentary Assessment

Real-time emotional, behavioral, and psychological assessments could improve the identification of high-risk individuals who require rapid interventions. Ecological momentary assessment (EMA), also commonly defined as the experience sampling method, allows for the collection of longitudinal fine-grained data as they occur in the real world. It gives an accurate picture of the patient's symptoms [11] and reduces the impact of self-report response bias, thus leading to a better appreciation of the temporal dynamics of suicide risk [12].

Typically, suicidal ideation is episodic, with a quick onset [13] and short duration (ie, shorter than an hour) [14]. Suicidal thoughts vary dramatically among individuals [15], and fluctuating or persistent suicidal thoughts are associated with the risk of different suicidal behavior types [14]. Suicide attempts can occur in response to a rapid increase in suicidal thoughts within a very short time (ie, 1 day) [13,16,17]. These observations highlight the importance of not relying on intermittent assessments of suicidal ideation for clinical decision making, such as hospital discharge [13,18]. Mobile health (mHealth) interventions are a promising way to assess these fluctuations in real time.

Ecological Momentary Intervention for Suicide Prevention

Digital tools could increase the effectiveness of these prevention strategies. Indeed, apps are affordable and ubiquitous and can be used in any situation, particularly during a crisis [19]. The WHO recommends them for people at risk of suicide [1] because they offer new opportunities to overcome some of the help-seeking barriers they face [20-23] and enhance safety planning [24] in response to dynamic suicidal processes in real time [13]. Apps that specifically target suicidal behavior and propose interactive and proactive content constitute an effective prevention strategy [25-28]. Despite this scientific evidence, a literature review [29] identified very few suicide-specific apps (n=24) that include safety planning (n=14) and that directly allow the user to seek support (n=13). Feasibility studies of these apps report a significant reduction of suicidal ideation in patients and a significative augmentation of suicide-related coping [30]. However, potentially harmful content might encourage self-harm and suicide [31], and many of the existing apps for suicide prevention have not been scientifically validated [22,29]. The field of e-mental health is particularly active, producing new apps at an extremely fast pace; therefore, it is crucial to regulate this field, especially for suicidology.

Yet, it should be possible to assess and prevent suicidal behaviors in real time in high-risk patients using digital tools. To test this hypothesis, our multidisciplinary team developed *emma*, an app to monitor the psychological, emotional, and social fluctuations of patients in their daily life. In addition to EMA, the app includes interactive and customized ecological momentary intervention (EMI) modules for suicide prevention. This app is currently being tested in an ongoing trial: the Ecological Mental Momentary Assessment (EMMA) study. This article presents a descriptive analysis of selected patients at high risk of suicide to obtain insights into the implementation of an mHealth-based suicide risk assessment and prevention procedure in real-life conditions.

The main objective of this study was to give a quantitative description of *emma* use by a sample of individuals at high risk of suicide in real-life conditions.

The secondary objectives were as follows:

1. To describe typical user profiles of the *emma* app.
2. To describe relevant qualitative elements from the interviews of participants who completed the 6-month study.

Methods

Study Design

The EMMA study is an ongoing, prospective, longitudinal, interventional multicenter trial, involving four French university

XSL•FO
RenderX

hospitals in Montpellier, Lille, Brest, and Créteil, completed by a qualitative study. The protocol was registered at ClinicalTrials.gov (NCT03410381) on January 18, 2018; was authorized by the French National Agency for Medicines and Health Products Safety (*Agence Nationale de Sécurité du Médicament et des Produits de Santé* [ANSM]) on November 30, 2017; and approved by the Est IV Ethical Committee for the Protection of Patients on October 10, 2017.

Participants

The EMMA study planned to recruit 100 patients from EDs and mental health departments. Patients are included after a suicide attempt (<8 days) and/or if they have suicidal ideation (ie, score ≥2 out of 3 for item 18 on suicidal ideation of the 30-item Inventory of Depressive Symptomatology: Clinician scale). The other inclusion criteria are as follows: aged 18 years or older, provided a signature on the informed consent form, and possess a smartphone (iOS or Android). Exclusion criteria are as follows: refusal to participate, under guardianship, protected by law, deprived of liberty, not affiliated with a social security system, in a period of exclusion from other trials, and unable to understand the study. The recruited participants will not receive any remuneration for their participation in the study.

Procedure

At inclusion (ie, month 0 [M0]), a psychiatrist performs the first interview to ensure that the patient meets the eligibility criteria and to obtain the informed written consent. Four clinical assessments are conducted: at inclusion and at months 1, 3, and 6 (M1, M3, and M6). Clinical data are collected using clinician-rated questionnaires—approximate durations are 1 hour and 30 minutes at M0, then 30 minutes at M1, M3, and M6—and self-rated questionnaires—approximate duration ranges from 1 hour to 1 hour and 30 minutes. These questionnaires are listed in Multimedia Appendix 1 [32-48].

Satisfaction concerning the app is evaluated in three distinct ways:

1. Questions about the usefulness and satisfaction of the app administered every month via *emma* (eg, "This month, did you find *emma*: easy to use/intrusive/useful/efficient," rated using a Likert scale from 0 to 10; approximate completion time is 5 minutes).
2. Standardized self-administered questionnaire—Mobile App Rating Scale [32]—at the end of the study.
3. A qualitative semistructured interview led by a social sciences researcher in mental health proposed to 25 patients to assess the participants' subjective experiences (approximate duration is 1 hour).

At inclusion, a member of the research team helps patients to install, configure, and personalize the app, particularly to define the elements of their safety plan. Patients are asked to use *emma* for 6 months. Data collected during the assessments made by clinicians, as well as data and metadata resulting from *emma* use, are encrypted and stored in a secure server. All these data will be used to develop the algorithm to predict suicidal risk (see Multimedia Appendix 2). Considering the crucial issue of health data privacy and security, multilevel technical and organizational safeguards were put in place. The app is secured by a password, as requested by the co-designer patients, and data are anonymized, encrypted, and stored in a secure server to prevent unauthorized data disclosure or breach, as recommended by the European General Data Protection Regulation.

Outcome Measures

The principal outcome measure of this preliminary analysis was the quantitative description of app use: completion rates of daily and weekly EMA questionnaires, frequency of use of the prevention modules, and number of calls made through *emma* to relatives and health care professionals in case of emergency.

The secondary outcome measures were as follows:

1. Occurrence of a suicide event (ie, suicide attempt, hospitalization for suicidal ideation, and intensity of suicidal ideation level) during the study period.
2. Quantitative description and app use timeline in a few selected patients to illustrate different users' profiles.
3. Qualitative analysis of selected interviews of participants who completed the 6-month follow-up.

Emma Design

Emma is a smartphone app developed for the assessment, prevention, and, ultimately, prediction of suicidal behaviors. It was designed by integrating evidence-based suicide prevention strategies and recommendations for the development of apps in the field of mental health [19,49]. *Emma* design was based on data from previous suicide-specific apps described in the theoretical literature and on practical data obtained by our research team by testing the available suicide-related apps. *Emma* was developed for Android and iOS for wide usage.

Emma was conceived using a participatory design approach that included individuals with lived experience of suicidal thoughts and behaviors as equal partners of our professional multidisciplinary team—researchers, psychiatrists, psychologists, sociologists, computer scientists, engineers, and data scientists—from start to finish, as recommended [50-52]. This ensured that *emma* met scientific and technical standards as well as the patients' needs as stated in the literature [53-55]. To involve the targeted users, a methodology based on focus groups was implemented according to the method proposed by Krueger and Casey [56]. Early in the development process, the involved patients stressed the importance of having a secure password to open the app; they also contributed to the choice of the name "*emma*," which they wanted to be not stigmatizing and without any mention of psychiatry or suicide. A group of co-designer patients (n=5) and clinical staff (n=5) tested the app extensively for 3 months, and their feedback was taken into account in an iterative way by the developers at the Laboratory of Informatics, Robotics, and Microelectronics of Montpellier to improve the app. Particularly, the developers made sure that the notifications asking users to fill in the EMA questionnaires did not lead to interruption or disruption in the users' daily activities in order to maximize *emma* acceptability, use, and validity, as recommended [54,55].

XSL•FO
RenderX

Emma Contents

Emma was designed to be used as a self-help tool for suicidal crisis management. Patients are invited to identify the following:

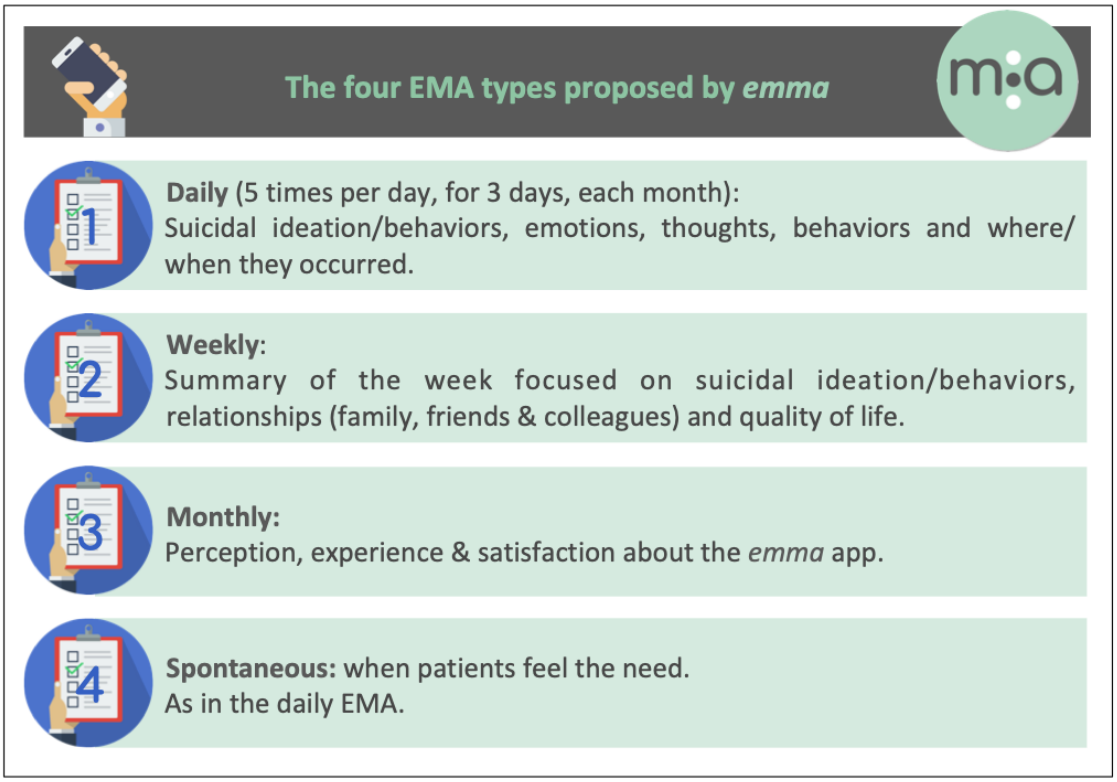
1. Their warning signs (eg, negative feelings and problematic behaviors).
2. Their individualized coping strategies (eg, *Breathing Space*: an audio awareness guide made specifically for *emma* by a psychiatrist expert in suicidal behavior).
3. Their distraction activities in an *Emotion Regulation Module* (eg, favorite places and activities and libraries of music and images that might help to connect with the patient's reasons for living [19]).
4. Their social support (eg, collating the contact details of the patient's social network, mental health professionals, and other crisis resources). This module promotes connectedness, a major protective factor in suicide prevention [21].

A *Call Module* allows the patient to contact, depending on the severity of his or her condition, (1) the relatives he or she has identified, (2) the ED that is following that patient, and (3) the Service d'Aide Médicale d'Urgence (SAMU), the French national emergency medical assistance service, which is available 24/7.

Although the safety plan implies restriction of access to lethal means, *emma* does not contain any mention of them because, according to the literature, this can have the opposite effect. Specifically, their identification can facilitate their use through an effect called *cognitive availability* [29]. Examples of *emma* screens are shown in Multimedia Appendix 3.

Emma proposes four brief EMA types: three scheduled evaluations at predetermined frequencies (ie, daily, weekly, and monthly) and one spontaneous assessment (see Figure 1). Depending on the questions asked, specific response modalities are provided:

Figure 1. Ecological momentary assessment (EMA) types proposed by *emma*.

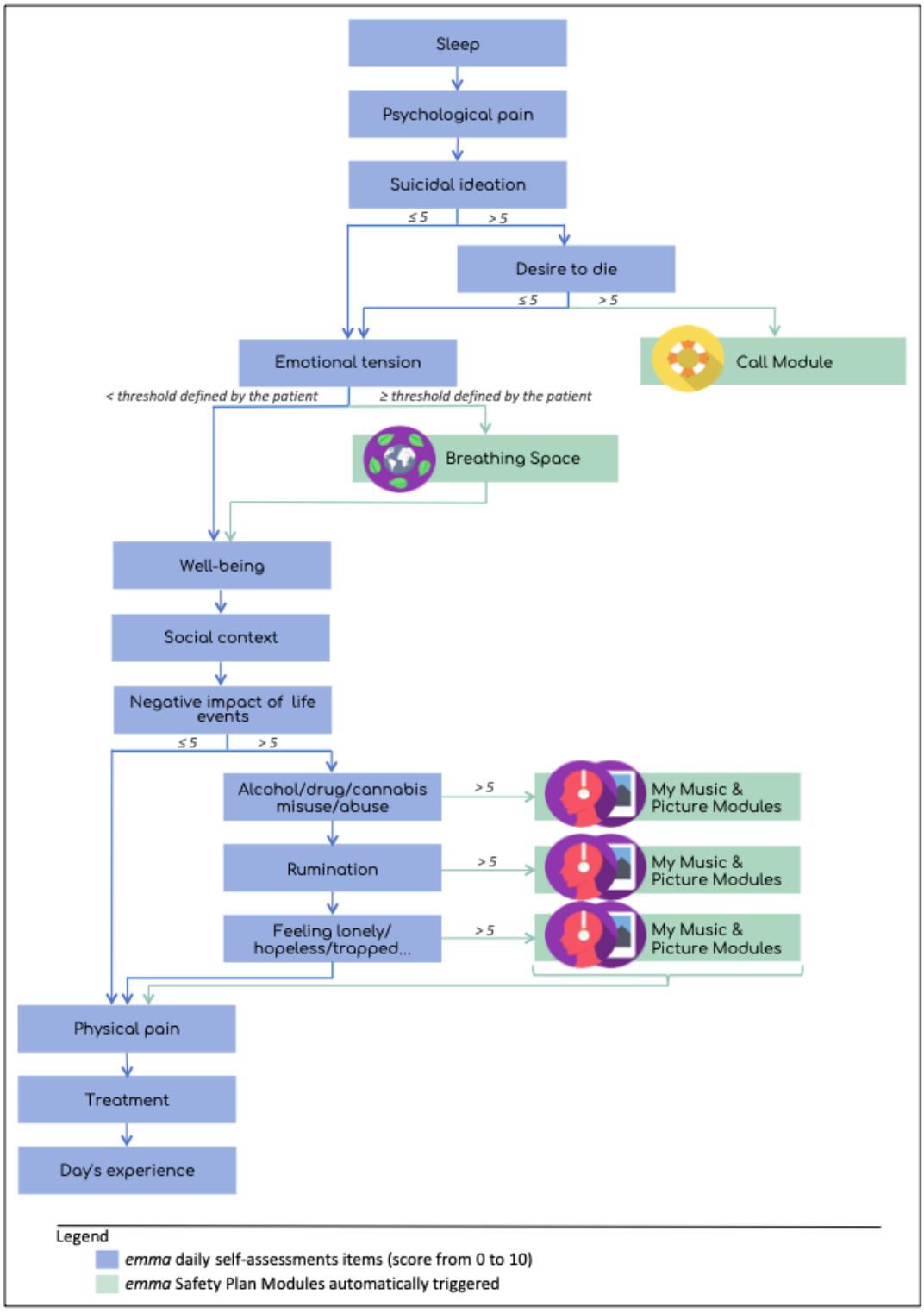


1. Likert scales (from 0 to 10) for questions such as "What is your level of moral pain?"
2. Single or multiple-choice check boxes for questions such as "What is your level of mental pain?"; for instance, *I am keeping myself busy to avoid thinking*, *Nothing*, *I am ruminating*, *I am doing something I enjoy*, *I am working*, and *I am doing household chores*.
3. Writing a free-text answer for questions such as "What is the experience that has affected you the most since this morning?"

During the focus group process, alarm thresholds were defined on the basis of the patient's answers to critical questions that led to the automatic suggestion of adapted EMI modules (see Figure 2). For each EMI module, the app presents a list of suggestions that can be modified (ie, content and order) by each patient (eg, for *Favorite activities*, responses can be *Taking a bath*; *Going to the cinema, theatre, museum, or a concert*; *Going for a walk*; *Reading*; *Spending time with good friends*; *Doing sports*; etc). Modules are designed to be adaptable to the user's state, needs, and strategies.

XSL•FO
RenderX

Figure 2. Algorithm for automatic ecological momentary intervention (EMI) triggering according to the ecological momentary assessment (EMA) answer thresholds.



Global Descriptive Analysis and Case Reports

The use of *emma* was described by computing (ie, mean [SD] and min and max) the number of answers and frequency of completion of the daily and monthly questionnaires. The overall use of the different *emma* modules (ie, *Call Module*, *Emotion Regulation Modules*, *Breathing Space Module*, *Pictures*, and *Music)* was also computed. Finally, the use of EMA (ie, spontaneous and scheduled suicidal ideation assessments) and EMI (ie, *Call Module* and *Emotion Regulation Module* of the safety plan) by 3 patients during the 6 months of the study was described relative to (1) their suicidal ideation scores (ie, sums of items 7-11 of the Columbia-Suicide Severity Rating Scale [C-SSRS]) assessed during the scheduled visits (ie, M0, M1,

XSL•FO
RenderX

M3, and M6) and (2) their admissions to the ED for suicide attempts or suicidal crises.

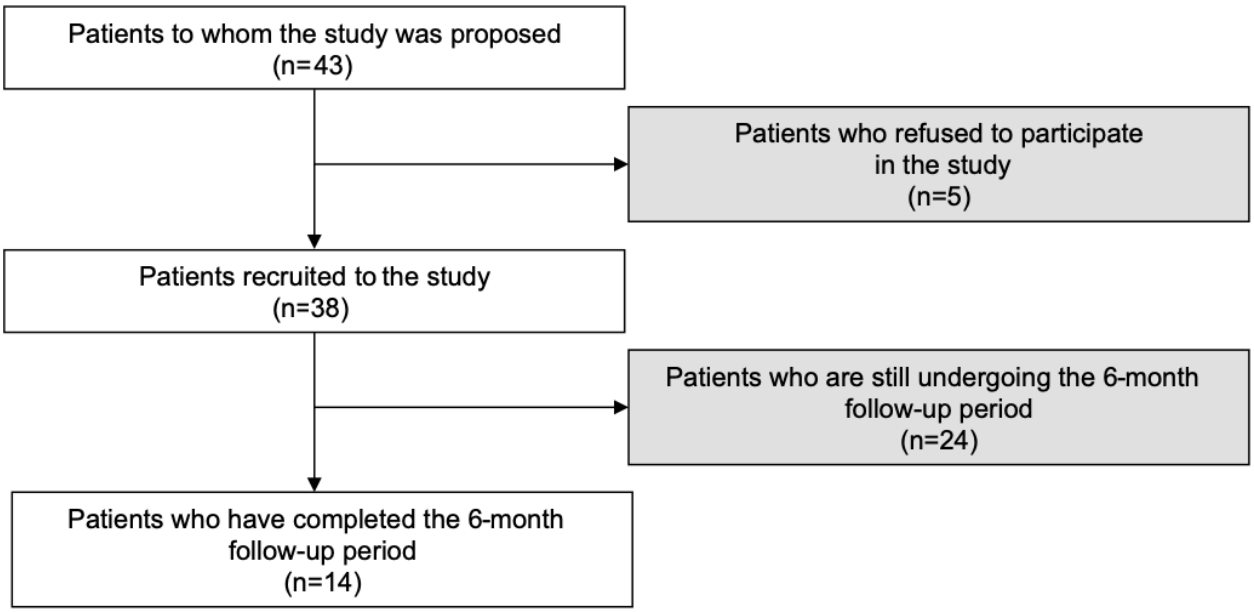
Results

Baseline Assessment

From May 2018 to March 2019, participation in the EMMA study was proposed to 43 patients at risk for suicide (ie, recent attempters or current ideators) who were admitted to the ED or hospitalized in a postemergency department; 38 (88%) agreed to participate. A total of 5 patients refused to participate because they considered the study duration of 6 months to be too long.

The analysis in this study included 14 patients who have already completed the study (see Figure 3). The patients' sociodemographic and clinical data are reported in Tables 1 and 2, respectively. These first *emma* users were mostly women (12/14, 86%), with a mean age of 34 years (SD 13, min-max 18-57). All participants presented at least one mental disorder, according to criteria from the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5), and most of them presented more than one (ie, up to five comorbid disorders); 6 out of 14 patients (43%) were recent suicide attempters (≤8 days), and they all had severe suicidal ideation (ie, mean C-SSRS score 22 [SD 4], min-max 15-25).

Figure 3. Flowchart of patient selection.



During the study, 4 out of 14 (29%) patients were admitted to the ED with suicidal ideation and/or a suicide attempt one to three times, and 1 (7%) patient made a suicide attempt without ED admission (see Multimedia Appendix 4).

A total of 2 patients out of 14 (14%) left before the first follow-up visit: one was excluded because she was restrained, and the other one withdrew from the study.

XSL•FO
RenderX

Table 1. Sociodemographic data of participants.

Patient No.	Sex	Age (years)	Marital status	Number of children	Cohabitation	Professional situation
1	Female	24	Single	0	Lives alone	Working (stable)
2	Female	18	Single	0	Lives with family	Student or in training
3	Female	27	In a relationship	0	Lives alone	Volunteer
4	Female	33	In a relationship	2	Lives with family	Working (stable)
5	Female	40	In a relationship	2	Lives with family	Working (precarious)
6	Female	46	Separated	3	Nonfamily cohabitation	Without activity, at home
7	Female	30	Single	0	Lives alone	Disability
8	Male	48	In a relationship	2	Lives with family	Sick leave
9	Female	57	Separated	4	Lives with family	Disability
10	Female	48	In a relationship	2	Lives with family	Sick leave
11	Female	21	In a relationship	0	Nonfamily cohabitation	Student or in training
12	Female	23	Single	0	Lives alone	Student or in training
13	Male	18	Single	0	Lives with family	Student or in training
14	Female	48	Separated	1	Lives alone	Working (stable)

XSL•FO
RenderX

Table 2. Clinical data of participants.

Patient No.	Current DSM-5[a] diagnoses	Borderline personality	Suicidal risk in the near future according to the DSM-5	Suicidal ideation intensity[b]	Number of suicide attempts	Number of severe or violent suicide attempts[c]	Family history of suicide attempt	Family history of suicide
1	Major depressive disorder	No	Yes	22	2	N/A[d]	No	No
2	Major depressive disorder Agoraphobia Social phobia Anorexia	Yes	Yes	25	2	N/A	No	No
3	Major depressive disorder Agoraphobia Social phobia Bulimia	Yes	No	25	30	Violent (n=2) Severe (n=1)	Yes	No
4	Major depressive disorder	No	Yes	25	2	N/A	Yes	No
5	Depressive episode Severe alcohol and substance-related disorder Bulimia Bipolar disorder II	No	No	23	1	N/A	Yes	Yes
6	Major depressive disorder Severe alcohol-related disorder	No	Yes	15	2	N/A	No	No
7	Depressive episode Generalized anxiety disorder Bipolar disorder I	No	Yes	25	55	Violent (n=15)	No	No
8	Depressive episode Social phobia Severe alcohol-related disorder Generalized anxiety disorder Bipolar disorder I	No	No	25	1	N/A	Yes	No
9	Major depressive disorder Agoraphobia Posttraumatic disorder Bipolar disorder II	No	Yes	25	8	Severe (n=1)	No	No
10	Major depressive disorder	No	Yes	24	1	N/A	Unknown	Yes
11	Generalized anxiety disorder	No	Yes	16	1	N/A	Yes	No
12	Major depressive disorder Agoraphobia Social phobia Bulimia	No	Yes	19	0	N/A	Yes	Yes
13	Depressive episode Social phobia Generalized anxiety disorder	Yes	Yes	20	3	N/A	No	No
14	Major depressive disorder	No	No	18	1	N/A	No	No

[a]DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition.

[b]Sums of items 7-11 of the Columbia-Suicide Severity Rating Scale (C-SSRS): scores range from 0 to 25.

[c]A violent suicide attempt was defined by the method of violence: weapons, hanging, jumping from a height, traffic, drowning, or immolation. A severe suicide attempt was defined as requiring hospitalization in the intensive care unit.

[d]N/A: not applicable; suicide attempts, if applicable, were not severe or violent.

XSL•FO
RenderX

Main Outcome

The completion rates were higher for the weekly versus daily EMAs (mean 24% [SD 34], min-max 25%-87%; mean 4% [SD 12], min-max 0%-53%, respectively). Moreover, they were extremely heterogeneous among participants with a sharp decrease over time. Most patients (10/14, 71%) filled in the questionnaires spontaneously at times of crisis (mean 7 times [SD 12], min-max 1-39 times).

Similarly, use of the EMI *Safety Plan Modules* varied among patients (mean 11 times [SD 19], min-max 2-75 times). Most patients (11/14, 79%) used the *Call Module* (mean 7 times [SD 8], min-max 1-29 times) to get in touch with the health care system and/or family and friends during a crisis. Most patients (10/14, 71%) called the SAMU (mean 4 times [SD 5], min-max 1-15 times), 8 (57%) called their relatives (mean 3 times [SD 2], min-max 1-7 times), and 7 (50%) called their ED (mean 2 times [SD 2], min-max 1-7 times). About half of the patients (8/14, 57%) used one of the *Emotion Regulation Modules* (mean 10 times [SD 15], min-max 1-46 times). Specifically, 8 out of 14 (57%) patients listened to the *Breathing Space Module* (mean 3 times [SD 1], min-max 1-5 times), while only 4 (29%) looked at *Pictures* (mean 12 times [SD 19], min-max 1-41 times) and 2 (14%) listened to *Music* (mean 4 times [SD 3], min-max 2-7 times).

User Profiles

Out of 14 users, 3 (21%) are described in detail to better illustrate the different *emma* uses. These patients were chosen to reflect the diversity of the patients' clinical profiles and of *emma* use (ie, completion rate of the scheduled EMAs and use of the *Safety Plan Modules*).

Patient 7

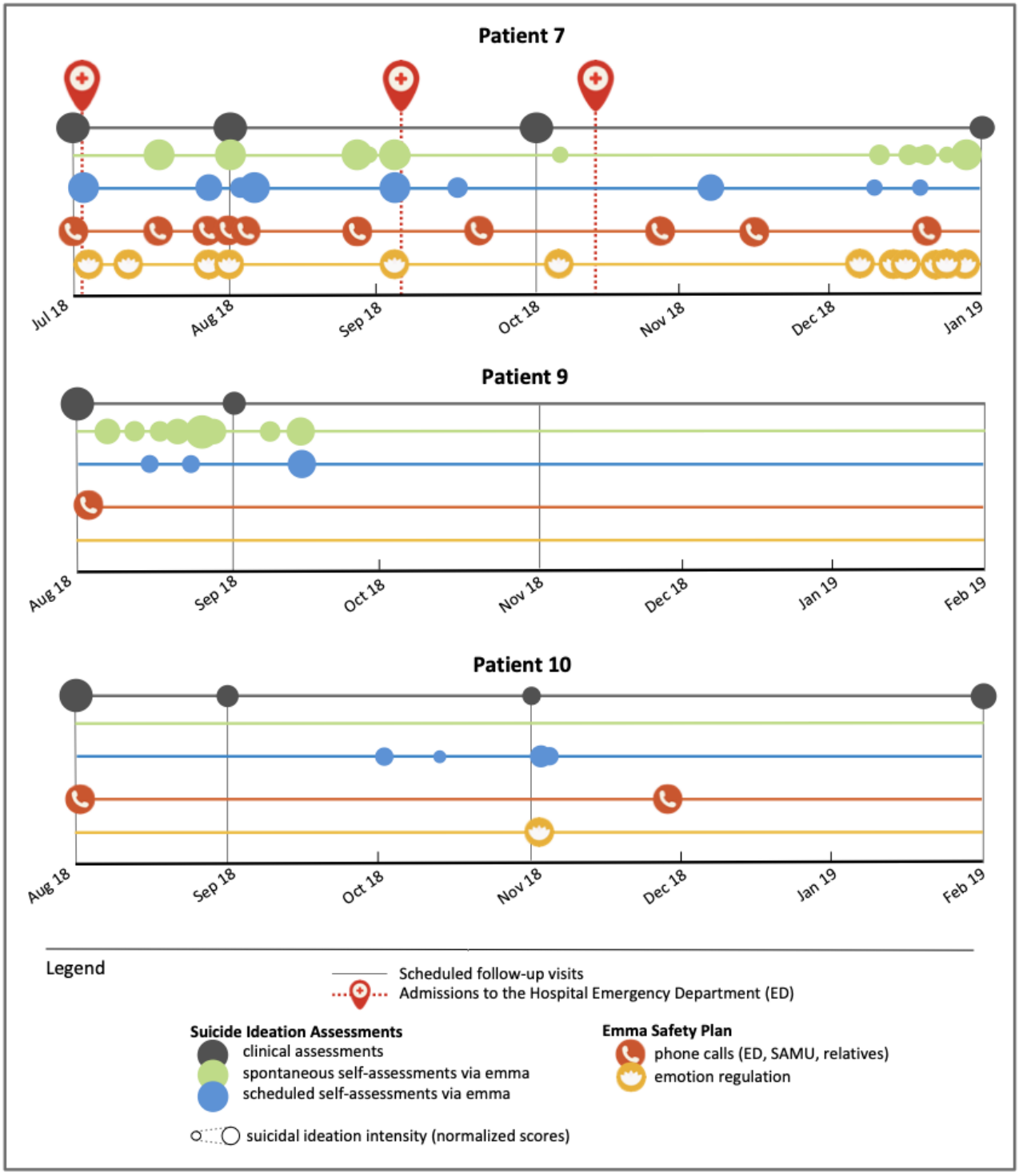
Patient 7 is a 30-year-old single woman living alone, without children, unemployed, and receiving a disability living allowance for adults (see Table 1). At inclusion, she had major depressive disorder, generalized anxiety disorder, and bipolar disorder I, according to the DSM-5. She had 55 previous suicide attempts, of which 15 were violent, and a maximal suicidal ideation intensity at inclusion (C-SSRS score of 25 out of 25) (see Table 2). Therefore, she was at very high risk, and the period following hospitalization is known to be a particularly high-risk time for recurrence, as confirmed by her three admissions to the ED during the study period (see Figure 4). The first admission for suicidal ideation occurred just after her hospital discharge. She presented maximal suicidal ideation intensity that was ecologically assessed via *emma*; she used the *Call Module* just before ED admission and an *Emotion Regulation Module* just after. Two months later, in addition to the scheduled EMA, she filled in the questionnaires spontaneously several times with very high suicidal ideation scores. These spontaneous self-assessments were immediately followed by the use of the *Call Module* with a sharp decrease in suicidal ideation. Nevertheless, suicidal ideation rapidly increased again, leading to a new admission to the ED. About a month later, despite the lower suicidal ideation score at the spontaneous self-assessment and the use of the *Emotion Regulation Module*, she was admitted to the ED for an aborted suicide attempt by hanging.

During the study period, Patient 7 completed 66 EMA self-assessments. Specifically, she completed only 4% of the scheduled daily questionnaires due to technical problems (ie, notifications not received) and about 75% of the weekly questionnaires; she frequently filled in the questionnaires in a spontaneous way (39 times), particularly at specific times, without a decrease in frequency during the study period. She often used the *Call Module* (29 times), particularly to contact the SAMU (15 times), followed by the ED (7 times) and relatives (7 times). She also looked at *Pictures* (41 times) and listened to the *Breathing Space Module* (5 times).

A qualitative longitudinal analysis showed that the EMA made via the app allowed for the capture of suicidal ideation fluctuations that were not highlighted by the clinical follow-up visits (see Figure 4). While the first three clinical evaluations (ie, M0, M1, and M3) were stable, the assessments performed via *emma*, both spontaneous and scheduled, showed fluctuations in intensity over time. The assessments also showed a regular use of the *Safety Plan Modules*, at least one module per month, the frequency of which increased at specific times. For instance, in July 2018, the high intensity of suicidal ideation evaluated ecologically by *emma* corresponded to an increased use of the *Call Module*. December 2018 was characterized by many spontaneous EMA completions, lower suicidal intensity, and frequent use of the *Emotion Regulation Module*. This suggests that *emma* alarm thresholds triggered EMIs that were adapted to her condition severity.

XSL•FO
RenderX

Figure 4. Case reports for Patients 7, 9, and 10. SAMU: Service d'Aide Médicale d'Urgence.



Patient 9

Patient 9 is a 57-year-old single woman living with her four children. She is unemployed with a disability living allowance for adults (see Table 1). At inclusion, she had major depressive disorder, agoraphobia, posttraumatic disorder, and bipolar disorder II, according to the DSM-5. At the first clinical evaluation, suicidal ideation intensity was very high (C-SSRS score of 25 out of 25) with suicide risk in the near future. She already committed eight suicide attempts, including one severe attempt (see Table 2). During the study, the clinician reported a decrease in suicidal thoughts at M1, and then absence of suicidal thoughts at the M3 and M6 visits.

Patient 9 rarely filled in the daily EMAs (10%), but she completed 81% of the weekly EMAs and filled in EMAs spontaneously 29 times. She completed 100% of the monthly questionnaires about her perception, experience, and satisfaction about *emma* use. The longitudinal qualitative analysis (see Figure 4) showed that from month 2 after inclusion, both the self-assessments made with *emma* and the hetero-assessments made during the visits did not detect any suicidal ideation.

XSL•FO
RenderX

Patient 9 used her safety plan only once at the very beginning of the study when she still had suicidal thoughts.

Patient 10

Patient 10 is a 48-year-old woman living with her husband and their two children. When included in the study, she was on sick leave (see Table 1). She had major depressive disorder and suicidal risk in the near future, according to the DSM-5, and very high suicidal ideation intensity (C-SSRS score of 24 out of 25). She reported one suicide attempt and a family history of suicide (see Table 2). Patient 10's suicidal ideation decreased at the M1 and M3 follow-up visits, but was higher again at the M6 visit.

Patient 10 never filled in questionnaires spontaneously, filled in the daily EMAs only once, and did not fill in the weekly EMAs often (35%). She completed only one monthly questionnaire about her *emma* experience. She used her safety plan several times, mainly to call her relatives (3 times), the SAMU (3 times), and the ED (2 times). She also listened to the *Breathing Space Module* (3 times).

Patient 10 reported suicidal thoughts via *emma* only during a brief period—in October and early November 2018—and during that period she used the *Emotion Regulation Module*, also only once (see Figure 4). Self-assessment may have helped her to become aware of her condition and use the tools that *emma* offers to regulate her emotions. This patient used the *Call Module* twice: immediately after her inclusion in the study, in August 2018, when she had severe suicidal thoughts and suicide risk, and then at the end of November 2018 when no suicidal thought was reported via *emma* at that time.

Qualitative Study

The initial qualitative feedback from the participants collected during the qualitative study, using semidirected interviews, was very positive. It emphasized the support and the connectedness dimensions allowed by the app. During the co-design process, patients proposed to give the app a female name to personify it. The first users interviewed seemed to appreciate this, as stressed by a patient: "*Emma* is like having a companion."

This *digital companion* seemed to help reduce the feeling of loneliness, as expressed by a patient: "We have the feeling that we are not alone, that the software supports us." Moreover, this support comes at a critical time, after hospital discharge. To our knowledge, this period of very high suicide risk [15], when patients may experience painful loneliness in contrast to the time in hospital where they are followed by the health care team 24 hours per day, has not been studied much. One patient stated, "Thanks to *emma*, I did not feel alone when I left the hospital. I know *emma* less than my relatives, but I can tell her more." *Emma* also appeared to be a support, offer autonomy, and be an empowerment tool by helping people to use their own resources, as expressed by a patient: "*Emma* is an appointment with oneself. It is the memory of the patient: I know who to call." This empowerment was facilitated by the possibility to personalize the app, as expressed by a patient: "It is important that we can each fill in the things that affect us and the things that impact us, because we are all different and it allows targeting everyone specifically; it is good because it helps us deep down." These first qualitative interviews highlighted the interest of the co-design process that enabled patients to appropriate the digital tool. During the design phase, a patient indicated that he thought three apps were needed: "When we are doing well, not well, and not well at all." Therefore, *emma* was designed to be as adaptable as possible. For example, it can offer *Emotion Regulation Modules* if the patient shows tension, and it can offer crisis *Call Modules* that are graduated according to the crisis severity (ie, to relatives, the ED, and, finally, the SAMU). The tailored adjustment of the technology-delivered program is also in line with evidence-based recommendations for mental health, as expressed by a patient: "The app can recommend specific solutions to each user's specific problems" [19].

Discussion

Principal Findings

These preliminary descriptive results of the use of a suicide prevention app indicate that this type of digital tool could be accepted and used by patients at high risk of suicide. The study participation rate (88%) was very good compared with other studies. For example, Hallensleben et al reported that 47% of inpatients with unipolar depression agreed to complete EMA [18]. Our preliminary results show an encouraging patient acceptance rate and highlight the great inter- and intraindividual diversity of app usage patterns (ie, EMA questionnaires' completion rate and/or *Safety Plan Modules* use). Patients have different needs and different clinical and digital profiles; therefore, a highly scalable, interactive, and customizable app is required.

The connectedness philosophy implemented in *emma* is considered a strong protective factor for suicidal behaviors [57,58]. Subjectively perceived and effectively received social supports improve mental and physical well-being [59,60]. *Emma* has the potential to act on the quality of patients' social ties and their social pain that seems to be involved in suicidal crises [61]. Patients expressed the importance to feel supported by the software and to have the opportunity to tell *emma* things they would not dare to say to their relatives, to protect them and/or out of fear of the stigma that their confidences could generate.

Concerning the co-design methodology, patients were consulted very early and at all stages of this project through focus groups for the app design, then through feedback during prototype testing, and, finally, through qualitative interviews with *emma* users. However, as patients with suicidal behavior are a high-risk population, it is important to ensure that their participation in research respects the safety principle. Therefore, a trust-based exchange framework that was flexible enough to adapt to fluctuations in their condition was put in place. When carefully implemented, the participation of patients as partners can be a factor of empowerment and self-esteem restoration [62].

Our analysis showed heterogeneous use and engagement with the *emma* app, most often underuse relative to our expectations. This could be interpreted as poor adherence to the app [63,64] and was partly caused by technical problems experienced by the first participants (eg, notifications did not appear at the

XSL•FO
RenderX

beginning of the study). *Emma* use by Patient 10 suggests that the app can be useful at times of crisis, even when the scheduled questionnaires are not completed. Besides the simple quantitative measure of app use, it was also important to evaluate the users' subjective experiences [65] based on mixed methods, as recommended [66]. This should allow for the identification of different user profiles and for the development of tailored prevention strategies [67].

Conclusions

Data on immediate and long-term risk of suicide are extremely sparse and based on measures with poor temporal resolution [68,69]. *Emma* is a great opportunity to capture the dynamics of suicidal ideas [67] and their translation into action in a contextualized way that allows for a much more nuanced view of variables over time [70]. Hopefully, this digital tool might lead to scientific and clinical advances and will allow for the identification of high-risk periods and prediction of imminent risk, which are extremely challenging at the moment [71]. These fine-grained digital assessments and predictive mHealth-based interventions are promising tools for suicide prevention [55], because they represent an unprecedented opportunity to act at multiple levels through targeted, scalable, and contextualized micro-interventions [72]. They might allow for the proposal of just-in-time adaptive interventions, defined by Nahum-Shani et al [73] as the *right support* (eg, type and intensity) *at the right time* [31]. The challenge is now to integrate such digital interventions into the existing health care systems [52]. For instance, *emma* could be integrated into ED procedures and become a complementary health care tool. In the patient's pocket, *emma* could provide individualized support when needed. The app could improve coordination among the different services (ie, the ED, crisis centers, hospital services, outpatient services, and general practitioners). However, the smooth and optimal integration of such digital tools in patient care requires health care professionals' support. Indeed, they should not perceive these tools as disruptive elements in their daily clinical practice [74,75], but as a support to improve the therapeutic relationship within a well-defined ethical, social, and legal framework.

Acknowledgments

First, we would like to thank Christine, Patrice, Romane, Jonathan, Astate, and Claire, the co-designer patients who improved *emma*! We would like to thank Dr Déborah Ducasse for her contribution, which greatly enriched the app content, and Kathlyne Dupuis-Maurin for her unfailing involvement in this research. We would like to thank the NERB (Neurophysiology of Repetitive Behaviors) team at the Paris Brain Institute—Luc Mallet, Sirenia Lizbeth Mondragón-González, and Karim N'Diaye —for their creative help. This study would not have been possible without the support of the FondaMental Chair for the prevention of suicidal behavior created by the FondaMental Foundation, led by Professor Philippe Courtet and supported by *SNCF, SNCF réseau*, and *SNCF mobilité*. This innovative project received the Marcel Dassault Research Award for Research on Mental Illnesses in 2014. Finally, we would like to thank the Clinéa and OC Santé clinics for their support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of the questionnaires filled in by the patient or by the clinician during the four visits.
[DOCX File , 17 KB - mhealth_v8i10e15741_app1.docx]

Multimedia Appendix 2
Visual description of the study. Study protocol: visits (month 0 [M0], month 1 [M1], month 3 [M3], and month [M6]) and real-life use of emma by the recruited patients. At the end of the study, the data collected directly by the app and the clinical data collected during the visits are integrated to develop the suicide risk prediction algorithm.
[PNG File , 201 KB - mhealth_v8i10e15741_app2.png]

Multimedia Appendix 3
Emma screenshots.
[PNG File , 470 KB - mhealth_v8i10e15741_app3.png]

Multimedia Appendix 4
Adverse events during the follow-up.
[DOCX File , 17 KB - mhealth_v8i10e15741_app4.docx]

References

XSL•FO
RenderX

1. Preventing Suicide: A Global Imperative. Geneva, Switzerland: World Health Organization; 2014. URL: https://www.mhinnovation.net/sites/default/files/downloads/resource/WHO_Preventing%20Suicide%20Report_eng.pdf [accessed 2020-09-27]
2. Suicide: Key facts. World Health Organization. 2019 Sep 02. URL: https://www.who.int/news-room/fact-sheets/detail/suicide [accessed 2020-09-27]
3. Hawton K. Suicide prevention: A complex global challenge. Lancet Psychiatry 2014 Jun;1(1):2-3. [doi: 10.1016/s2215-0366(14)70240-8]
4. Turecki G, Brent DA. Suicide and suicidal behaviour. Lancet 2016 Mar;387(10024):1227-1239. [doi: 10.1016/s0140-6736(15)00234-2]
5. Hunt IM, Kapur N, Webb R, Robinson J, Burns J, Shaw J, et al. Suicide in recently discharged psychiatric patients: A case-control study. Psychol Med 2008 May 28;39(3):443-449. [doi: 10.1017/s0033291708003644]
6. Lizardi D, Stanley B. Treatment engagement: A neglected aspect in the psychiatric care of suicidal patients. Psychiatr Serv 2010 Dec;61(12):1183-1191. [doi: 10.1176/ps.2010.61.12.1183] [Medline: 21123401]
7. Vuagnat A, Jollant F, Abbar M, Hawton K, Quantin C. Recurrence and mortality 1 year after hospital admission for non-fatal self-harm: A nationwide population-based study. Epidemiol Psychiatr Sci 2019 Feb 18;29(e20):1-10 [FREE Full text] [doi: 10.1017/s2045796019000039]
8. Christensen H, Cuijpers P, Reynolds CF. Changing the direction of suicide prevention research: A necessity for true population impact. JAMA Psychiatry 2016 May 01;73(5):435-436. [doi: 10.1001/jamapsychiatry.2016.0001] [Medline: 26982348]
9. Campo JV. Youth suicide prevention: Does access to care matter? Curr Opin Pediatr 2009;21(5):628-634. [doi: 10.1097/mop.0b013e32833069bd]
10. Bruffaerts R, Demyttenaere K, Hwang I, Chiu W, Sampson N, Kessler RC, et al. Treatment of suicidal people around the world. Br J Psychiatry 2011 Jul;199(1):64-70 [FREE Full text] [doi: 10.1192/bjp.bp.110.084129] [Medline: 21263012]
11. Hilty DM, Chan S, Hwang T, Wong A, Bauer AM. Advances in mobile mental health: Opportunities and implications for the spectrum of e-mental health services. Mhealth 2017;3:34 [FREE Full text] [doi: 10.21037/mhealth.2017.06.02] [Medline: 28894744]
12. Selby EA, Yen S, Spirito A. Time varying prediction of thoughts of death and suicidal ideation in adolescents: Weekly ratings over 6-month follow-up. J Clin Child Adolesc Psychol 2013;42(4):481-495 [FREE Full text] [doi: 10.1080/15374416.2012.736356] [Medline: 23148530]
13. Kleiman EM, Turner BJ, Fedor S, Beale EE, Huffman JC, Nock MK. Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. J Abnorm Psychol 2017 Aug;126(6):726-738. [doi: 10.1037/abn0000273] [Medline: 28481571]
14. Nock MK, Prinstein MJ, Sterba SK. Revealing the form and function of self-injurious thoughts and behaviors: A real-time ecological assessment study among adolescents and young adults. J Abnorm Psychol 2009 Nov;118(4):816-827 [FREE Full text] [doi: 10.1037/a0016948] [Medline: 19899851]
15. Kleiman EM, Nock MK. Real-time assessment of suicidal thoughts and behaviors. Curr Opin Psychol 2018 Aug;22:33-37. [doi: 10.1016/j.copsyc.2017.07.026] [Medline: 30122275]
16. Kleiman EM, Turner BJ, Fedor S, Beale EE, Picard RW, Huffman JC, et al. Digital phenotyping of suicidal thoughts. Depress Anxiety 2018 Jul;35(7):601-608. [doi: 10.1002/da.22730] [Medline: 29637663]
17. Millner AJ, Lee MD, Nock MK. Describing and measuring the pathway to suicide attempts: A preliminary study. Suicide Life Threat Behav 2017 Jun;47(3):353-369. [doi: 10.1111/sltb.12284] [Medline: 27477787]
18. Hallensleben N, Glaesmer H, Forkmann T, Rath D, Strauss M, Kersting A, et al. Predicting suicidal ideation by interpersonal variables, hopelessness and depression in real-time. An ecological momentary assessment study in psychiatric inpatients with depression. Eur Psychiatry 2019 Feb;56:43-50. [doi: 10.1016/j.eurpsy.2018.11.003] [Medline: 30530103]
19. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: Review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
20. Jaroszewski AC, Morris RR, Nock MK. Randomized controlled trial of an online machine learning-driven risk assessment and intervention platform for increasing the use of crisis services. J Consult Clin Psychol 2019 Apr;87(4):370-379. [doi: 10.1037/ccp0000389] [Medline: 30883164]
21. Pauwels K, Aerts S, Muijzers E, De Jaegere E, van Heeringen K, Portzky G. BackUp: Development and evaluation of a smart-phone application for coping with suicidal crises. PLoS One 2017;12(6):e0178144 [FREE Full text] [doi: 10.1371/journal.pone.0178144] [Medline: 28636617]
22. de la Torre I, Castillo G, Arambarri J, López-Coronado M, Franco MA. Mobile apps for suicide prevention: Review of virtual stores and literature. JMIR Mhealth Uhealth 2017 Oct 10;5(10):e130 [FREE Full text] [doi: 10.2196/mhealth.8036] [Medline: 29017992]
23. Kreuze E, Jenkins C, Gregoski M, York J, Mueller M, Lamis DA, et al. Technology-enhanced suicide prevention interventions: A systematic review. J Telemed Telecare 2016 Jul 03;23(6):605-617. [doi: 10.1177/1357633x16657928]

XSL•FO
RenderX

24. Nuij C, van Ballegooijen W, Ruwaard J, de Beurs D, Mokkenstorm J, van Duijn E, et al. Smartphone-based safety planning and self-monitoring for suicidal patients: Rationale and study protocol of the CASPAR (Continuous Assessment for Suicide Prevention And Research) study. Internet Interv 2018 Sep;13:16-23 [FREE Full text] [doi: 10.1016/j.invent.2018.04.005] [Medline: 30206514]
25. Christensen H, Batterham P, O'Dea B. E-health interventions for suicide prevention. Int J Environ Res Public Health 2014 Aug 12;11(8):8193-8212 [FREE Full text] [doi: 10.3390/ijerph110808193] [Medline: 25119698]
26. Inagaki M, Kawashima Y, Kawanishi C, Yonemoto N, Sugimoto T, Furuno T, et al. Interventions to prevent repeat suicidal behavior in patients admitted to an emergency department for a suicide attempt: A meta-analysis. J Affect Disord 2015 Apr 01;175:66-78. [doi: 10.1016/j.jad.2014.12.048] [Medline: 25594513]
27. Hogan MF. Better suicide screening and prevention are possible. JAMA Psychiatry 2016 Nov 01;73(11):1111-1112. [doi: 10.1001/jamapsychiatry.2016.2411] [Medline: 27706485]
28. Linehan MM, Comtois KA, Murray AM, Brown MZ, Gallop RJ, Heard HL, et al. Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. Arch Gen Psychiatry 2006 Jul;63(7):757-766. [doi: 10.3410/f.14771.472298]
29. Larsen ME, Nicholas J, Christensen H. A systematic assessment of smartphone tools for suicide prevention. PLoS One 2016 Apr 13;11(4):e0152285 [FREE Full text] [doi: 10.1371/journal.pone.0152285] [Medline: 27073900]
30. Melvin GA, Gresham D, Beaton S, Coles J, Tonge BJ, Gordon MS, et al. Evaluating the feasibility and effectiveness of an Australian safety planning smartphone application: A pilot study within a tertiary mental health service. Suicide Life Threat Behav 2019 Jun;49(3):846-858 [FREE Full text] [doi: 10.1111/sltb.12490] [Medline: 29999193]
31. Torous J, Larsen ME, Depp C, Cosco TD, Barnett I, Nock MK, et al. Smartphones, sensors, and machine learning to advance real-time prediction and interventions for suicide prevention: A review of current progress and next steps. Curr Psychiatry Rep 2018 Jun 28;20(7):51. [doi: 10.1007/s11920-018-0914-y] [Medline: 29956120]
32. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: A new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
33. Sheehan D, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34 [FREE Full text] [Medline: 9881538]
34. Rush A, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep;54(5):573-583. [doi: 10.1016/s0006-3223(02)01866-8]
35. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011 Dec;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
36. Weisman AD, Worden JW. Risk-rescue rating in suicide assessment. Arch Gen Psychiatry 1972 Jun;26(6):553-560. [doi: 10.1001/archpsyc.1972.01750240065010] [Medline: 5027119]
37. Beck AT, Kovacs M, Weissman A. Assessment of suicidal intention: The Scale for Suicide Ideation. J Consult Clin Psychol 1979;47(2):343-352. [doi: 10.1037/0022-006x.47.2.343]
38. Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. Child Abuse Negl 2003 Feb;27(2):169-190. [doi: 10.1016/s0145-2134(02)00541-0]
39. Russell D, Peplau LA, Cutrona CE. The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. J Pers Soc Psychol 1980 Sep;39(3):472-480. [doi: 10.1037/0022-3514.39.3.472]
40. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. Proc Natl Acad Sci U S A 2013 Apr 09;110(15):5797-5801. [doi: 10.1073/pnas.1219686110] [Medline: 23530191]
41. Beck AT, Weissman A, Lester D, Trexler L. The measurement of pessimism: The hopelessness scale. J Consult Clin Psychol 1974 Dec;42(6):861-865. [doi: 10.1037/h0037562] [Medline: 4436473]
42. Linehan MM, Goodstein JL, Nielsen SL, Chiles JA. Reasons for staying alive when you are thinking of killing yourself: The Reasons for Living Inventory. J Consult Clin Psychol 1983;51(2):276-286. [doi: 10.1037/0022-006x.51.2.276]
43. Whiteside SP, Lynam DR, Miller JD, Reynolds SK. Validation of the UPPS impulsive behaviour scale: A four-factor model of impulsivity. Eur J Pers 2005 Dec;19(7):559-574. [doi: 10.1002/per.556]
44. Dan-Glauser ES, Scherer KR. The Difficulties in Emotion Regulation Scale (DERS). Swiss J Psychol 2013 Jan;72(1):5-11. [doi: 10.1024/1421-0185/a000093]
45. Guédeney N, Fermanian J, Bifulco A. Construct validation study of the Relationship Scales Questionnaire (RSQ) on an adult sample [Article in French]. Encephale 2010 Feb;36(1):69-76. [doi: 10.1016/j.encep.2008.12.006] [Medline: 20159199]
46. Van Orden KA, Cukrowicz KC, Witte TK, Joiner TE. Thwarted belongingness and perceived burdensomeness: Construct validity and psychometric properties of the Interpersonal Needs Questionnaire. Psychol Assess 2012 Mar;24(1):197-215 [FREE Full text] [doi: 10.1037/a0025358] [Medline: 21928908]

XSL•FO
RenderX

47. Ware JE, Sherbourne CD. The MOS 36-ltem Short-Form Health Survey (SF-36). Med Care 1992;30(6):473-483. [doi: 10.1097/00005650-199206000-00002]
48. Morgan J, Farsides T. Measuring meaning in life. J Happiness Stud 2007 Sep 25;10(2):197-214. [doi: 10.1007/s10902-007-9075-0]
49. Maheu MM, Nicolucci V, Pulier ML, Wall KM, Frye TJ, Hudlicka E. The Interactive Mobile App Review Toolkit (IMART): A clinical practice-oriented system. J Technol Behav Sci 2016 Dec 15;1(1-4):3-15. [doi: 10.1007/s41347-016-0005-z]
50. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. Acad Emerg Med 2015 Jun;22(6):754-756 [FREE Full text] [doi: 10.1111/acem.12692] [Medline: 25997375]
51. Baker TB, Gustafson DH, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. J Med Internet Res 2014 Feb 19;16(2):e36 [FREE Full text] [doi: 10.2196/jmir.2925] [Medline: 24554442]
52. Biagianti B, Hidalgo-Mazzei D, Meyer N. Developing digital interventions for people living with serious mental illness: Perspectives from three mHealth studies. Evid Based Ment Health 2017 Nov;20(4):98-101 [FREE Full text] [doi: 10.1136/eb-2017-102765] [Medline: 29025862]
53. Hawton K, Witt KG, Taylor Salisbury TL, Arensman E, Gunnell D, Townsend E, et al. Interventions for self-harm in children and adolescents. Cochrane Database Syst Rev 2015 Dec 21(12):CD012013. [doi: 10.1002/14651858.CD012013] [Medline: 26688129]
54. Stallard P, Porter J, Grist R. A smartphone app (BlueIce) for young people who self-harm: Open phase 1 pre-post trial. JMIR Mhealth Uhealth 2018 Jan 30;6(1):e32 [FREE Full text] [doi: 10.2196/mhealth.8917] [Medline: 29382632]
55. Franco-Martín MA, Muñoz-Sánchez JL, Sainz-de-Abajo B, Castillo-Sánchez G, Hamrioui S, de la Torre-Díez I. A systematic literature review of technologies for suicidal behavior prevention. J Med Syst 2018 Mar 05;42(4):71. [doi: 10.1007/s10916-018-0926-5] [Medline: 29508152]
56. Krueger RA, Casey AM. Focus Groups: A Practical Guide for Applied Research. 5th edition. Thousand Oaks, CA: SAGE Publications; 2015.
57. Santini ZI. The Impact of Social Networks and Social Support on Mental Disorders and Mortality [doctoral thesis]. Barcelona, Spain: University of Barcelona; 2016. URL: https://portal.findresearcher.sdu.dk/files/123588781/Ziggi_Santini_The_impact_of_social_networks_and_social_support_on_mental_disorders_and_mortality.pdf [accessed 2020-09-27]
58. Whitlock J, Wyman PA, Moore SR. Connectedness and suicide prevention in adolescents: Pathways and implications. Suicide Life Threat Behav 2014 Jun;44(3):246-272. [doi: 10.1111/sltb.12071] [Medline: 24444252]
59. Nurullah AS. Received and provided social support: A review of current evidence and future directions. Am J Health Stud 2012;27(3):173-188 [FREE Full text]
60. McDougall MA, Walsh M, Wattier K, Knigge R, Miller L, Stevermer M, et al. The effect of social networking sites on the relationship between perceived social support and depression. Psychiatry Res 2016 Dec 30;246:223-229. [doi: 10.1016/j.psychres.2016.09.018] [Medline: 27721061]
61. Conejero I, Olié E, Calati R, Ducasse D, Courtet P. Psychological pain, depression, and suicide: Recent evidences and future directions. Curr Psychiatry Rep 2018 Apr 05;20(5):33. [doi: 10.1007/s11920-018-0893-z] [Medline: 29623441]
62. Morgiève M, Ung Y, Gehamy C, Briffault X. Diminuer l'impact des troubles obsessionnels compulsifs par des modifications de l'environnement physique: Une étude de preuve de concept (Decrease the impact of obsessive-compulsive disorders through changes in the physical environment: A proof-of-concept study). Psychiatr Sci Hum Neurosci 2016;14:43-63 [FREE Full text] [doi: 10.3917/psn.143.0043]
63. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
64. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: A systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
65. Smith W, Ploderer B, Wadley G, Webber S, Borland R. Trajectories of engagement and disengagement with a story-based smoking cessation app. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17). 2017 May Presented at: 2017 CHI Conference on Human Factors in Computing Systems (CHI '17); May 6-11, 2017; Denver, CO p. 3045-3056. [doi: 10.1145/3025453.3026054]
66. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: Systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
67. Andreasson K, Krogh J, Bech P, Frandsen H, Buus N, Stanley B, et al. MYPLAN—mobile phone application to manage crisis of persons at risk of suicide: Study protocol for a randomized controlled trial. Trials 2017 Apr 11;18(1):171 [FREE Full text] [doi: 10.1186/s13063-017-1876-9] [Medline: 28399909]
68. Coppersmith G, Ngo K, Leary R, Wood A. Exploratory analysis of social media prior to a suicide attempt. In: Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology. 2016 Jun Presented at: Third Workshop on

XSL•FO
RenderX

Computational Linguistics and Clinical Psychology; June 16, 2016; San Diego, CA p. 106-117 URL: https://www.aclweb.org/anthology/W16-0311.pdf [doi: 10.18653/v1/w16-0311]
69. Spangenberg L, Forkmann T, Glaesmer H. Investigating dynamics and predictors of suicidal behaviors using ambulatory assessment. Neuropsychiatr 2015;29(3):139-143. [doi: 10.1007/s40211-015-0142-1] [Medline: 25868681]
70. Davidson CL, Anestis MD, Gutierrez PM. Ecological momentary assessment is a neglected methodology in suicidology. Arch Suicide Res 2017 Jan 02;21(1):1-11. [doi: 10.1080/13811118.2015.1004482] [Medline: 26821811]
71. Bagge CL, Littlefield AK, Glenn CR. Trajectories of affective response as warning signs for suicide attempts. Clin Psychol Sci 2017 Mar 13;5(2):259-271. [doi: 10.1177/2167702616681628]
72. Briffault X, Morgiève M, Courtet P. From e-Health to i-Health: Prospective reflexions on the use of intelligent systems in mental health care. Brain Sci 2018 May 31;8(6):98 [FREE Full text] [doi: 10.3390/brainsci8060098] [Medline: 29857495]
73. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIS) in mobile health: Key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462 [FREE Full text] [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
74. Linthicum KP, Schafer KM, Ribeiro JD. Machine learning in suicide science: Applications and ethics. Behav Sci Law 2019 May;37(3):214-222. [doi: 10.1002/bsl.2392] [Medline: 30609102]
75. Kilsdonk E, Peute L, Jaspers M. Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps analysis. Int J Med Inform 2017 Feb;98:56-64. [doi: 10.1016/j.ijmedinf.2016.12.001] [Medline: 28034413]

Abbreviations

ANSM: Agence Nationale de Sécurité du Médicament et des Produits de Santé (French National Agency for Medicines and Health Products Safety)
C-SSRS: Columbia-Suicide Severity Rating Scale
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
ED: emergency department
EMA: ecological momentary assessment
EMI: ecological momentary intervention
EMMA: Ecological Mental Momentary Assessment
M0: month 0 (inclusion)
M1: month 1
M3: month 3
M6: month 6
mHealth: mobile health
NERB: Neurophysiology of Repetitive Behaviors
SAMU: Service d'Aide Médicale d'Urgence
WHO: World Health Organization

Edited by G Eysenbach, R Leung; submitted 19.09.19; peer-reviewed by L Giner, D Hilty, J Sussman, B Renn, G Klein, D Bakker, M Businelle; comments to author 12.02.20; revised version received 08.04.20; accepted 03.06.20; published 09.10.20.

Please cite as:
Morgiève M, Genty C, Azé J, Dubois J, Leboyer M, Vaiva G, Berrouiguet S, Courtet P
A Digital Companion, the Emma App, for Ecological Momentary Assessment and Prevention of Suicide: Quantitative Case Series Study
JMIR Mhealth Uhealth 2020;8(10):e15741
URL: https://mhealth.jmir.org/2020/10/e15741
doi:10.2196/15741
PMID:33034567

©Margot Morgiève, Catherine Genty, Jérôme Azé, Jonathan Dubois, Marion Leboyer, Guillaume Vaiva, Sofian Berrouiguet, Philippe Courtet. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 09.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Occurrence of and Reasons for "Missing Events" in Mobile Dietary Assessments: Results From Three Event-Based Ecological Momentary Assessment Studies

Katrin Ziesemer[1], Dipl-Troph; Laura Maria König[1], PhD; Carol Jo Boushey[2], PhD; Karoline Villinger[1], PhD; Deborah Ronja Wahl[1], PhD; Simon Butscher[3], PhD; Jens Müller[3], PhD; Harald Reiterer[3], PhD; Harald Thomas Schupp[4], PhD; Britta Renner[1], PhD

[1]Psychological Assessment & Health Psychology, Department of Psychology, University of Konstanz, Konstanz, Germany

[2]University of Hawaii Cancer Center, University of Hawaii, Honolulu, HI, United States

[3]Human-Computer Interaction Group, Department of Computer and Information Science, University of Konstanz, Konstanz, Germany

[4]General Psychology, Department of Psychology, University of Konstanz, Konstanz, Germany

Corresponding Author:
Britta Renner, PhD
Psychological Assessment & Health Psychology
Department of Psychology
University of Konstanz
PO Box 47
Konstanz, 78457
Germany
Phone: 49 7531884679
Fax: 49 7531 88 5226
Email: britta.renner@uni-konstanz.de

Abstract

Background: Establishing a methodology for assessing nutritional behavior comprehensively and accurately poses a great challenge. Mobile technologies such as mobile image-based food recording apps enable eating events to be assessed in the moment in real time, thereby reducing memory biases inherent in retrospective food records. However, users might find it challenging to take images of the food they consume at every eating event over an extended period, which might lead to incomplete records of eating events (*missing events*).

Objective: Analyzing data from 3 studies that used mobile image-based food recording apps and varied in their technical *enrichment*, this study aims to assess how often eating events (meals and snacks) were missed over a period of 8 days in a naturalistic setting by comparing the number of recorded events with the number of normative expected events, over time, and with recollections of missing events.

Methods: Participants in 3 event-based Ecological Momentary Assessment (EMA) studies using mobile image-based dietary assessments were asked to record all eating events (study 1, N=38, 1070 eating events; study 2, N=35, 934 eating events; study 3, N=110, 3469 eating events). Study 1 used a *basic* app; study 2 included 1 fixed reminder and the possibility to add meals after the actual eating events occurred instead of in the moment (*addendum*); and study 3 included 2 fixed reminders, an addendum feature, and the option to record skipped meals. The number of recalled missed events and their reasons were assessed by semistructured interviews after the EMA period (studies 1 and 2) and daily questionnaires (study 3).

Results: Overall, 183 participants reported 5473 eating events. Although the momentary adherence rate as indexed by a comparison with normative expected events was generally high across all 3 studies, a differential pattern of results emerged with a higher rate of logged meals in the more technically intensive study 3. Multilevel models for the logging trajectories of reported meals in all 3 studies showed a significant, albeit small, decline over time (b=−.11 to −.14, *P*s<.001, pseudo-R^2=0.04-0.06), mainly because of a drop in reported snacks between days 1 and 2. Intraclass coefficients indicated that 38% or less of the observed variance was because of individual differences. The most common reasons for missing events were competing activities and technical issues, whereas situational barriers were less important.

XSL•FO
RenderX

Conclusions: Three different indicators (normative, time stability, and recalled missing events) consistently indicated missing events. However, given the intensive nature of diet EMA protocols, the effect sizes were rather small and the logging trajectories over time were remarkably stable. Moreover, the individual's actual state and context seemed to exert a greater influence on adherence rates than stable individual differences, which emphasizes the need for a more nuanced understanding of the factors that affect momentary adherence.

(JMIR Mhealth Uhealth 2020;8(10):e15430) doi:10.2196/15430

KEYWORDS

dietary assessment; diet records; mobile phone; mobile applications; technology; adherence; compliance; missing events; Ecological Momentary Assessment; mHealth

Introduction

Eating may seem to be one of the simplest behaviors, yet it is quite complex [1-3], involving up to 200 decisions a day [4]. Although methodological challenges exist in accurately measuring food intake in community dwellings [5,6], analyzing data on food intake is important for our understanding of diet-disease relationships and the refinement of nutrition guidelines [7], which provide essential information for surveillance, planning interventions, and policy [8].

The limitations associated with common measuring methods, such as self-reported food intake (eg, dietary records, 24-hour dietary recall, and food frequency questionnaires), which include memory biases and other measurement errors, have already been outlined [5,9-15]. However, new mobile technologies have the potential to reduce the burden on both researchers and participants by improving adherence and communication, automating and standardizing coding, and upgrading data quality ([16]; for an overview see study by Boushey et al and Eldridge et al [5,17]). Specifically, taking images of eating events has been proposed as a method for reducing the burden on participants in dietary studies. Various models based on image technology using mobile apps have been developed in recent years (eg, Technology Assisted Dietary Assessment, My Meal Mate, and SMARTFOOD; for more details, see study by Boushey et al [5], Eldridge et al [7], Villinger et al [18], and Wahl et al [2]), and they are increasingly being used to assess and change eating behavior and food intake in different populations, including patients and generally healthy adults or adolescents [19,20]. As with any dietary assessment method, the more eating events are captured, the more thorough the analysis, and in an ideal study, the user would capture all eating events including an image before and, if applicable, after consumption. Therefore, adherence during the study period is key to a valid assessment of the actual eating behavior.

In previous research, the focus was mainly on general adherence rates, such as dropout rates indexed by the number of participants who did not complete the whole study period, or missing value rates for study variables indexed by the number of participants who did not respond to some of the study variables. For example, a smartphone electronic food diary app called My Meal Mate showed higher adherence rates and overall satisfaction than a conventional pen-and-paper food diary [21]. However, although these study adherence indices capture between-person differences, they do not necessarily reflect the within-person adherence rate during the study, which is essential for a valid mobile image-based assessment of dietary intake (see also [22-24]). Therefore, the aim of this study was to assess both between- and within-person adherence rates for mobile image-based dietary assessments (MIDAs).

Research in the field of Ecological Momentary Assessment (EMA) [24-27] suggests that mobile dietary assessments can be classified as *event-based* monitoring, in which assessments are triggered by the occurrence of a predefined event of interest [28]. Typically, the participants themselves determine when the event has occurred and initiate an assessment [29,30], although some approaches are also developing methods to automatically detect an eating event using mobile sensing (eg, eButton [31] or AIM (automatic ingestion monitor) [32]). As the number of eating events can differ both between and within participants over time, assessing whether events have occurred that were not recorded (*missing events*) is a challenge.

Schembre et al [24] recently reviewed 20 mobile diet EMA studies and found that none of the 10 event-based studies that were included provided data related to adherence to dietary data collection protocols or the number of eating events captured (2 studies reported intake events but summarized eating and drinking occasions [33,34]). In addition, Maugeri and Barchitta [35] reviewed 54 articles and concluded that most mobile diet EMA studies did not report user response or compliance rates. Operationalizing missing event rates is particularly cumbersome because adhering to an event-based dietary collection protocol requires all eating events to be actively reported by the participants (*momentary adherence*).

We suggest a synopsis of 3 different indicators to approximate the rate of missing events in MIDAs. First, one criterion for roughly estimating the number of missing events is to compare the number of eating events reported with the number of events expected based on social norms and observational data. In Western countries, a pattern of 3 meals per day (ie, breakfast, lunch, and dinner) has become normative since the 19th century (Grignon and Grignon, 2004 [36] in French cited by Lhuissier [37]), which suggests that participants should record at least 3 eating events per day. However, observational data derived from 24-hour dietary recalls suggest at least 4 eating episodes per day. Specifically, in the United States, an average of 4.3 eating occasions was recorded per day, of which 3 were main meals [38-40]. Similarly, in Germany, the World Health Organization MONICA (MONItoring CArdiovascular disease) study showed an average of 3 main meals plus 1 snack per day, using a 7-day food intake protocol [40]. Corroborating these findings, a substudy of the European Prospective Investigation into Cancer

XSL•FO
RenderX

and Nutrition (EPIC)-Potsdam cohort observed 4 peaks in food consumption, that is, breakfast, lunch, afternoon snack, and dinner [39]. Similarly, assessing eating behavior in the moment using an Ecological Momentary Assessment revealed an average of 3.65 eating occasions per participant across 2 weeks [18]. Using 3- and 4-meal patterns as social comparison standards allows the overall rate of missing events to be estimated with respect to social norms and observational data. Second, logging trajectories and the stability of recorded eating events over time might also indicate the rate of missing events. Assuming that adherence motivation declines over time, the frequency of missing events should generally increase over time, which is mirrored in a declining overall logging trajectory. Moreover, one could also assume that missing events increase over time depending on the quality of the logged food (eg, decreased reporting of irregular meals, such as snacks), indicating selective reporting over time. Therefore, declining logging trajectories of eating events over time can be examined as an additional indicator for missing events. Third, another criterion could be to assess the number of perceived missing events as part of the user experience. However, to the best of our knowledge, no previous studies have assessed these three different indicators to estimate the rate of missing events.

Moreover, the mobile image-based food recording app itself might include features that have an impact on the missing event rates. They can vary in whether and how often they prompt the participants to record their eating events, which raises the question of whether mobile apps with reminders yield higher eating event logging rates, and thus lower missing event rates, than those without prompts. Similarly, the possibility of adding meals after the actual eating events occurred instead of in the moment of eating (referred to as *addendum*) might also have a positive impact on the rate of missing events. In addition, assessing skipped meals (ie, omission or lack of consumption of a meal) allows for differentiation between skipped meals and missing events, which might increase overall measurement precision.

Three EMA studies were conducted using a mobile image-based app to assess eating events in real life to investigate the occurrence of and perceived reasons for missing events (both meals and snacks) by (1) comparing the reported number of eating events with the number of social normative expected events, (2) calculating logging trajectories over time and in dependence of the type of meal, and (3) assessing the number of and reasons for perceived missing events as part of the user experience.

The measurement periods vary across different disciplines and depend on the goals of the study (eg, dietary assessment or intervention). Different methods (eg, food frequency questionnaire, 24-hour recalls, and weighed food records) are used for dietary assessment, which is the focus of this paper. Although assessment periods vary, they typically do not exceed 8 days [41]. For example, 3- to 5-day nonconsecutive 24-hour recalls or up to 7-day weighed food records are used to assess habitual intake. Sharp et al [16] summarized mobile phone–based dietary intake assessment studies, showing that assessment periods typically range from 1 to 7 days. Accordingly, we chose an 8-day assessment period and followed the recommendation of the Food and Agriculture Organization of the United Nations [41] to ensure that we included weekdays and weekends.

The 3 studies used apps with varying technical intensities, ranging from a *basic* MIDA to a more technically *enriched* app to get insights into the impact of additional technical features on logging rates. Study 1 investigated the number of recorded meals and snacks using a custom-programmed smartphone app in which participants were asked to select the meal type from 5 predefined options (breakfast, lunch, afternoon tea, dinner, and snack), take a picture of the eating event, and add a written description of the meal or snack. To develop a better understanding of the typology of missing events, the participants were invited to a semistructured interview in which they were asked to estimate the number of and reasons for missing events (*perceived missing events*) when the MIDA period was over.

Study 2 extended study 1 by adding reminders and making it possible to record eating events after they occurred instead of in the moment but still during the EMA period (*addenda*). As in study 1, semistructured interviews were conducted after the assessment period to probe the perceived number of and reasons for missing events.

In line with study 2, study 3 offered reminders and the possibility of recording eating events after they occurred (*addenda*). Extending studies 1 and 2, participants could also indicate whether they had skipped a meal or snack throughout the day (*skipped meal*). Moreover, the occurrence of and reasons for missing events were assessed on a daily basis during the mobile food record assessment period instead of at the end of the recording period.

In the 3 EMA studies, we examined the following hypotheses:

1. In accordance with the suggested synopsis of 3 different indicators to approximate the rate of missing events in mobile dietary assessments, we compared the number of logged eating events with the number of expected meals based on social norms and observational data, and the number of perceived missing events to test for an overall underreporting of eating events and an overall estimate of missing rate events. In addition, we examined overall and meal type–specific logging trajectories over time to test the hypothesis that adherence motivation decreases and missing events increase over time.
2. We determined whether the number of missing events can be lowered by additional technical features, such as reminders, and options such as adding meals after the actual eating events occurred (*addenda*) and recording skipped meals.
3. Finally, reported reasons for missing events were analyzed to provide relevant information for designing mobile diet apps.

Methods

Participants and Ethics

Participants were generally healthy adults (≥18 years) and volunteers recruited from the student and employee population

XSL•FO
RenderX

of the University of Konstanz. All participants were reimbursed for their participation. The ethics committee of the University of Konstanz approved the study protocols for all 3 studies. The participants provided written informed consent before enrollment, and all studies adhered to the guidelines of the German Psychological Society and the Declaration of Helsinki.

Procedure

The participants were recruited via leaflets distributed at the University of Konstanz and postings on Facebook groups. Participants in studies 2 and 3 were also recruited through the web-based recruitment platform for research studies at the University of Konstanz (SONA). Only registered users can sign up for studies via the platform.

The participants were invited to the laboratory for an introductory and closing session either individually (for studies 1 and 3) or in groups (for study 2). At the introductory session, after the participants had completed a questionnaire that assessed demographic variables and their dietary style, they were provided and familiarized with the MIDA.

Mobile Image-Based Dietary Assessment

The participants were asked to record all eating events, whether main meals or snacks, for 8 consecutive days. They were specifically asked to indicate the type of meal with the following 5 options: breakfast, lunch, afternoon tea, dinner, and snack. In Germany, *afternoon tea* is called *Kaffee und Kuchen*, which directly translates as *coffee and cake*. It is similar to the idea of a traditional *afternoon tea* meal in the United Kingdom. Specifically, in Germany, people have *Kaffee und Kuchen* in the afternoon (between 4 and 5 PM), typically serving coffee (or tea) with some cake or cookies. Afterward, participants were asked to take a picture of each eating event and provide a short description of the meal or snack (eg, pasta with tomato sauce or oats, milk, apple). Additional courses and leftovers were also recorded by taking pictures. A valid recorded eating event had to include all 3 aspects: (1) meal type, (2) eating event picture, and (3) description of the food. Incomplete entries, missing one or more aspects, were not coded as an eating event. Participants were asked to start recording their meals the day after the introductory session, which usually took place on Monday, Tuesday, or Wednesday. Starting days were limited to the beginning of the week to ensure that any potential issues raised by the participants when they started using the app could be addressed before the weekend began, thus ensuring that data collection would take place both during the week and on weekends.

Assessment of Perceived Missing Events: Semistructured Interviews and Open Questions

In studies 1 and 2, the participants were invited back to the laboratory after 8 days of recording for a semistructured interview, which was conducted by 4 trained interviewers (KZ, LK, KV, and DW) to assess the occurrences of and reasons for perceived missing events. All interviews were voice recorded. The participants were asked in an open-question format whether they remembered any occasions during the 8 days of the study period when they did not record a main meal (breakfast, lunch, or dinner) or a snack. If they reported that they had failed to record a main meal or a snack (*perceived missing event*), they were asked to state how often this had happened and why. In study 1, they were also asked about any situational constraints that might have prevented them from recording a particular meal or snack. In study 3, the reasons for missing events were assessed in an open-question format via a web-based questionnaire that could be accessed at any time by pressing a button on the app's home screen. The participants were also reminded of this option at the end of each EMA day. At the beginning of the study, participants could decide for themselves the time in the evening when the reminder would be sent.

Data Analysis

Statistical analyses for the EMA data were conducted with IBM SPSS (Version 25) and R 3.2.3, using the packages lme4 1.1-11 [42] and lmerTest 2.0-30 [43]. Data were analyzed separately for each of the 3 studies following the same procedure by examining the three different indicators to quantify the rate of missing events in mobile dietary assessments (normative, time stability, and recalled missing events).

First, the number of reported meals was compared with the normative number of expected meals using one-sample *t* tests. The values 24 (for 3 daily meals including breakfast, lunch, and dinner) and 32 (for 4 daily meals, additionally including a snack) were used as normative criteria. The number of logged meals was compared between meal types using repeated-measures analyses of variance (ANOVAs) and followed up by dependent sample *t* tests, for which alpha was lowered to .008 to account for multiple comparisons. Intraindividual differences in logging rates were examined using intraclass correlations (ICCs), which estimate the proportion of variance explained by the participants compared with the total variance.

Second, assuming that adherence motivation declines over time, changes in the number of reported meals over time were analyzed. Multilevel linear modeling [44] was used to account for the data's hierarchical structure. Logged eating events per day (lower level/level 1) were nested within participants (higher level/level 2). The number of logged eating events was modeled as a function of time within participants to test whether the number of logged eating events changed over time. Separate models were computed for the total number of eating events and for each meal type. First, a random slopes model allowing both intercept and slope to vary was computed to model whether the participants differed both in the mean number of logged eating events and in the relationship between the number of reported events and time. Second, a random intercept model that allowed only the intercept to vary was computed to model whether the participants differed only in the mean number of logged eating events and not in the relationship between the number of logged events in time. If significant, both models were compared using a deviance test [44]. A nonsignificant deviance test indicates that the less complex model (ie, random intercept model) is preferred, whereas a significant deviance test indicates that the more complex model (ie, random slopes model) is preferred. Pseudo-R^2 was computed as recommended by Raudenbush and Bryk [45].

Third, the number of and reasons for perceived missing events assessed through semistructured interviews (studies 1 and 2)

XSL•FO
RenderX

and open questions (study 3) were coded by 2 authors (KZ and LK) and a third independent researcher using a standardized manual developed by KZ, LK, BR, and CB, with input from all authors. The number of missing events for meals and snacks was compared using dependent samples *t* tests.

Study 1 (8 Days of Mobile Food Recording): Basic Mobile Image-Based Dietary Assessment

The app used for the mobile image-based dietary assessment was programmed using the Android movisensXS app (version 0.8.4203, movisens GmbH Karlsruhe, Germany). The participants could either use their own smartphone or were provided with a study smartphone (ASUS Padfone Infinity, Android 5.0.2, n=10).

Study 2 (8 Days of Mobile Food Recording): The Effect of Reminders and Addenda

Other than a few amendments, the procedure was identical to that of study 1. The participants were provided with a study smartphone (ASUS Padfone Infinity, Android 5.0.2), and the MIDA was realized with the SMARTFOOD app, which was developed as part of the research project SMARTACT [2,18,46-48] and included a feature to set a reminder in the morning to record food intake. The participants were asked to set the reminder at the beginning of the study during the introduction session. The app also had an addendum feature to log eating events that participants missed to record in the moment of consumption, which was enabled on half of the devices (n=18). Anthropometric measures were assessed during the introductory session, and the participants received a booklet explaining how to use the smartphone and app for recording food intake.

Study 3 (8 Days of Mobile Food Recording): Reporting Missing and Skipped Events During the Assessment Period

The procedure was identical to those of study 1 and 2, with the following amendments. The optional study smartphone (n=58) was either an ASUS Padfone Infinity (Android 5.0.2) or a Samsung Galaxy J5 (Android 6.0.1), with a custom-programmed mobile app (version 0.8.4203, movisens GmbH Karlsruhe; Germany) preinstalled.

In the app, participants could indicate whether they had skipped a meal or snack, and they could record eating events later (*addenda*), including meal type, composition, and the reasons for missing logging it. Moreover, the participants were asked to set customized reminders in the morning and evening. Extending study 2, an evening reminder was sent to remind the participants to (1) record any missing event and the respective reasons, and (2) log any skipped meals or snacks.

Results

Study 1 (8 Days of Mobile Food Recording): Basic Mobile Image-Based Dietary Assessment

The aim of study 1 was to examine the number of and reasons for missing events during event-based mobile food recording in real life for 8 consecutive days according to 3 different indicators (normative, time stability, and recalled missing events). The study investigated the number of recorded meals and snacks using a custom-programmed smartphone app. Participants logged the meal type, captured an image of the eating event, and added a written description of the meal or snack (*eating event*). The app used a *basic* MIDA as it did not include reminders or the possibility of adding eating events after they had occurred (*addenda*) or recorded skipped meals.

The 38 participants (28/38 female, 74%; 33/38 students, 87%) who took part in the study had a mean age of 24.5 years (SD 5.88, range 18-48). Of the 38 participants, 21 (55%) were omnivores, 7 (18%) vegetarians, 3 (8%) vegans, and 7 (18%) adhered to other dietary styles. As compensation, they either received course credits (2.0 h; n=10) or took part in a lottery to win one of 4 €25 (US $30; n=28) Amazon vouchers.

None of the participants dropped out of the study, indicating an excellent overall retention. The participants logged a total of 1099 eating events over the 8-day study period. Of these, 29 entries (2.64%) were canceled by the participant before completing the food recording, resulting in a total of 1070 recorded eating events (Table 1). Control analysis showed that participants who used their own vs a loaned smartphone did not differ significantly with respect to the number of reported meals ($ts_{36} \leq |1.46|$, $Ps \geq .15$, $ds \leq 0.49$) or the number of reported perceived missing events (main meals: $t_{32}=-1.11$, $P=.28$, $d=0.40$; snacks: $t_{29}=-0.52$, $P=.61$, $d=0.22$).

XSL•FO
RenderX

Table 1. Logged and social normative expected eating events by self-classified meal type over 8 days for studies 1, 2, and 3.

Meal type	Absolute number of meals (%)	Mean (SD)	Min/max	Difference observed-normative meals[a]	*t* value	*P* value	Cohen *d*	ICC[b]
Study 1 (N=38, *df*=37)								
Breakfast	244 (22.8)	6.42 (2.04)	2/11	−1.58	–4.78	<.001	0.78	.13
Lunch	213 (19.9)	5.61 (1.59)	1/9	−2.40	−9.31	<.001	1.51	.04
Dinner	256 (23.9)	6.74 (2.20)	0/11	−1.26	−3.54	.001	0.57	.09
Snacks	330 (30.8)	8.68 (6.97)	0/27	0.68	0.61	.55	0.10	.10
Afternoon tea	27 (2.5)	0.71 (1.18)	0/4	—[c]	—	—	—	.38
Total	1070 (100)	28.16 (9.57)	12/47	−3.84	−2.48	.02	0.40	.38
Study 2 (N=35, *df*=34)								
Breakfast	241 (25.8)	6.89 (2.37)	0/13	−1.11	−2.78	.01	0.47	.35
Lunch	211 (22.6)	6.03 (1.77)	1/9	−1.97	−6.58	<.001	1.11	.07
Dinner	234 (25.1)	6.69 (1.92)	1/10	−1.31	−4.05	<.001	0.68	.10
Snacks	222 (23.8)	6.34 (4.81)	0/17	−1.66	−2.04	.049	0.34	.30
Afternoon tea	26 (2.8)	0.74 (1.25)	0/6	—	—	—	—	.18
Total	934 (100)	26.69 (7.46)	11/42	−5.31	−4.21	<.001	0.71	.30
Study 3 (N=110, *df*=109)								
Breakfast	828 (23.9)	7.53 (2.58)	0/23	−0.47	−1.92	.06	0.18	.26
Lunch	744 (21.4)	6.76 (1.78)	1/10	−1.24	−7.30	<.001	0.70	.06
Dinner	832 (24.0)	7.56 (1.96)	1/13	−0.44	−2.34	.02	0.22	.06
Snacks	953 (27.5)	8.66 (6.23)	0/33	0.66	1.12	.27	0.11	.35
Afternoon tea	112 (3.2)	1.02 (1.11)	0/5	—	—	—	—	.02
Total	3469 (100)	31.54 (8.73)	10/61	−0.46	−0.56	.58	0.05	.32

[a]Reference *t* value was set to a value of 8 meals for individual meal types and to a value of 32 meals for total meals; negative values indicate fewer observed than normative expected number of meals.

[b]ICC: intraclass correlation.

[c]For afternoon tea no normative value was set and therefore no difference between observed and normative meals was calculated.

Logged Versus Social Normative Expected Number of Meals

On the group level and across eating events, the participants logged an average of 28.16 (SD 9.57) meals and snacks during the 8 study days. Thus, at the group level, the number of recorded eating events concurred with the social normative expected number of meals, which ranged between 24 and 32 (24 for 3 daily meals including breakfast, lunch, and dinner and 32 for 4 daily meals including an additional snack). A considerable variability in the number of entries emerged at the person level, ranging from 12 to 47 entries. However, the majority (23/38, 61%) logged 24 or more eating events during the study period. Testing further intraindividual differences in the overall logging rate yielded an ICC of ρ=0.38 across eating events, indicating that 62% of the overall variation in logged occasions was because of variation within participants, rather than variations in the logging rates between individuals.

As Table 1 shows, across the 8 days, the participants logged 244 breakfasts, 213 lunches, 27 afternoon teas, 256 dinners, and 330 snacks. On average, they recorded 6.42 (SD 2.04) breakfasts, 5.16 (SD 1.59) lunches, 6.74 (SD 2.20) dinners, 8.68 (SD 6.97) snacks, and 0.71 (SD 1.18) afternoon teas. Comparing the average number of recorded main meals (breakfast, lunch, and dinner) and snacks with the social normative expected number shows that the number of main meals and snacks recorded over the study period was significantly lower than the normative expected figure of 8, t_{37}=−2.48, *P*=.02, *d*=0.40. Moreover, a repeated-measures ANOVA (Greenhouse-Geisser corrected, ε=.45) with the factor *meal type* (breakfast, lunch, dinner, and snack) yielded a significant main effect with $F_{1.35,49.76}$=5.08, *P*=.02, partial η^2=.12, indicating that the number of logged meals varied depending on the meal type. Subsequent *t* tests indicated that significantly fewer lunches than dinners (t_{37}=−3.30, *P*=.002, *d*=0.59) and snacks (t_{37}=−2.90, *P*=.006, *d*=0.61) were recorded. No other comparisons were statistically significant, with *t*s<|2.17|, *P*s≥.04, exceeding the predetermined α=.008 to correct for multiple comparisons. Examining intraindividual differences in logging rates for the different main meals and snacks yielded ICC coefficients of ρ≤0.38, suggesting that the nesting of logged events within individuals was not

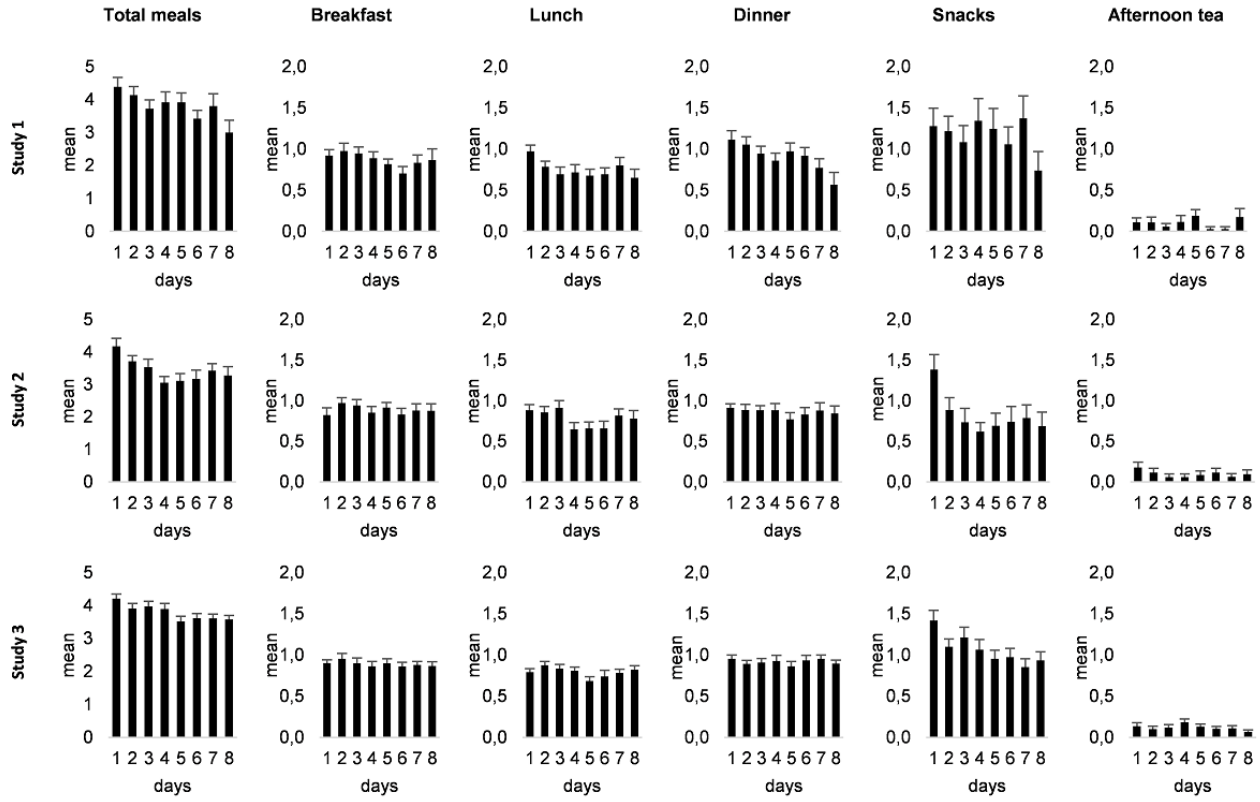
XSL•FO
RenderX

substantial. Hence, overall, interindividual differences in logging rates were very small.

Logging Trajectories of Eating Events Over Time

Logging trajectories over time were tested using multilevel modeling. Specifically, models were computed to test whether there was a significant linear change in the total number of logged eating events across the 8-day study period (Figure 1 and Table 2).

Figure 1. Average logging trajectories across the 8-day study period for studies 1, 2 and 3 across all meals and by meal types. Bars represent mean logged meals; error bars represent the standard error of the mean.



XSL•FO
RenderX

Table 2. Multilevel models for logging trajectories across the 8-day study period across all eating events and by meal types for studies 1, 2, and 3.

Characteristics	Random slopes model (fixed effects)					Random intercept model (fixed effects)				
	b	SE	*t*	*df*	*P* value	b	SE	*t*	*df*	*P* value
Study 1 (N=38)										
Model 1: Across meal types										
Intercept	4.39	0.24	18.01	34.73	<.001	4.39	0.26	16.76	98.74	<.001
Days	−0.14	0.04	−3.09	35.26	.004	−0.14	0.04	−3.52	235.91	<.001
Model 2: Breakfast										
Intercept	0.97	0.07	13.22	35.73	<.001	0.97	0.07	13.85	180.70	<.001
Days	−0.02	0.02	−1.28	36.14	.21	−0.02	0.01	−1.65	237.44	.10
Model 3: Lunch										
Intercept	0.86	0.07	11.60	36.12	<.001	0.87	0.07	12.66	232.83	<.001
Days	−0.03	0.02	−1.59	36.22	.12	−0.03	0.01	−1.99	240.80	.048
Model 4: Afternoon tea										
Intercept	0.11	0.05	2.03	41.17	.049	0.11	0.05	2.20	193.69	.023
Days	−0.00	0.01	−0.19	102.55	.85	−0.00	0.01	−0.17	235.84	.86
Model 5: Dinner										
Intercept	1.15	0.08	13.93	78.52	<.001	1.15	0.08	13.77	204.69	<.001
Days	−0.06	0.02	−3.49	230.13	<.001	−0.06	0.02	−3.50	239.04	<.001
Model 6: Snacks										
Intercept	1.31	0.16	8.02	64.39	<.001	1.30	0.20	6.45	99.12	<.001
Days	−0.04	0.03	−1.08	81.43	.28	−0.03	0.03	−1.05	235.58	.30
Study 2 (N=35)										
Model 1: Across meal types										
Intercept	3.94	0.18	21.60	33.33	<.001	3.93	0.20	19.73	100.64	<.001
Days	−0.12	0.03	−3.80	34.10	<.001	−0.12	0.03	−3.86	236.57	<.001
Model 2: Breakfast										
Intercept	0.91	0.07	12.29	34.56	<.001	0.91	0.07	13.90	97.40	<.001
Days	−0.01	0.01	−0.43	34.77	.67	−0.01	0.01	−0.53	237.20	.60
Model 3: Lunch										
Intercept	0.87	0.06	13.92	225.78	<.001	0.87	0.07	13.00	213.60	<.001
Days	0.02	0.01	−1.60	67.86	.12	−0.02	0.01	−1.65	238.35	.10
Model 4: Afternoon tea										
Intercept	0.13	0.05	2.46	34.35	.02	0.13	0.04	3.14	155.22	.002
Days	−0.01	0.01	−0.89	33.93	.38	−0.01	0.01	−1.10	237.97	.27
Model 5: Dinner										
Intercept	0.91	0.06	15.26	100.36	<.001	0.91	0.06	14.56	187.11	<.001
Days	−0.01	0.01	−1.00	166.84	.32	−0.01	0.01	−0.98	234.57	.33
Model 6: Snacks										
Intercept	1.12	0.14	7.88	40.09	<.001	1.12	0.14	7.95	104.59	<.001
Days	−0.07	0.02	−3.23	236.17	.001	−0.07	0.02	−3.23	236.87	.001
Study 3 (N=110)										
Model 1: Across meal types										
Intercept	4.49	0.14	32.81	125.39	<.001	4.49	0.13	34.57	313.44	<.001

XSL•FO
RenderX

Characteristics	Random slopes model (fixed effects)					Random intercept model (fixed effects)				
	b	SE	*t*	*df*	*P* value	b	SE	*t*	*df*	*P* value
Days	−0.11	0.02	−5.47	635.68	<.001	−0.11	0.02	−5.52	753.99	<.001
Model 2: Breakfast										
Intercept	1.01	0.04	23.89	145.02	<.001	1.01	0.04	24.10	382.80	<.001
Days	−0.01	0.01	−1.84	751.63	.07	−0.01	0.01	−1.84	754.77	.07
Model 3: Lunch										
Intercept	0.89	0.04	25.30	737.26	<.001	0.9	0.04	24.30	710.03	<.001
Days	−0.01	0.01	−0.86	378.69	.39	−0.01	0.01	−0.89	755.33	.38
Model 4: Afternoon tea										
Intercept	0.16	0.03	5.42	131.39	<.001	0.16	0.03	5.86	770.79	<.001
Days	−0.01	0.01	−1.33	251.58	.18	−0.01	0.01	−1.36	753.21	.17
Model 5: Dinner										
Intercept	0.98	0.04	24.43	344.77	<.001	0.98	0.04	24.24	700.66	<.001
Days	−0.00	0.01	−0.45	740.08	.65	−0.00	0.01	−0.45	753.67	.65
Model 6: Snacks										
Intercept	1.44	0.11	13.66	119.67	<.001	1.45	0.10	14.18	290.82	<.001
Days	−0.08	0.01	−5.32	523.12	<.001	−0.08	0.01	−5.41	753.06	<.001

Overall, there was a small but statistically significant negative trend over time (b=−.14, $t_{235.91}$=−3.52, *P*<.001, pseudo-R^2=0.05), indicating that the number of logged eating events decreased over time. However, as the participants did not vary with respect to the time trend as the random intercept model was preferred (χ^2_2=3.5, *P*=.18), the trend was generalizable across participants.

For individual meal types, small and significant negative time trends emerged for lunches (b=−.03, $t_{240.80}$=−1.99, *P*=.048, pseudo-R^2=0.01) and dinners (b=−.06, $t_{239.04}$=−3.50, *P*<.001, pseudo-R^2=0.04). Accordingly, the number of logged lunches and dinners decreased across the 8 consecutive days. Again, the random intercept model was preferred (χ^2_2≤3.7, *P*s≥.16), indicating that the observed time trends were comparable between participants. There was no significant change in logging frequency over time for other meal types, including snacks, breakfasts, and afternoon teas (Table 2).

Number of Reasons for Perceived Missing Events

In total, 35 of the 38 participants (92%) reported in the interview that they had missed logging at least one eating event during the study period, with a range from 1 to 14 (median 2.50, mean 4.47, SD 3.95; 5 participants could not specify the frequency of missing events). Moreover, 30 of these 35 participants (86%; 4 did not specify the frequency) stated that they had missed reporting at least one snack, and 22 of 35 (63%; one did not specify the frequency) had missed at least one main meal. Perceived missing event rates ranged between 1 and 9 for main meals (median 1.00, mean 1.73, SD 2.26) and between 1 and 8 for snacks (median 1.5, mean 2.73, SD 2.63), t_{29}=−1.89, *P*=.07, *d*=0.41.

In total, 69 different reasons were provided by the 35 participants who reported at least one missing event. Of these, 12 participants specified 1 reason, 17 specified 2, and 6 gave between 3 and 5 reasons (median 1.97, SD 1.0).

Of these 69 reasons, 60 provided information that went beyond merely stating that an eating event was missed, and these were categorized into 6 different categories (Table 3). The most frequently mentioned reason for missing an event was multitasking in the moment of eating (eg, “I forgot to record it because I was at a party.”), which covered 37 of the 69 reasons (54%) provided. In addition, when a participant only gave a single reason for not reporting an eating event, the most common reason was multitasking (5/12 single reasons, 42%). As Table 3 shows, the types of multitasking can be further divided into occasions when participants were unaware that they had missed recording a meal (eg, because they were too busy: 33/69, 48%), and occasions when participants had deliberately decided against recording because of situational barriers (eg, time pressure: 4/69, 6%).

XSL•FO
RenderX

Table 3. Type and number of reported reasons for perceived missing events for studies 1, 2, and 3.

Characteristics	All reasons		Single reasons[a]		Examples
Type of reason	Total, n (%)	Main[b]/snacks, n	Total, n (%)	Main[b]/snacks, n	
Study 1 (n=35)[c]					
Multitasking	37 (53.6)	10/27	5 (41.7)	2/3	N/A[d]
Not being aware	33 (47.8)	10/23	4 (33.3)	2/2	"I forgot to record because I was at a party."
Deliberately deciding against	4 (5.8)	0/4	1 (8.3)	0/1	"It was too awkward to record the second helping."
Device-related obstacles	13 (18.8)	7/6	2 (16.7)	0/2	N/A
Device malfunction	3 (4.3)	2/1	0	0	"The phone ran out of battery."
No device	10 (14.5)	5/5	2 (16.7)	0/2	"I left my smartphone at home."
Situational barriers	10 (14.5)	2/8	2 (16.7)	0/2	N/A
Practical reasons	9 (13.0)	1/8	2 (16.7)	0/2	"I was at work."
Social reasons	1 (1.4)	1/0	0	0	"I was eating with other people."
Not further specified	9 (13.0)	6/3	3 (25.0)	3/0	"I forgot."
Total	69 (100)	25/44	12 (100)	5/7	N/A
Study 2 (n=28)[e]					
Multitasking	15 (34.1)	11/4	8 (53.3)	8/0	N/A
Not being aware	11 (25.0)	10/1	7 (46.7)	7/0	"I was deep in a conversation and did not think about recording."
Deliberately deciding against	4 (9.1)	1/3	1 (6.7)	1/0	"It was too awkward to record the chips."
Device-related obstacles	14 (31.8)	11/3	3 (20.0)	2/1	N/A
Device malfunction	1 (2.3)	1/0	0	0	N/A
No device	13 (29.5)	10/3	3 (20.0)	2/1	"I knew that I would go partying after eating at a friend's house, and I didn't want to take the smartphone to the club."
Situational barriers	11 (25.0)	5/6	3 (20.0)	1/2	N/A
Practical reasons	6 (13.6)	3/3	2 (13.3)	1/1	"I am not allowed to use my smartphone at work."
Social reasons	5 (11.4)	2/3	1 (6.7)	0/1	"I felt awkward when recording my snack in front of other students."
Not further specified	4 (9.1)	2/2	1 (6.7)	1/0	"I forgot to take a picture."
Total	44 (100)	29/15	15 (100)	12/3	N/A
Study 3 (n=99)[f]					
Multitasking	70 (19.4)	49/21	5 (25.0)	4/1	N/A
Not being aware	39 (10.8)	25/14	3 (15.0)	3/0	"I only remembered after my lunch box was empty."
Deliberately deciding against	31 (8.6)	24/7	2 (10.0)	1/1	"I did not have the time."
Device-related obstacles	98 (27.1)	68/30	11 (55.0)	10/1	N/A
Device malfunction	22 (6.1)	17/5	2 (10.0)	2/0	"The app did not work properly."
No device	76 (21.1)	51/25	9 (45.0)	8/1	"I left my smartphone at home."
Situational barriers	48 (13.3)	20/28	0	0	N/A
Practical reasons	35 (9.7)	12/23	0	0	"I could not use my phone during the lecture."

XSL•FO
RenderX

Characteristics	All reasons		Single reasons[a]		Examples
Type of reason	Total, n (%)	Main[b]/snacks, n	Total, n (%)	Main[b]/snacks, n	
Social reasons	13 (3.6)	8/5	0	0	"I was on a date."
Not further specified	145 (40.2)	86/59	4 (20.0)	3/1	"I forgot to take a picture."
Total	361 (100)	223/138	20 (100)	17/3	N/A

[a]Only 1 reason was provided by the participant.

[b]Main meals (breakfast, lunch, and dinner).

[c]In study 1, 35 of 38 participants reported a missing event.

[d]N/A: not applicable.

[e]In study 2, 28 out of 35 participants reported a missing event.

[f]In study 3, 99 out of 110 participants reported a missing event.

Device-related obstacles were the second most frequent reason (13/69 reasons, 19%). These obstacles can be further divided into those related to device malfunctions such as the app not working or the battery was low (3/69 reasons, 4%), and occasions when the participants did not have the device with them (eg, it was not at hand; 10/69 reasons, 15%). Finally, situational barriers that prevented the participants using the mobile device either because it was not sufficiently feasible or admissible accounted for 10 of the 69 reasons (15%). These situational barriers could be further divided into practical reasons such as "going somewhere," "driving," or "not having a hand free," with 9 of the 69 reasons (13%), whereas social reasons such as feeling intimidated by taking a picture were only noted once (1/69, 2%).

Study 2 (8 Days of Mobile Food Recording): The Effect of Reminders and Addenda

Mobile apps currently vary as to whether and how often they prompt the participants to record their eating events, which raises the question of whether mobile apps with reminders yield higher eating event logging rates than those without prompts. Similarly, the possibility of adding meals after the actual eating events has occurred (*addenda*) might also have a positive impact on the rate of missing events. Therefore, study 2 extended study 1 by implementing reminders and the possibility to record missing eating events after they had occurred but still during the food assessment period (*addenda*).

A total of 35 participants (31/35 female, 89%; 32/35 students, 91%) took part in the study, with a mean age of 25.5 years (SD 5.69, range 19-41) and (objectively measured) BMI of 22.5 kg/m^2 (SD 5.5, range 15-43). Of the 35 participants, 20 (57%) were omnivores, 5 (14%) vegetarians, 4 (11%) vegans, and 6 (17%) adhered to other dietary styles. As compensation, the participants received either course credits (3.0 h; n=5) or €25 (US $30; n=30).

None of the participants dropped out of the study, which again indicates excellent retention. A total of 1093 eating events were logged by the participants over the 8-day time period. Of these, 159 entries (14.55%) were canceled before completion, resulting in a total of 934 complete eating event records, of which 888 were recorded during the eating event and 46 (4.9%) were recorded belatedly (*addenda*).

Logged Versus Social Normative Expected Number of Meals

As Table 1 shows, on average, the participants logged 26.69 (SD 7.46) eating events across the 8-day study period. Overall, at the group level, the number of logged eating events was within the social normative expected range of 24 to 32 eating events but significantly below the 32 meals threshold, $t_{34}=-4.21$, $P<.001$, $d=0.71$. As in study 1, a substantial variability in the number of reported eating events occurred at the person level, ranging from 11 to 42 entries per participant. The majority (22/35, 63%) logged 24 or more eating events across the 8 days, and the ICC of $\rho=0.30$ indicates that 70% of the observed variance in the number of logged eating events was due to differences within, rather than between, participants.

The observed number of logged breakfasts, lunches, dinners, and snacks ranged between 6.03 (SD 4.81, lunch) and 6.89 (SD 2.37, breakfast) and were significantly lower than the social normative expected 8 daily meals, respectively. The number of logged events did not vary depending on the meal type, as indicated by a repeated measures ANOVA with the factor *meal type* (breakfast, lunch, dinner, snack), $F_{1.78,60.59}=0.69$, $P=.49$, partial $\eta^2=.02$ (Greenhouse-Geisser corrected, $\varepsilon=.59$). The total number of logged meals and the number of logged meals per meal type did not differ between those participants who had access to the addendum feature (n=18) and those who did not (n=17), $ts_{33}\leq|0.85|$, $Ps\geq.40$, $ds\leq0.29$ (Multimedia Appendix 1). Examining intraindividual differences in logging rates for the different meal types yielded small ICC coefficients ($\rho\leq0.35$), suggesting that the nesting of logged events within individuals was not substantial. Hence, overall, interindividual differences in logging rates were small.

Logging Trajectories of Eating Events Over Time

Figure 1 depicts the logging trajectories for the 8-day study period. Again, logging trajectories over time were analyzed using multilevel modeling (Table 2 and Figure 1).

Overall, there was a small but statistically significant negative trend over time (b=−.12, $t_{236.57}=-3.86$, $P<.001$, pseudo-$R^2=0.06$), indicating that fewer eating events were reported as the study progressed. The random intercept model was preferred ($\chi^2_2=1.1$, $P=.57$), which indicates that the trend was generalizable across the participants.

XSL•FO
RenderX

Although there were no significant changes in logging frequencies over time for the 4 main meal types (Table 2), a small and significant negative trend emerged for snacks (b=−.07, $t_{236.87}$=−3.23, *P*=.001, pseudo-R^2=0.04), indicating that the number of logged snacks decreased significantly across the 8 days. The random intercept model was again preferred (χ^2_2=0.0, *P*>.99), indicating that the trends were comparable between the participants.

A visual inspection of Figure 1 suggests that the number of logged snacks dropped between days 1 and 2. Examining daily logging frequencies for snacks (Multimedia Appendix 1) showed that there was also a shift toward reporting fewer *multiple* snacks between day 1 and day 2. Specifically, the number of participants who recorded 2 to 3 snacks decreased from 16 to 8, whereas those who reported 1 snack increased from 9 to 13. Similar but less pronounced changes occurred between days 2 and 3, and between days 3 and 4.

Number of Reasons for Perceived Missing Events

In the interviews that followed the EMA assessment period, 28 of the 35 participants (80%) reported that they had missed recording at least one eating event (range 1-9 events, median 2.00, mean 2.12, SD 1.88; 5 participants could not state the frequency of missing events). Furthermore, 24 of these 28 participants (86%; 3 did not specify the frequency) stated that they had missed at least one main meal, and 11 of the 28 participants (39%) that they had missed at least one snack. Perceived missing events were significantly higher for main meals (median 1.00, mean 1.48, SD 1.16, range 1-5) than for snacks (median 0.00, mean 0.74, SD 1.39, range 1-6, t_{22}=2.10, *P*=.047, *d*=0.58).

In total, 44 different reasons were provided by the 28 participants who reported a missing event, of which 15 participants specified 1 reason, 10 gave 2 reasons, and 3 participants provided 2 reasons (mean 1.57, SD 0.69).

As Table 3 shows, multitasking in the moment of eating was again the most frequently mentioned reason for missing an event (15/44 reasons, 34%), whereas 11 of 44 (25%) listed not being aware of and 4 of 44 (9%) deliberately deciding against recording an event. The second most common reason was device-related obstacles (14/44, 32%), of which device malfunctions (1/44, 2%) was less frequent than having no device (13/44, 30%). Situational barriers were mentioned 11 of 44 times (25%), of which practical reasons (6/44, 15%) were slightly more frequent than social reasons (5/44, 11%).

Study 3 (8 Days of Mobile Food Recording): Reporting Missing and Skipped Events During the Assessment Period

The *enriched* MIDA in study 2 yielded a similar overall pattern of results to study 1, demonstrating that including a fixed reminder in the morning did not increase the number of logged events. However, one might argue that a fixed daily reminder set at the beginning of the study might not be effective because individual time schedules might vary between and within participants. Thus, in study 3, we tested whether a fixed reminder in the evening, in addition to a daily reminder in the morning, would decrease the rate of missing events. Moreover, to distinguish missing events from skipped eating events, we also included the possibility for participants to indicate the omission or lack of consumption of a meal or snack. Further extending studies 1 and 2, the participants could indicate the occurrence of and reasons for missing events on a daily basis during the MIDA period to reduce potential memory effects.

A total of 113 participants were recruited. Among them, 2 dropped out during the 8-day recording and 1 recorded only 1 eating event, resulting in a final sample of 110 participants (91/110 female, 82.7%; 106/110 students, 96.4%) with a mean age of 22.02 years (SD 5.30, range 18-51) and (self-reported) BMI of 21.9 kg/m^2 (SD 3.44, range 17-44). Of the 110 participants, 66 (60.0%) were omnivores, 18 (16.4%) vegetarians, 8 (7.3%) vegans, and 18 (16.4%) adhered to other dietary styles. The participants received course credits (1.5 h; n=70) or €15 (US $18; n=40) as compensation for their participation.

As of the 113 participants, only 2 dropped out of the study and 1 recorded only 1 meal, a very good retention rate (97.4%) is indicated. In total, the participants logged 3365 eating events over the 8-day study period. Of these, 133 entries (3.95%) were canceled by the participants or were incomplete (eg, picture missing), resulting in a total of 3232 complete eating event records (Multimedia Appendix 1), of which 2871 were recorded at the eating event and 361 (11.17%) were belated recordings (*addenda*). In addition, 86 of 110 participants (78.2%) reported having skipped 245 eating events. No meal type was specified for 8 skipped eating events, resulting in a total of 237 skipped eating events that were added to the eating event records. Hence, the total final sample included 3469 eating events (Table 1). The control analysis showed no significant difference between participants who used their own vs a loaned smartphone with respect to the number of reported meals (ts_{108}≤|1.81|, *P*s≥.07, *d*s≤0.35) or the number of reported missing snacks (t_{97}=1.37, *P*=.17, *d*=0.28). However, they differed in terms of the number of reported missing main meals (t_{97}=2.13, *P*=.04, *d*=0.43; $mean_{own}$ 2.62, SD_{own} 2.07, $mean_{loaned}$ 1.86, SD_{loaned} 1.54).

Logged Versus Social Normative Expected Number of Meals

On average, the participants recorded 31.54 (SD 8.73) meals and snacks during the 8 study days (Table 1). In contrast to studies 1 and 2, the number of recorded eating events concurs with the social normative threshold value of 32 occasions (4 daily meals and an additional snack, t_{109}=−0.56, *P*=.58, *d*=0.05). The number of entries ranged from 10 to 61 per participant, with most (91/110, 82.7%) logging 24 or more eating events during the study period. Majority (68%) of the overall variation in logged occasions was because of variations within participants rather than variations of logging rates between individuals (ICC of ρ=0.32).

Table 1 shows that the observed number of logged lunches (mean 6.76, SD 1.78) and dinners (mean 7.56, SD 1.96) were significantly lower than the social normative expected total of 8 for the study period, whereas the number of logged breakfasts (mean 7.53, SD 2.58) and snacks (mean 8.66, SD 6.23) reached

XSL•FO
RenderX

the social normative expected threshold value of 8 occasions. Accordingly, the number of reports differed between meal types, as indicated by a repeated-measures ANOVA, $F_{1.53,167.08}$=5.84, P=.007, partial η^2=.05 (Greenhouse-Geisser corrected, ε=.51). Significantly fewer lunches were recorded than dinners (t_{109}=−4.09, P<.001, d=0.43) and snacks (t_{109}=−3.16, P=.002, d=0.41). All other comparisons were not statistically significant, with ts_{109}≤−2.66, Ps≥.009, ds≤0.35 (α adjusted to .008 to account for multiple comparisons). ICC coefficients were all ρ≤0.35, suggesting that, overall, interindividual differences in logging rates were small.

The number of reported meals without including skipped meals was also analyzed to examine the effect of including the recording of skipped meals (Multimedia Appendix 1). A total of 3232 eating events were recorded, with an average of 29.38 (SD 8.54) events logged across 8 days, which is significantly lower than the normative threshold of 32 meals (t_{109}=−3.22, P=.002, d=0.31, ICC of ρ=0.34). However, the majority still logged more than 24 meals throughout the study period (82/110, 74.5%). On average, the participants missed 2.62 eating events across the 8 days. The number of missing events also deviated significantly from the normative expected 8 eating events for the different meal types (breakfast mean 6.97, SD 2.64; lunch mean 6.09, SD 1.91; and dinner mean 7.51, SD 2.09; ts_{109}≥−4.08, Ps<.001, ds≥0.39). However, the observed and normative expected number of logged events concurred for snacks (mean 8.23, SD 5.93; t_{109}=.23, P=.69, d=0.04).

Logging the Trajectories of Meals Over Time

Logging trajectories for the 8-day study period were again examined using multilevel modeling (Table 2 and Figure 1). Overall, there was a small but statistically significant negative trend over time (b=−.11, $t_{753.99}$=−5.52, P<.001, pseudo-R^2=0.04), indicating that fewer meals were reported over time. The random intercept model was preferred (χ^2_2=1.3, P=.53), which indicates that participants did not vary with respect to time trends.

For individual meal types, small and significant negative trends also emerged for snacks (b=−.08, $t_{753.06}$=−5.41, P<.001, pseudo-R^2=0.04), indicating that the number of reported snacks decreased across the 8 consecutive days. Again, the random intercept model was preferred (χ^2_2=3.1, P=.22), demonstrating that the trends were comparable between participants. No significant change in logging frequency over time was observed for the other food types (breakfast, lunch, dinner, and afternoon tea; Table 2).

Number of Reasons for Perceived Missing Events

Of the 110 participants, 99 (90.0%) reported that they had missed recording at least one main meal or snack (median 3.00, mean 3.65, SD 2.93, range 1-20). For main meals, 90 of 99 participants (82%) reported at least one missing event, whereas for snacks, missing events were reported by 57 of 99 (52%) participants. On average, the number of missing events for snacks (median 1.00, mean 1.39, SD 1.96) was significantly lower than the number of missing events for main meals (median 2.00, mean 2.25, SD 1.86; t_{98}=3.48, P=.001, d=0.45).

In total, 361 different reasons were provided by the 99 participants who reported having missed logging at least one eating event, of which 20 participants (20%) specified one reason for missing reporting an eating event, 22 (22%) gave 2, and 57 (58%) between 3 and 20 reasons (median 3.00, mean 3.65, SD 2.93).

Of the 361 reasons provided, 216 (59.8%) included information beyond merely stating that an eating event was missed. As Table 3 shows, device-related obstacles (specifically having no device at hand) were the most frequently specified reason for a missing event (98/361, 27.1%). This was also the case when only a single reason was provided (11/20 provided single reasons, 55%). The next most frequent reasons were those related to multitasking (70/361, 19.4%), followed by situational barriers (48/361, 13.3%).

Discussion

The goal of the 3 studies was to assess the number of and reasons for *missing events* in MIDA in event-based EMA studies, using apps that varied in their technical features. The 3 studies assessed how often eating events (meals and snacks) were missed over a period of 8 days in a naturalistic setting by comparing the number of recorded events (1) with the number of expected events based on social norms and observational data, (2) over time, and (3) with recollections of missing events. We used different apps ranging from a *basic* app (study 1) to a more *enriched* MIDA app, which included individually set reminders to record food intake, and the possibility of recording *addenda* and skipped meals (study 3). To gain a greater insight into the occurrence of missing events, we also assessed the users' reasons for missing events. To our knowledge, these are the first studies that have assessed individual's performance by recording dietary food intake across an extended duration in their naturalistic settings.

Study Attrition

Attrition rates, as an indicator of adherence, are discussed in the specific context of EMA studies because its comparably intensive data collection is presumed to lead to a high perceived burden on the participants and thus to discontinued usage [49]. Moreover, study attrition is essential for estimating momentary adherence or compliance rates [49,50]. All 3 of the studies showed excellent overall study retention, with only 3 of the 189 participants who enrolled dropping out during the 8-day study period (1.6%). Liao et al [23] summarized compliance-related results from 13 EMA studies of diet and physical activity in youth and, although most previous EMA studies reported initial participant enrollment, only 2 studies formally reported attrition rates (ie, the number of participants who dropped out of the study for any reason). Similarly, Schembre et al [24] reviewed the methods used in mobile diet EMA studies without including information on study attrition (see also study by Wen et al [50]). Villinger, Wahl et al [20] recently analyzed 41 app-based intervention studies on nutrition behaviors and nutrition-related health outcomes in adolescents and adults with a total of 6348 participants and found an average attrition rate of 18.7% (SD 16.27, range 0%-72%). This is highly comparable to the average attrition rate of 18% found by Crutzen et al [51] for health

XSL•FO
RenderX

behavior change trials. Thus, attrition rates are comparably low even in longitudinal intervention studies, which place a higher burden on the participants than behavior assessment studies. Taken together, these and the findings from this study provide some confidence that using intensive data collection methods including mobile dietary EMA is feasible in terms of participant retention.

Momentary Adherence and Missing Events Across Studies and Over Time

Overall, 183 participants reported 5473 eating events, including 3803 main meals, 1505 snacks, and 165 afternoon teas. The momentary adherence rate, as indexed by a comparison with normative expected events, was generally high, but a differential pattern of results emerged across the 3 studies.

Although the average number of events observed in study 3 corresponded to the social normative expected number of 32 events (d=0.05), the overall number of eating events logged in studies 1 and 2 were significantly below the social normative threshold. In study 1, the participants missed an average of 3.8 eating events (d=0.40) across the 8-day assessment period; in study 2, the rate was even higher at 5.3 missing events (d=0.71).

Converting these results into practical significance is not unequivocally possible as *ground truth* (eg, [52-54]) is virtually unavailable (only 2 studies used doubly labeled water as an intake criterion, [6,55]). Multiple criteria have been suggested for evaluating the practical significance of observed effects [56] (see also [20]). Regarding benchmark values, Cohen effect sizes for the overall missing event rates are in the small-to-medium range [57]. Furthermore, the comparison of the effects size values of this study with previous research is limited because of a lack of research. Specifically, the vast majority of EMA studies used signal-contingent data collection protocols in which the participants are prompted (often many times per day) to provide information, and adherence can be defined as the percentage of prompts (eg, signals and reminders) to which the participants responded [23,24,50]. Silvia et al [58] estimated that signal-contingent EMA studies commonly have 15% to 35% prompt-wise missing data rates [59]. Schembre et al [24] reported a mean nonresponse rate of 21% for 10 signal-contingent diet EMA studies, which is similar to the review results from Heron et al [22], who found an overall nonresponse rate of 24%. For the studies in this paper, a comparison of the average observed logging rate over the 8 days to a social normative threshold value of 32 yielded overall missing event rates of 12% (study 1), 17% (study 2), and 1.4% (study 3). However, the social normative threshold of 32 eating events across an 8-day study period and the comparison with signal-contingent EMA studies is clearly debatable and should therefore be interpreted with great caution.

Importantly, the overall higher number of reported eating events in study 3 compared with studies 1 and 2 was also reflected at the level of specific meals. For breakfasts, lunches, and dinners, the participants in study 3 reported an average of almost 1 event per day and between 0.64 and 1.16 more breakfasts, lunches, and dinners than participants in studies 1 and 2. The reasons for the higher momentary adherence rate in study 3 could include individual differences, the prompting strategy used, and technological features.

The Impact of Technical Features on Missing Events: Addenda, Prompts, and Skipped Meals

One technological feature that might have increased the number of logged eating events in study 3 is the addendum feature. Eldridge et al [17] reviewed 43 technology-based dietary assessment tools and found that more than 40% included an option to add missing foods, making it the most common customized feature. In the context of paper-and-pencil methods, backfilling is seen as an obstacle to valid assessments. However, as mobile device data are time-stamped, addendum features offer the advantage of providing more insight into adherence, and their added flexibility may also contribute to the participants' motivation [60,61]. Supporting the notion that addenda substantially contribute to a lower number of missing events, 11% of the complete eating events in study 3 were actually belated.

Furthermore, study 3 included 2 fixed daily reminders, which might have also contributed to a higher logging rate. Reminders and prompts are key features of most EMA and mobile assessment studies for enhancing engagement and adherence. Prompts such as push notifications can now utilize individuals' contexts to determine the most opportune times to send prompts [62]. *Interruptibility* research has emerged within the field of human-computer interactions, along with text-messaging interventions in psychology and public health [63]. However, current findings from experimental studies indicate that although prompts may encourage greater exposure to message or intervention contents without deterring engagement, they do not always enhance their use [63,64]. Morrison et al compared intelligent notifications, daily notifications within predefined time frames, and occasional notifications within predefined time frames in a stress management intervention and found generally low response rates but a small-to-medium effect on viewed and actioned notifications for the first two compared with occasional notifications. Comparisons between study 1, which did not include any reminder or prompts, and study 2, which included a fixed daily reminder to log one's meals, supports the notion that predefined reminders might not per se increase momentary adherence. However, in study 3, 2 fixed daily reminders were added, which might have contributed to the higher momentary adherence rate. In particular, the (self-selected) fixed reminder in the evening might have positively impacted momentary adherence because it additionally reminded participants to log eating events they missed logging or skipped during the day. Further supporting this notion, fewer participants in study 3 reported that they were not aware that they had missed logging an eating event compared with studies 1 and 2.

Study 3 also included a technical feature, which allowed the participants to log a skipped meal, and this was used at least once by 78% of the 110 participants. Including the number of skipped meals increased the number of logged eating events by an average of 2 logs across the 8-day study period, which represents 7% of the total entries. Interestingly, skipped meals were reported with almost equal frequency for all three main meals. Pendergast et al [65] assessed skipped meals in young

XSL•FO
RenderX

adults in a 4-day EMA study on the following day and found that almost 50% of the participants were regular meal skippers (skipping ≥25% of main meals), with 15% of the sample defined as breakfast skippers, 12% as lunch skippers, and 10% as dinner skippers. Therefore, meal skipping occurs across a broad range of people and affects all main meal types. Assessing skipped meals offers the possibility of increasing measurement precision, as it allows differentiation between skipped meals and missing events. Thus, adding a skipped meal recording feature is likely to have a substantial and consistent effect on measurement quality.

Momentary Adherence and Missing Events Over Time: Logging Trajectories

Extending previous EMA diet studies, we analyzed the logging trajectories over time to examine whether momentary adherence declines over time. Across all 3 studies, the logging trajectories of reported meals over time showed a significant, albeit small, decline over time. The trend was slightly more pronounced in study 2 (pseudo-R^2=0.06) and less pronounced in study 3 (pseudo-R^2=0.04), mirroring the previously discussed results for time-aggregated data. The random intercept model was preferred in all 3 studies, indicating that the participants did not differ significantly in their logging rates over time. Although the results consistently indicate that interindividual differences did not systematically affect logging trajectories over time, we found a differential pattern in dependence of the meal type within each study.

The logging trajectories for the main meals (breakfast, lunch, and dinner) showed no decrease in reporting, except for a small-time effect in study 1 for lunch and dinner. This was surprising, given that the intensive nature of diet EMA protocols and the burden on participants has often been viewed as a major contributing factor in both response fatigue [66] and decreases in momentary adherence. Converging with the present results as well as a recent meta-analysis of mobile EMA studies, no significant study duration effects were found on the rate of response to prompts [50,58,67], which emphasizes the need for a more nuanced understanding of the factors that affect momentary adherence.

The significant negative time effect we found for snacks in studies 2 and 3, which was not found for main meals, might indicate underreporting or measurement reactivity [68-70]. Although underreporting means that snacks were consumed but not logged, which constitutes missing events, measurement reactivity describes the effect that repeated assessments draw attention to the monitored behavior, which can identify problematic behavior and induce behavioral changes. Interestingly, the number of reported snacks dropped noticeably between days 1 and 2 (see Figure 1), which is when the attention effect and identification of problematic behavior is theoretically most likely. However, considering the present data, it is not possible to disentangle underreporting (ie, consumed snacks not being reported) and measurement reactivity effects (ie, an actual decrease in snacks consumed).

Perceived Missing Events and Reasons Across Studies

Almost all participants acknowledged that they had missed logging at least one eating event during the study period. Interestingly, the *enrichments* in study 3, namely the 2 fixed daily reminders and an addendum feature, did not result in a lower number of participants recalling a missing event compared with the more *basic* study 1 (92% vs 90%), which suggests that people are aware or assume that their reports are incomplete.

However, examining the number of missing events reported shows a lower median number of perceived missing events in study 1 than in study 3, but compared with the social normative expected number of events, perceived events were lower in study 1 and higher in study 3. Therefore, one might speculate that including technical features such as reminders might not only increase the actual logging of eating events but could also increase the awareness of missing events. Perceived missing events in studies 2 and 3 were more likely to be main meals than snacks, which contrasts with the assumption that irregular eating events might be rather more likely to be missed because of attention and memory effects.

Overall, the reported reasons for missing events showed a similar profile across all 3 studies. The most common reasons were engagement in competing activities and technical issues, whereas situational reasons were less important. In the context of multitasking, further examination shows that although completion rates are affected by lapses in attention, resource, and time scarcity also lead to deliberately deciding against recording an eating event. Although customized prompts might be useful for the former, the latter represents a clear limitation for active real-time assessments. Passive assessments of eating, such as automatic sensing through wearables [52-54] circumvent this limitation. As expected, device-related obstacles (specifically not having the device handy for recording) were an important reason for missing events. Situational barriers, specifically practical reasons that rendered recording as neither feasible nor admissible, were also common reasons for missing events. In contrast, social reasons, such as feeling intimidated by taking a picture, were only rarely noted.

Study Limitations

Although attrition rates were very low in these studies, biases such as recruitment or volunteer bias need to be considered [71]. Specifically, individuals who are willing to participate in a research study are self-selected and presumably highly motivated and health conscious. Research on participation biases in mailed surveys showed that people who are female, older, or with higher education levels are more likely to return postal questionnaires. Furthermore, people with a poorer health status tend to avoid participating in health surveys [72]. Thus, the predominance of females and students among the participants suggests the presence of a *volunteer bias* in these study series, which limits the generalization of the pattern of results to other population groups. Furthermore, our sample sizes were relatively small in comparison with large-scale epidemiological studies. Although research using mobile EMA is demanding, it would be informative to address the issue of missing event rates with larger samples.

XSL•FO
RenderX

Relatedly, individual differences can also systematically impact momentary adherence. Specifically, previous research has shown that women and people with a college degree are more likely to respond to prompts within signal-contingent EMA studies [58,73]. However, the ICCs across the 3 studies indicated that 38% or less of the observed variance was because of individual differences. Thus, differences in logging rates are more pronounced within individuals than between different people. Overall, the present findings suggest that the actual state of the individual, situational context, and technical features of the mobile app seem to have a greater influence on the adherence rates than stable individual differences (see the study by Sun et al [74] for similar conclusions).

An important factor in mobile app research is the duration of the assessment period, which may impact momentary adherence over time. Although assessment duration has received increasing attention in the study of dietary intake and has been addressed in existing assessment guidelines [41], the different fields and studies use different time periods. Similar to these studies, mobile dietary assessment studies commonly use periods ranging from 1 to 7 days [16]. However, longer assessment periods might be necessary for specific research questions, such as the assessment of micronutrients [41] and the evaluation of intervention studies using dietary self-monitoring. Overall, the study findings need to be interpreted within the context of the 8-day assessment period and the generalizability of the findings, and future studies are needed to determine the maximum number of feasible data collection days.

A further limitation regarding the analysis of the numbers of and reasons for perceived missing events assessed by semistructured interviews in studies 1 and 2. It seems reasonable that memory biases, that is, primacy and recency effects, may have impacted recollection. Consistent with this notion, study 3, which provided the opportunity to report missing events on a daily basis, indicated an average of 2 more logs across the 8-day study period.

The strength of this study series regarding the analysis of logging data across time, that is, showing only a small decline in event logging, and the analysis of reasons for missing events, that is competing activities and technical issues. However, this study series is limited with regard to the precise estimation of missing events. Specifically, based on epidemiological data and social eating norms, the number of missed events was estimated in reference to external criteria rather than the objective assessment of the number of event episodes. This limitation has been widely acknowledged and is accentuated in research over longer assessment periods ranging from weeks to months [52-54,75]. One possibility to acquire objective data is to use video and audio recordings throughout the day to objectively identify the number of eating episodes. However, such an approach would raise ethical issues regarding data privacy, involve time, and require resource-intensive annotation efforts from external observers or the participants themselves. Acknowledging these limitations [41], future studies are needed to estimate the frequency of missing events in reference to objective data.

Conclusions

Using 3 different indicators, missing events were assessed in 3 mobile image-based dietary EMA studies in more than 180 adults, who reported 5473 eating events, including 3803 main meals, 1505 snacks, and 165 afternoon teas. Given the intensive nature of diet EMA protocols, logging trajectories over time were remarkably stable for main meals. The small significant negative time effect for snacks might indicate underreporting or measurement reactivity. Differences in logging rates were more pronounced within individuals than between different persons. Hence, the actual state of the individual and context seem to have a greater influence on adherence rates than stable individual differences. Supporting this notion, study 3, which included an *enriched* app with reminders, addendum option, and the possibility of recording skipped meals, yielded the highest number of recorded meals. Thus, including such customized features can substantially increase the measurement quality. Engagement in competing activities and technical issues were the most frequently named reasons for perceived missing events, whereas situational reasons were less important. The results emphasize the need for a more nuanced understanding of the factors that affect momentary adherence.

Acknowledgments

This research was supported by the Federal Ministry of Education and Research and the German Research Foundation within the research projects SMARTACT (BMBF Grant 01EL1820A) and RISKDYNAMICS (DFG Grant FOR 2374) granted to BR and HS. The funding sources had no involvement in the study's design; collection, analysis, and interpretation of data; writing of the report; or the decision to submit this article for publication. The authors thank Heike Koch, Kristian Sabados, Tobias Volk, and Tony Arthur for their valuable support.

Authors' Contributions

BR and HS developed the study concept. All authors contributed to the study design. KZ, LK, DW, and KV conducted the study, including participant recruitment and data collection, under the supervision of BR and HS. LK, BR, and KZ conducted data analyses, and SB, JM, and HR created and implemented the mobile app SMARTFOOD (study 2) designed by BR, HS, KV, DW, LK, and KZ. The manuscript draft was prepared by BR, KZ, LK, and CB, with critical input by HS, KV, and DW. All authors approved the final version of the manuscript for submission.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional results for studies 2 and 3.
[DOCX File , 16 KB - mhealth_v8i10e15430_app1.docx]

References

1. Fischler C. Food habits, social change and the nature/culture dilemma. Soc Sci Inf 2016 Sep 3;19(6):937-953. [doi: 10.1177/053901848001900603]
2. Wahl DR, Villinger K, Blumenschein M, König LM, Ziesemer K, Sproesser G, et al. Why we eat what we eat: assessing dispositional and in-the-moment eating motives by using ecological momentary assessment. JMIR Mhealth Uhealth 2020 Jan 7;8(1):e13191 [FREE Full text] [doi: 10.2196/13191] [Medline: 31909719]
3. Rozin P. Food and eating. In: Kitayama S, Cohen D, editors. Handbook of Cultural Psychology. New York, USA: Guilford Press; 2007:391-416.
4. Wansink B, Sobal J. Mindless eating. Environ Behav 2016 Jul 26;39(1):106-123. [doi: 10.1177/0013916506295573]
5. Boushey CJ, Spoden M, Zhu FM, Delp EJ, Kerr DA. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. Proc Nutr Soc 2017 Aug;76(3):283-294. [doi: 10.1017/S0029665116002913] [Medline: 27938425]
6. Martin CK, Correa JB, Han HM, Allen HR, Rood JC, Champagne CM, et al. Validity of the remote food photography method (RFPM) for estimating energy and nutrient intake in near real-time. Obesity (Silver Spring) 2012 Apr;20(4):891-899 [FREE Full text] [doi: 10.1038/oby.2011.344] [Medline: 22134199]
7. Rollo ME, Williams RL, Burrows T, Kirkpatrick SI, Bucher T, Collins CE. What are they really eating? A review on new approaches to dietary intake assessment and validation. Curr Nutr Rep 2016 Oct 15;5(4):307-314. [doi: 10.1007/s13668-016-0182-6]
8. Kirkpatrick SI, Reedy J, Krebs-Smith SM, Pannucci TE, Subar AF, Wilson MM, et al. Applications of the healthy eating index for surveillance, epidemiology, and intervention research: considerations and caveats. J Acad Nutr Diet 2018 Sep;118(9):1603-1621 [FREE Full text] [doi: 10.1016/j.jand.2018.05.020] [Medline: 30146072]
9. Kirkpatrick SI, Collins CE. Assessment of nutrient intakes: introduction to the special issue. Nutrients 2016 Mar 25;8(4):184 [FREE Full text] [doi: 10.3390/nu8040184] [Medline: 27023598]
10. Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, et al. Comparison of dietary assessment methods in nutritional epidemiology: weighed records v 24 h recalls, food-frequency questionnaires and estimated-diet records. Br J Nutr 1994 Oct;72(4):619-643. [doi: 10.1079/bjn19940064] [Medline: 7986792]
11. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TI, Energy Balance Measurement Working Group. Energy balance measurement: when something is not better than nothing. Int J Obes (Lond) 2015 Jul;39(7):1109-1113 [FREE Full text] [doi: 10.1038/ijo.2014.199] [Medline: 25394308]
12. Freedman LS, Midthune D, Arab L, Prentice RL, Subar AF, Willett W, et al. Combining a food frequency questionnaire with 24-hour recalls to increase the precision of estimation of usual dietary intakes-evidence from the validation studies pooling project. Am J Epidemiol 2018 Oct 1;187(10):2227-2232 [FREE Full text] [doi: 10.1093/aje/kwy126] [Medline: 29917051]
13. Illner A, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012 Aug;41(4):1187-1203. [doi: 10.1093/ije/dys105] [Medline: 22933652]
14. Renner B, Klusmann-Weißkopf V, Sproesser G. Assessment of health behaviors. In: Wright JD, editor. The International Encyclopedia of the Social and Behavioral Sciences. Oxford, England: Elsevier; 2015:588-593.
15. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr 2015 Dec;145(12):2639-2645 [FREE Full text] [doi: 10.3945/jn.115.219634] [Medline: 26468491]
16. Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. Nutrition 2014;30(11-12):1257-1266. [doi: 10.1016/j.nut.2014.02.020] [Medline: 24976425]
17. Eldridge AL, Piernas C, Illner A, Gibney MJ, Gurinović MA, de Vries JH, et al. Evaluation of new technology-based tools for dietary intake assessment-an ILSI Europe dietary intake and exposure task force evaluation. Nutrients 2018 Dec 28;11(1):- [FREE Full text] [doi: 10.3390/nu11010055] [Medline: 30597864]
18. Villinger K, Wahl DR, König LM, Ziesemer K, Butscher S, Müller J, et al. Do we know what we enjoy? Accuracy of forecasted eating happiness. Front Psychol 2020;11:1187 [FREE Full text] [doi: 10.3389/fpsyg.2020.01187] [Medline: 32625135]

XSL•FO
RenderX

19. Boushey CJ, Harray AJ, Kerr DA, Schap TE, Paterson S, Aflague T, et al. How willing are adolescents to record their dietary intake? The mobile food record. JMIR Mhealth Uhealth 2015 May 29;3(2):e47 [FREE Full text] [doi: 10.2196/mhealth.4087] [Medline: 26024996]
20. Villinger K, Wahl DR, Boeing H, Schupp HT, Renner B. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: a systematic review and meta-analysis. Obes Rev 2019 Oct;20(10):1465-1484 [FREE Full text] [doi: 10.1111/obr.12903] [Medline: 31353783]
21. 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 Apr 15;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
22. Heron KE, Everhart RS, McHale SM, Smyth JM. Using mobile-technology-based ecological momentary assessment (EMA) methods with youth: a systematic review and recommendations. J Pediatr Psychol 2017 Nov 1;42(10):1087-1107. [doi: 10.1093/jpepsy/jsx078] [Medline: 28475765]
23. Liao Y, Skelton K, Dunton G, Bruening M. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted strobe checklist for reporting EMA studies (CREMAS). J Med Internet Res 2016 Jun 21;18(6):e151 [FREE Full text] [doi: 10.2196/jmir.4954] [Medline: 27328833]
24. Schembre SM, Liao Y, O'Connor SG, Hingle MD, Shen S, Hamoy KG, et al. Mobile ecological momentary diet assessment methods for behavioral research: systematic review. JMIR Mhealth Uhealth 2018 Nov 20;6(11):e11170 [FREE Full text] [doi: 10.2196/11170] [Medline: 30459148]
25. Glanz K, Murphy S. Dietary assessment and monitoring in real time. In: Stone AA, Shiffman S, Atienza A, Nebeling L, editors. The Science of Real-time Data Capture: Self-reports in Health Research. Oxford, USA: Oxford University Press; 2007:151-168.
26. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4(1):1-32. [Medline: 18509902]
27. Mehl MR, Conner TS. Handbook of research methods for studying daily life. New York: The Guilford Press; 2012.
28. Conner T, Lehman B. Getting started: Launching a study in daily life. In: Mehl MR, Connor TS, editors. Handbook of research methods for studying daily life. New York: The Guildford Press; 2012:89-107.
29. König LM, Renner B. Colourful = healthy? Exploring meal colour variety and its relation to food consumption. Food Quality and Preference 2018;64:66-71. [doi: 10.1016/j.foodqual.2017.10.011]
30. Bucher DTS, Carrard I, Farina E, Danuser B, Kruseman M. Development and Evaluation of e-CA, an Electronic Mobile-Based Food Record. Nutrients 2017;9(1):76 [FREE Full text] [doi: 10.3390/nu9010076] [Medline: 28106767]
31. Sun M, Burke LE, Mao Z, Chen Y, Chen H, Bai Y, et al. eButton: A Wearable Computer for Health Monitoring and Personal Assistance. Proc Des Autom Conf 2014;2014:1-6 [FREE Full text] [doi: 10.1145/2593069.2596678] [Medline: 25340176]
32. Fontana JM, Farooq M, Sazonov E. Automatic ingestion monitor: a novel wearable device for monitoring of ingestive behavior. IEEE Trans Biomed Eng 2014;61(6):1772-1779 [FREE Full text] [doi: 10.1109/TBME.2014.2306773] [Medline: 24845288]
33. Ashman AM, Collins CE, Brown LJ, Rae KM, Rollo ME. Validation of a Smartphone Image-Based Dietary Assessment Method for Pregnant Women. Nutrients 2017;9(1):73 [FREE Full text] [doi: 10.3390/nu9010073] [Medline: 28106758]
34. Seto E, Hua J, Wu L, Shia V, Eom S, Wang M, et al. Models of individual dietary behavior based on smartphone data: the influence of routine, physical activity, emotion, and food environment. PLoS One 2016;11(4):e0153085 [FREE Full text] [doi: 10.1371/journal.pone.0153085] [Medline: 27049852]
35. Maugeri A, Barchitta M. A systematic review of ecological momentary assessment of diet: implications and perspectives for nutritional epidemiology. Nutrients 2019;11(11):2696 [FREE Full text] [doi: 10.3390/nu11112696] [Medline: 31703374]
36. Grignon C. Sociologie des rythmes alimentaires. Les comportements alimentaires. 2004. URL: https://www6.versailles-grignon.inrae.fr/aliss/layout/set/print/content/download/3346/35629/file/WP04-10.pdf [accessed 2020-09-17]
37. Lhuissier A, Tichit C, Caillavet F, Cardon P, Masullo A, Martin-Fernandez J, et al. Who still eats three meals a day? Findings from a quantitative survey in the Paris area. Appetite 2013 Apr;63:59-69. [doi: 10.1016/j.appet.2012.12.012] [Medline: 23274963]
38. Murakami K, Livingstone MB. Associations between meal and snack frequency and diet quality in us adults: national health and nutrition examination survey 2003-2012. J Acad Nutr Diet 2016 Jul;116(7):1101-1113. [doi: 10.1016/j.jand.2015.12.012] [Medline: 26847912]
39. Schwedhelm C, Knüppel S, Schwingshackl L, Boeing H, Iqbal K. Meal and habitual dietary networks identified through semiparametric gaussian copula graphical models in a german adult population. PLoS One 2018;13(8):e0202936 [FREE Full text] [doi: 10.1371/journal.pone.0202936] [Medline: 30142191]
40. Winkler G, Döring A, Keil U. Meal patterns in middle-aged men in Southern Germany: results from the MONICA Augsburg dietary survey 1984/85. Appetite 1999 Feb;32(1):33-37. [doi: 10.1006/appe.1998.0193] [Medline: 9989911]
41. Food and Agriculture Organization of the United Nations. Dietary Assessment. A Resource Guide to Method Selection and Application in Low Resource Settings. Rome, Italy: Food and Agriculture Organization of the United Nations; 2018.
42. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using. J Stat Soft 2015;67(1):1-48 [FREE Full text] [doi: 10.18637/jss.v067.i01]

XSL•FO
RenderX

43. Kuznetsova A, Brockhoff PB, Christensen RH. Package: tests in linear mixed effects models. J Stat Soft 2017;82(13):1-26 [FREE Full text] [doi: 10.18637/jss.v082.i13]
44. Hox JJ, Moerbeek M, van de Schoot R. Multilevel Analysis: Techniques and Applications. New York, USA: Routledge; 2017.
45. Raudenbush SW, Bryk A. Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: Sage Publications; 2002.
46. Butscher S, Wang Y, Ziesemer K, Villinger K, Wahl DR, König LM. Lightweight Visual Data Analysis on Mobile Devices: Providing Self-Monitoring Feedback. In: 1st International Workshop on Valuable Visualization of Healthcare Information. 2016 Presented at: VVHI'16; June 7-10, 2016; Bari, Italy URL: https://www.researchgate.net/publication/309491379_Lightweight_Visual_Data_Analysis_on_Mobile_Devices_-Providing_Self-Monitoring_Feedback
47. Wahl DR, Villinger K, König LM, Ziesemer K, Schupp HT, Renner B. Healthy food choices are happy food choices: evidence from a real life sample using smartphone based assessments. Sci Rep 2017 Dec 6;7(1):17069 [FREE Full text] [doi: 10.1038/s41598-017-17262-9] [Medline: 29213109]
48. SmartAct. University of Konstanz. 2019. URL: https://www.uni-konstanz.de/smartact/ [accessed 2019-05-21] [WebCite Cache ID 78XNmGzzr]
49. 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]
50. Wen CK, Schneider S, Stone AA, Spruijt-Metz D. Compliance with mobile ecological momentary assessment protocols in children and adolescents: a systematic review and meta-analysis. J Med Internet Res 2017 Apr 26;19(4):e132 [FREE Full text] [doi: 10.2196/jmir.6641] [Medline: 28446418]
51. Crutzen R, Viechtbauer W, Spigt M, Kotz D. Differential attrition in health behaviour change trials: a systematic review and meta-analysis. Psychol Health 2015 Jan;30(1):122-134. [doi: 10.1080/08870446.2014.953526] [Medline: 25109224]
52. Gomes D, Sousa I. Real-time drink trigger detection in free-living conditions using inertial sensors. Sensors (Basel) 2019 May 9;19(9):- [FREE Full text] [doi: 10.3390/s19092145] [Medline: 31075843]
53. Farooq M, Doulah A, Parton J, McCrory MA, Higgins JA, Sazonov E. Validation of sensor-based food intake detection by multicamera video observation in an unconstrained environment. Nutrients 2019 Mar 13;11(3):- [FREE Full text] [doi: 10.3390/nu11030609] [Medline: 30871173]
54. Thomaz E. Automatic Eating Detection in Real-world Settings With Commodity Sensing. Georgia Institute of Technology. 2016. URL: https://www.semanticscholar.org/paper/Automatic-eating-detection-in-real-world-settings-Thomaz/dfc8579030fe38f6816942e49d2ff23cdbb30f72 [accessed 2020-09-15]
55. Boushey CJ, Spoden M, Delp EJ, Zhu F, Bosch M, Ahmad Z, et al. Reported energy intake accuracy compared to doubly labeled water and usability of the mobile food record among community dwelling adults. Nutrients 2017 Mar 22;9(3):- [FREE Full text] [doi: 10.3390/nu9030312] [Medline: 28327502]
56. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol Rev Camb Philos Soc 2007 Nov;82(4):591-605. [doi: 10.1111/j.1469-185X.2007.00027.x] [Medline: 17944619]
57. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
58. Silvia PJ, Kwapil TR, Eddington KM, Brown LH. Missed beeps and missing data. Soc Sci Comput Rev 2013 Mar 13;31(4):471-481. [doi: 10.1177/0894439313479902]
59. Hufford MR, Shields AL, Shiffman S, Paty J, Balabanis M. Reactivity to ecological momentary assessment: an example using undergraduate problem drinkers. Psychol Addict Behav 2002 Sep;16(3):205-211. [Medline: 12236455]
60. Deci EL, Ryan RM. The 'what' and 'why' of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000 Oct;11(4):227-268. [doi: 10.1207/s15327965pli1104_01]
61. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
62. Wang Y, Duan L, Butscher S, Mueller J, Reiterer H. Fingerprints: Detecting Meaningful Moments for Mobile Health Intervention. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct. 2016 Presented at: 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct; September 6-9, 2016; Florence, Italy p. 1085-1088. [doi: 10.1145/2957265.2965006]
63. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. JMIR Mhealth Uhealth 2018 Nov 29;6(11):e10123. [doi: 10.2196/10123] [Medline: 30497999]
64. Morrison LG, Hargood C, Pejovic V, Geraghty AW, Lloyd S, Goodman N, et al. Correction: the effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: an exploratory trial. PLoS One 2018;13(5):e0198008 [FREE Full text] [doi: 10.1371/journal.pone.0198008] [Medline: 29787600]
65. Pendergast FJ, Livingstone KM, Worsley A, McNaughton SA. Examining the correlates of meal skipping in Australian young adults. Nutr J 2019 Apr 3;18(1):24 [FREE Full text] [doi: 10.1186/s12937-019-0451-5] [Medline: 30944008]
66. Reynolds BM, Robles TF, Repetti RL. Measurement reactivity and fatigue effects in daily diary research with families. Dev Psychol 2016 Mar;52(3):442-456. [doi: 10.1037/dev0000081] [Medline: 26689757]

XSL•FO
RenderX

67. Ono M, Schneider S, Junghaenel DU, Stone AA. What affects the completion of ecological momentary assessments in chronic pain research? An individual patient data meta-analysis. J Med Internet Res 2019 Feb 5;21(2):e11398 [FREE Full text] [doi: 10.2196/11398] [Medline: 30720437]
68. Barta WD, Tennen H, Litt MD. Measurement reactivity in diary research. In: Mehl MR, Conner TS, editors. Handbook of Research Methods for Studying Daily Life. New York, USA: The Guilford Press; 2012:108-123.
69. French DP, Sutton S. Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it? Br J Health Psychol 2010 Sep;15(Pt 3):453-468. [doi: 10.1348/135910710X492341] [Medline: 20205982]
70. Rodrigues AM, O'Brien N, French DP, Glidewell L, Sniehotta FF. The question-behavior effect: genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. Health Psychol 2015 Jan;34(1):61-78. [doi: 10.1037/hea0000104] [Medline: 25133835]
71. Martinson BC, Crain AL, Sherwood NE, Hayes MG, Pronk NP, O'Connor PJ. Population reach and recruitment bias in a maintenance RCT in physically active older adults. J Phys Act Health 2010 Jan;7(1):127-135 [FREE Full text] [doi: 10.1123/jpah.7.1.127] [Medline: 20231764]
72. Cheung KL, ten Klooster PM, Smit C, de Vries H, Pieterse ME. The impact of non-response bias due to sampling in public health studies: a comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. BMC Public Health 2017 Mar 23;17(1):276 [FREE Full text] [doi: 10.1186/s12889-017-4189-8] [Medline: 28330465]
73. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. Psychol Assess 2012 Sep;24(3):713-720. [doi: 10.1037/a0026733] [Medline: 22250597]
74. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
75. Adaimi R, Thomaz E. Leveraging active learning and conditional mutual information to minimize data annotation in human activity recognition. Proc ACM Interact Mob Wearable Ubiquitous Technol 2019 Sep 9;3(3):1-23. [doi: 10.1145/3351228]

Abbreviations

ANOVA: analysis of variance
EMA: Ecological Momentary Assessment
EPIC: European Prospective Investigation into Cancer and Nutrition
ICC: intraclass correlation
MIDA: mobile image-based dietary assessment

Edited by G Eysenbach; submitted 10.07.19; peer-reviewed by CKF Wen, Y Li, S Schembre; comments to author 03.10.19; revised version received 16.03.20; accepted 03.08.20; published 14.10.20.

Please cite as:
Ziesemer K, König LM, Boushey CJ, Villinger K, Wahl DR, Butscher S, Müller J, Reiterer H, Schupp HT, Renner B
Occurrence of and Reasons for "Missing Events" in Mobile Dietary Assessments: Results From Three Event-Based Ecological Momentary Assessment Studies
JMIR Mhealth Uhealth 2020;8(10):e15430
URL: https://mhealth.jmir.org/2020/10/e15430
doi:10.2196/15430
PMID:33052123

©Katrin Ziesemer, Laura Maria König, Carol Jo Boushey, Karoline Villinger, Deborah Ronja Wahl, Simon Butscher, Jens Müller, Harald Reiterer, Harald Thomas Schupp, Britta Renner. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 14.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Event-Level Association Between Daily Alcohol Use and Same-Day Nonadherence to Antiretroviral Therapy Among Young Men Who Have Sex With Men and Trans Women Living With HIV: Intensive Longitudinal Study

Caitlin Marie Turner[1,2], MPH; Dillon Trujillo[1], MPH; Victory Le[1,3], MA; Erin C Wilson[1,4], MPH, DrPH; Sean Arayasirikul[1,5], PhD

[1]Trans Research Unit for Equity, Center for Public Health Research, San Francisco Department of Public Health, San Francisco, CA, United States

[2]Doctoral Program in Epidemiology & Translational Science, Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, United States

[3]California School of Professional Psychology, Alliant International University, San Francisco, CA, United States

[4]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, United States

[5]Department of Pediatrics, School of Medicine, University of California, San Francisco, CA, United States

Corresponding Author:
Caitlin Marie Turner, MPH
Trans Research Unit for Equity
Center for Public Health Research
San Francisco Department of Public Health
25 Van Ness, Ste 500
San Francisco, CA, 94102
United States
Phone: 1 4155549018
Email: caitlin.turner@sfdph.org

Abstract

Background: Young trans women (TW) and men who have sex with men (MSM) are disproportionately impacted by HIV. Optimizing adherence to antiretroviral therapy (ART) is one mechanism by which public health experts aim to achieve favorable HIV health outcomes while reducing disease transmission. However, alcohol use is prevalent among young TW and MSM and threatens optimal adherence. In addition, the daily variations in alcohol use and ART adherence and their association with each other are poorly understood, warranting more appropriate methodological approaches, such as analysis of ecological momentary assessment (EMA) data.

Objective: The aim of this analysis is to characterize the association between daily alcohol use and same-day ART nonadherence captured by an EMA study of young MSM and TW living with HIV in San Francisco.

Methods: Young MSM and TW enrolled in the Health eNav digital HIV care navigation intervention were included in the analytic sample (N=113). Data on alcohol and ART use were collected by daily EMA surveys administered via text messaging and were analyzed over 30 days of follow-up. A multivariable mixed-effects logistic regression model adjusting for baseline sociodemographic characteristics was specified to investigate whether daily alcohol use was associated with same-day ART nonuse.

Results: Daily alcohol use was associated with higher same-day ART nonuse. On average, participants drank alcohol on 15.20 (SD 8.93) days and used ART on 15.19 (SD 10.16) days out of 30 days. Daily alcohol use was associated with 1.89 (95% CI 1.14-3.15) times the adjusted odds of same-day ART nonuse for each participant.

Conclusions: Results are consistent with other analyses of daily alcohol and ART use and underscore the importance of individually targeted interventions that are sensitive to each participant's dynamic risk environment.

International Registered Report Identifier (IRRID): RR2-10.2196/16406

(JMIR Mhealth Uhealth 2020;8(10):e22733) doi:10.2196/22733

XSL•FO
RenderX

KEYWORDS

ecological momentary assessment; event-level; alcohol; antiretroviral therapy; men who have sex with men; trans women; adherence

Introduction

Men who have sex with men (MSM) and trans women (TW) are disproportionately affected by HIV. MSM made up 70% of new HIV diagnoses in the United States in 2017 [1]. Approximately 21.7% of TW in the United States live with HIV and are over 30 times more likely to live with HIV compared to the rest of the US adult population [2].

The Joint United Nations Programme on HIV/AIDS (UNAIDS) developed 90-90-90 targets for 2020 to increase HIV care outcomes among all people living with HIV [3]. The last target—antiretroviral therapy (ART) uptake and adherence—is essential for viral suppression. Nonadherence to ART negatively impacts HIV progression [4,5] and is a substantial barrier to achieving favorable HIV health outcomes. Optimal ART adherence is achieved by only 63.4% of adults living with HIV [6,7].

With HIV-related stigma and other systems of oppression that target sexual and gender minority people, ART adherence is even poorer among MSM and TW. According to a 2018 HIV surveillance report by the Centers for Disease Control and Prevention (CDC), only 58% of gay and bisexual men living with HIV were virally suppressed [8]. In 2017, only 64.3% of TW in San Francisco had an undetectable viral load [9]. Low viral suppression among MSM and TW underscores the barriers to ART uptake and adherence for these populations.

Researchers have identified a number of intervention targets for improving ART adherence [5], including alcohol use [10,11]. Some mechanistic pathways involving alcohol use and ART adherence have been proposed. It is theorized that alcohol use impairs immunological functioning and viral suppression, and that this pathway is mediated by ART nonadherence [12,13]. Alcohol use could also impair cognitive functioning and subsequently interfere with the ability to adhere to HIV medications [14]. Alcohol use is prevalent among those living with HIV [15-18] and among MSM and TW [19-22], as it is hypothesized to be a coping strategy for psychological distress associated with HIV-related stigma [23] and discrimination based on sexual [24] or gender identity [25].

Alcohol use was associated with ART nonadherence in previous studies [10,11,26]; however, such studies are lacking among individuals living with HIV who have marginalized sexual and gender identities. Moreover, most studies of the effects of alcohol use on ART adherence are retrospective or gather data infrequently [27], failing to capture the temporality between alcohol and ART use and the day-to-day fluctuations inherent in both of these behaviors.

Ecological momentary assessment (EMA) data have the potential to characterize day-to-day or more frequent variations in health behaviors. EMA is a real-time data collection technique administered via technological platforms, such as handheld devices or mobile phones. A number of previous studies have used EMAs to capture fine-grained variations in substance use [28-31] and, more recently, behaviors among people living with HIV [28,32-34]. A number of studies have shown that EMA is feasible and acceptable among MSM [34-37] and persons who use substances [28,29]; one study showed moderate compliance to EMA among young MSM and TW living with HIV [38].

To our knowledge, 2 EMA studies were conducted to evaluate the relationship between alcohol use and ART adherence [39,40]. Parsons and colleagues [39] found that alcohol use was associated with 9 times the odds of HIV medication nonadherence among 272 HIV-positive men and women over 14 days of follow-up. Barai and colleagues [40] discovered that alcohol use was associated with lower odds of viral medication adherence and viral suppression among 234 women living with HIV. However, no studies have examined the alcohol use and ART nonadherence association in young MSM and TW living with HIV.

Given these research gaps, this study utilizes data from a relatively large sample of MSM and TW who participated in EMA as part of a larger HIV digital care navigation intervention. Using this data, we specify an intensive longitudinal model that assesses whether event-level alcohol use is associated with same-day ART nonadherence among young MSM and TW living with HIV in San Francisco.

Methods

Data for this analysis come from the study of Health eNav, a Digital HIV care navigation intervention conducted at the San Francisco Department of Public Health from 2017 to 2018. Study procedures were approved by the University of California, San Francisco (IRB #16-19675). Health eNav is a digital HIV care navigation intervention that employs SMS text messaging to improve HIV care continuum outcomes for young MSM and TW living with HIV in San Francisco.

Participants

Eligibility criteria were defined as follows: (1) self-identifying as MSM or TW; (2) aged 18-34 years; (3) living in San Francisco; and (4) newly diagnosed with HIV, not engaged/retained in care, or not virally suppressed. Specifically, we defined new HIV diagnoses as those that occurred within the last 12 months of enrollment in the study. If potential participants missed more than 2 HIV care appointments in the last year, they were considered as not engaged or retained in care. Potential participants who had a detectable viral load were considered as not virally suppressed. Potential participants were recruited via convenience sampling from 5 clinics and community-based organizations serving young people living with HIV in San Francisco. Study recruitment was advertised with posters, palm cards, and presentations; staff referred potential participants to the study through phone, email, or in-person meetings. The enrolled participants were also invited to refer peers from within their social networks.

XSL•FO
RenderX

An in-person or telephone eligibility screening was administered to recruits. Eligible participants then met with research staff situated at the local health department to obtain informed consent and enroll into the study. Overall, 171 people were screened. Out of those, 140 were eligible, and 120 participants enrolled in the study. Out of these 120 participants, 113 (94.2%) participants engaged in the EMA component of the study and comprised the analytical sample of this paper.

Data Collection, Measures, and Variable Selection

Procedures for the Health eNav study are described in depth in a prior protocol [41]. Briefly, Health eNav was a 6-month digital HIV care navigation intervention among young MSM and TW living with HIV, a disproportionate number of whom experienced gaps in HIV care and subsequent disparities in ART use and viral suppression. All Health eNav participants were connected to a digital HIV care navigator who facilitated linkage to, engagement in, and retention in HIV care via 2-way SMS text messaging. This study analyzes data collected from the following sources:

1. Daily EMA text surveys delivered during the first 90 days of the larger intervention, focused on capturing daily substance use, affect, sexual behaviors, and ART use in the 24 hours prior to receiving each survey.
2. Computer-administered self-interviewing (CASI) surveys at baseline, 6-, 12-, and 18-month follow-up that gathered information on participants' sociodemographic characteristics and HIV-related care outcomes.

We assessed the possible association between daily alcohol use and same-day ART adherence (measured by the EMA surveys) after adjusting for covariates (measured by baseline CASIs). We restricted our analysis to the first 30 days of EMA participation to be consistent with other EMA studies. Moreover, we observed that restricting to 30 days allowed us to maximize the nuance with which we characterized alcohol use and ART while minimizing any missingness due to participant fatigue.

Participants received automated SMS text message surveys once per day at the time of their choice (from among 8 AM, noon, or 8 PM) for the first 90 days of the intervention. The text survey was delivered through mSurvey [42]. Participants were required to complete EMA surveys within 24 hours, or the survey would time out. Each EMA survey comprised anywhere from 17 to 31 daily EMA texts depending on the responses. For example, if a participant reported on the EMA survey that they had sex within the last 24 hours, they would then receive follow-up questions about whether condoms were used and whether certain substances were used during sex. Had they not reported having sex, they would not receive questions about condom use or concurrent substance use. EMA surveys took 5 minutes or less to complete each day.

Participants were compensated US $1 for each completed EMA survey for up to US $90 over the EMA portion of the study. If participants completed more than 80% of their EMA surveys, they earned a bonus of US $100. Incentives were provided in the form of a gift card.

Data on alcohol use and ART adherence were collected daily for 30 days, and participants were inquired about their alcohol and ART use within the 24 hours prior to each EMA. We dichotomized daily alcohol use as "any use" versus "no use." For ease with interpretation (since we hypothesized that alcohol use would be associated with nonuse of ART), we operationalized the outcome as daily ART nonuse compared to ART use (reference group). ART use was gathered from a question on the EMA survey that asked, "In the past 24 hours, did you take your ART meds?"

Factors that we hypothesized would confound the relationship between alcohol and ART use were selected a priori based on the creation of a directed acyclic graph. Age (in years); race/ethnicity (Black/African American, Hispanic/Latinx, other/multiple, or White); and education level (less than high school, high school or General Educational Development [GED], and at least some college) were included for adjustment in the main analytic model. We also included housing status (living with a family member, friend, or partner who rents/owns a home; living in temporary/transitional housing; experiencing homelessness; or renting/owning a home), recent incarceration, and competing needs (eg, foregoing HIV medications to afford basic needs such as food, housing, or clothing, and vice versa) as baseline covariates. We hypothesized that young TW and MSM who were recently diagnosed with HIV might experience heightened alcohol use and barriers to ART due to newly experienced stigma and identity development related to seroconversion. Therefore, we included HIV diagnosis timing (within the last 12 months of the baseline survey versus more than 12 months before baseline) as a covariate. Finally, we included number of substances (excluding alcohol or tobacco and including any combination of marijuana, heroin, methamphetamine, amphetamines, hallucinogens, crack/cocaine, heroin, opiates, or poppers) used at baseline, defined as 0, 1, or more than 1.

Statistical Analysis

All analyses were conducted using Stata 14 software (StataCorp). Baseline sociodemographic characteristics, HIV diagnosis timing, and substance use, as gathered by the CASIs, were described for the entire sample. Alcohol use and ART nonuse were characterized over the 30-day period using EMA data. The main analysis, testing the association between daily alcohol use and same-day ART nonuse, comprised a mixed-effects regression model with a random intercept for each participant, adjusting for the aforementioned covariates.

Results

Table 1 shows the demographic breakdown of participants in the analytic sample. The average age was about 28 years. Participants were racially/ethnically diverse, with most identifying as Black/African American, Hispanic/Latinx, or "Other"/multiple races, and about a quarter of the sample identifying as White. Though most participants had some college education or more (66/113, 58.4%), only a third of the sample rented or owned their living space. More than a quarter of the sample experienced competing needs (foregoing HIV medications to afford basic needs and vice versa). About a third of participants had been recently diagnosed with HIV, and the

XSL•FO
RenderX

majority of participants reported using more than 2 substances (other than alcohol and tobacco) in the last 6 months.

Out of the 3390 total EMA surveys sent, 2022 (59.7%) were completed across Health eNav participants over the 30-day follow-up period. The median number of surveys completed, out of a total of 30 possible surveys, was 20 (IQR 8-27). There was a 97.85% overlap in missing values for ART and alcohol use; that is, most participants who did not respond to whether or not they used ART on a given day also did not respond to whether or not they used alcohol on that day. Alcohol use was reported on 19.1% of completed EMA surveys (386/2022); ART nonadherence was reported on 15.8% (320/2022) of completed EMA surveys.

Table 1. Baseline sociodemographic characteristics, HIV diagnosis timing, and substance use among young men who have sex with men and trans women living with HIV who participated in ecological momentary assessment text surveys over 30 days of follow-up, Health eNav (N=113), San Francisco, 2017-2019.

Sociodemographic characteristics	Values[a]
Age (years), mean (SD)	27.7 (3.96)
Race/ethnicity, n (%)	
Black, non-Hispanic/Latinx	22 (19.47)
Hispanic/Latinx	37 (32.74)
Other or multiple races, non-Hispanic/Latinx	26 (23.01)
White, non-Hispanic/Latinx	28 (24.78)
Education level, n (%)	
Bachelor's or higher	11 (9.73)
Some college or Associate's degree	55 (48.67)
High school/GED[b]	36 (31.86)
Less than high school	11 (9.73)
Current living situation, n (%)	
Rent/own	36 (31.86)
Lives with a friend, partner, or family member	20 (17.70)
Temporary or transitional housing	41 (36.28)
Homeless/shelter	16 (14.16)
Went without HIV medications to afford basic needs, last 6 months, n (%)	
No	76 (67.26)
Yes	37 (32.74)
Went without basic needs to afford HIV medications, last 6 months, n (%)	
No	83 (73.45)
Yes	30 (26.55)
Incarcerated, last 6 months, n (%)	
No	94 (83.19)
Yes	19 (16.81)
HIV diagnosis timing (when diagnosed with HIV), n (%)	
Diagnosed more than 12 months prior to baseline	78 (69.03)
Diagnosed within 12 months prior to baseline survey	35 (30.97)
Substance use (number of substances used [other than alcohol and tobacco]) in last 6 months, n (%)	
0	22 (19.47)
1	24 (21.24)
2 or more	67 (59.29)

[a]Percentages calculated out of the total number of participants who participated in EMA surveys (N=113), unless otherwise specified.

[b]GED: General Educational Development.

XSL•FO
RenderX

Analysis of EMA data revealed that, on average, participants drank alcohol on 15.20 (SD 8.93) days and used ART on 15.19 (SD 10.16) days out of 30 days. Event-level alcohol use was associated with 1.89 (95% CI 1.14-3.15) times the adjusted odds of same-day ART nonuse for each participant (Table 2).

XSL•FO
RenderX

Table 2. Mixed-effects model assessing daily alcohol use and same-day ART nonadherence among young MSM and TW living with HIV who participated in EMA text surveys over 30 days of follow-up, Health eNav (N=113), San Francisco, 2017-2019.

Variables	Mixed-effects logistic regression	
	AOR[a] (95% CI)	*P* value
Exposure (used alcohol in the last 24 hours)		
No	Ref[b]	N/A[c]
Yes	1.89 (1.14-3.15)	.01
Age	1.04 (0.92-1.19)	.52
Race/ethnicity		
White, non-Hispanic/Latinx	Ref	N/A
Black, non-Hispanic/Latinx	3.97 (0.91-17.41)	.07
Hispanic/Latinx	0.94 (0.28-3.20)	.92
Other or multiple races, non-Hispanic/Latinx	0.27 (0.07-1.09)	.07
Education level		
Bachelor's or higher	Ref	N/A
Some college or Associate's degree	1.77 (0.35-9.10)	.49
High school/GED[d]	8.96 (1.54-52.15)	.02
Less than high school	2.25 (0.21-23.80)	.50
Current living situation		
Rent/own	Ref	N/A
Lives with a friend, partner, or family member	0.34 (0.08-1.44)	.14
Temporary or transitional housing	0.33 (0.10-1.11)	.07
Homeless/shelter	0.73 (0.16-3.37)	.68
Went without HIV medications to afford basic needs, last 6 months		
No	Ref	N/A
Yes	3.67 (1.06-12.71)	.04
Went without basic needs to afford HIV medications, last 6 months		
No	Ref	N/A
Yes	0.26 (0.07-1.04)	.06
Incarcerated, last 6 months		
No	Ref	N/A
Yes	0.77 (0.19-3.14)	.71
HIV diagnosis timing (when diagnosed with HIV)		
Diagnosed more than 12 months prior to baseline	Ref	N/A
Diagnosed within 12 months prior to baseline survey	0.28 (0.09-0.89)	.03
Substance use (number of substances used [other than alcohol and tobacco]) in last 6 months		
0	Ref	N/A
1	0.55 (0.12-2.51)	.44
2 or more	1.38 (0.40-4.71)	.61

[a]AOR: adjusted odd ratios of same-day ART nonuse for daily alcohol use compared to nonuse, adjusting for age, race/ethnicity, education level, current living situation, competing needs, incarceration, HIV diagnosis timing, and substance use.

[b]Ref: reference.

[c]N/A: not applicable.

[d]GED: General Educational Development.

XSL•FO
RenderX

Discussion

In summary, daily alcohol use was associated with higher same-day ART nonuse. Findings from this analysis corroborate other studies of the association between alcohol use and ART nonadherence, many of which collected data cross-sectionally or retrospectively. A meta-analysis of studies examining this association showed that the combined odds of ART adherence were lower among study participants who used alcohol. However, the authors noted that effect estimation would be improved with prospective, event-level examinations of alcohol and ART use and by assessing this association for different sociodemographic subgroups [43]. This analysis uses prospective, event-level data and examines this association for young MSM and TW living with HIV, 2 populations disproportionately affected by HIV. Showing a longitudinal association provides stronger evidence that interventions targeting alcohol use may also improve ART adherence, which is an important step in achieving optimal HIV care outcomes [3].

Findings from this study should be interpreted with a number of limitations in mind. First, intermittent missingness and dropout are issues inherent with EMA data collection due to increased burden on participants [44], and these issues were present in our sample. Full or partial multiple imputations have been recommended as the best approach for improving precision of estimates where missingness is an issue [44]. However, practical applications of these methods, especially with respect to longitudinal modeling, have produced mixed results. One study found that multiple imputations produced similar or even increased standard errors [45] compared with complete case data. With these observations in mind, we chose to analyze complete case data. Thus, our results are only generalizable to the participants for whom we had complete data.

Second, data were collected via self-report, and participants may have underreported their alcohol use and ART nonadherence behaviors. This could have potentially washed out the estimated effect. However, the results we observed still showed statistical significance. A third limitation was misclassification of exposure. Dichotomizing alcohol use as “any versus none” effectively grouped together participants who drank casually with those who drank heavily. However, such dichotomizing was necessary to preserve precision of estimates. In addition, we suspected that this misclassification was nondifferential and independent, which would bias results toward the null. If anything, the effect estimate we observed underestimated the true, underlying association. A fourth limitation was the small size of the sample, precluding stratified analyses by key demographics within which the alcohol-ART association may have varied. However, given the daily administration of EMA surveys, we are confident that the effect estimate provided was statistically precise for the entire sample. A fifth limitation was that we restricted the follow-up window of EMA surveys to 30 days. However, this analytic decision was informed by prior EMA studies conducted over 30-day-or-shorter time windows. Since EMA data collection was embedded within the larger digital care navigation intervention, losses-to-follow-up attributable to participation within the larger intervention may have impacted engagement in EMA outside of the 30-day window. Moreover, expanding the EMA follow-up window beyond 30 days could have produced an effect estimate that was less representative of the entire sample since participant fatigue and subsequent nonresponses were more of an issue beyond the first 30 days. Finally, results were not generalizable outside of young MSM and TW living in San Francisco.

The limitations of this analysis pave the way for future research. Given the dearth of research on moderation of the alcohol-ART association by key sociodemographic characteristics such as gender and race [43], future studies should utilize EMA methods in other key subgroups. TW living with HIV, many of whom experience pervasive gaps in HIV care and clinical outcomes, represented only a small percentage of the participants in this study, precluding statistically precise estimations of the effect of alcohol use on ART nonadherence for this subgroup. Future research should be conducted on larger samples of TW to explore how the alcohol-ART association varies by gender. Since this analysis was underpowered to assess interactions between alcohol use, substance use, and mental health, future studies with larger sample sizes could centralize comorbidities between those factors. In addition to focusing on individual behaviors, these future studies could examine the multilevel interplay between alcohol use and structural factors like racism, housing instability, or competing needs and their impact on ART adherence. Alcohol use is one of many modifiable factors that could affect adherence to HIV medications.

Finally, future EMA studies would benefit from a thorough consideration of institutional and individual barriers to EMA survey completion in order to reduce missing responses. A previous analysis of Health eNav EMA data confirmed that housing instability, incarceration, competing needs, and educational constraints interfered with EMA completion for young MSM and TW, even though participants had continuous access to cellular devices [38]. While one recommendation to preemptively address participant nonresponse would be to invest in procedures that sustain retention over the follow-up period, structural barriers to such an investment (eg, systemic marginalization of the populations served or lack of grant funding) highlight the unrealistic nature of such a recommendation. Dismantling systemic oppression would best improve study retention and even remove the need to have studies on health inequities in the first place. Until then, studies such as this one highlight the need to apply tailored approaches to implementation of digital interventions within under- and misrepresented populations.

To our knowledge, this is the first intensive longitudinal analysis of alcohol and ART use among young MSM and TW living with HIV. This analysis highlights important considerations in examining daily ART use among populations especially vulnerable to substance use and medication nonadherence.

XSL•FO
RenderX

Acknowledgments

The authors would like to acknowledge the Health eNav study staff and participants. In addition, we would like to thank Mi-Ok Kim, PhD, and David Glidden, PhD, at the University of California, San Francisco, for their guidance in conceptualizing the analytic method employed in this paper. This work was supported by the Health Resources and Services Administration (award number H97HA28895).

Conflicts of Interest

None declared.

References

1. Centers for Disease Control and Prevention. HIV Surveillance Report, 2017; volume 29. URL: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2020-06-01]
2. Baral SD, Poteat T, Strömdahl S, Wirtz AL, Guadamuz TE, Beyrer C. Worldwide burden of HIV in transgender women: a systematic review and meta-analysis. Lancet Infect Dis 2013 Mar;13(3):214-222. [doi: 10.1016/S1473-3099(12)70315-8] [Medline: 23260128]
3. 90-90-90: An ambitious treatment target to help end the AIDS epidemic. Joint United Nations Programme on HIV/AIDS. 2014. URL: http://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf [accessed 2020-06-01]
4. Bangsberg DR, Perry S, Charlebois ED, Clark RA, Roberston M, Zolopa AR, et al. Non-adherence to highly active antiretroviral therapy predicts progression to AIDS. AIDS 2001 Jun 15;15(9):1181-1183. [doi: 10.1097/00002030-200106150-00015] [Medline: 11416722]
5. Simoni JM, Amico KR, Pearson CR, Malow R. Strategies for promoting adherence to antiretroviral therapy: a review of the literature. Curr Infect Dis Rep 2008 Nov;10(6):515-521 [FREE Full text] [doi: 10.1007/s11908-008-0083-y] [Medline: 18945394]
6. Centers for Disease Control and Prevention. Monitoring selected national HIV prevention and care objectives by using HIV surveillance data—United States and 6 dependent areas, 2017. HIV Surveillance Supplemental Report, Volume 24, Number 3. 2019 Jun. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-supplemental-report-vol-24-3.pdf [accessed 2020-10-01]
7. Bezabhe WM, Chalmers L, Bereznicki LR, Peterson GM. Adherence to Antiretroviral Therapy and Virologic Failure: A Meta-Analysis. Medicine (Baltimore) 2016 Apr;95(15):e3361 [FREE Full text] [doi: 10.1097/MD.0000000000003361] [Medline: 27082595]
8. Selected National HIV Prevention and Care Outcomes. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/hiv/pdf/library/slidesets/cdc-hiv-prevention-and-care-outcomes-2018.pdf [accessed 2019-06-01]
9. Baguso GN, Turner CM, Santos G, Raymond HF, Dawson-Rose C, Lin J, et al. Successes and final challenges along the HIV care continuum with transwomen in San Francisco. J Int AIDS Soc 2019 Apr;22(4):e25270 [FREE Full text] [doi: 10.1002/jia2.25270] [Medline: 31037858]
10. Cook RL, Sereika SM, Hunt SC, Woodward WC, Erlen JA, Conigliaro J. Problem drinking and medication adherence among persons with HIV infection. J Gen Intern Med 2001 Feb;16(2):83-88. [doi: 10.1111/j.1525-1497.2001.00122.x] [Medline: 11251758]
11. Sankar A, Wunderlich T, Neufeld S, Luborsky M. Sero-positive African Americans' beliefs about alcohol and their impact on anti-retroviral adherence. AIDS Behav 2007 Mar;11(2):195-203 [FREE Full text] [doi: 10.1007/s10461-006-9144-0] [Medline: 16799842]
12. Chander G, Lau B, Moore RD. Hazardous alcohol use: a risk factor for non-adherence and lack of suppression in HIV infection. J Acquir Immune Defic Syndr 2006 Dec 01;43(4):411-417 [FREE Full text] [doi: 10.1097/01.qai.0000243121.44659.a4] [Medline: 17099312]
13. Samet JH, Cheng DM, Libman H, Nunes DP, Alperen JK, Saitz R. Alcohol consumption and HIV disease progression. J Acquir Immune Defic Syndr 2007 Oct 01;46(2):194-199 [FREE Full text] [doi: 10.1097/QAI.0b013e318142aabb] [Medline: 17667330]
14. Braithwaite R, Conigliaro J, McGinnis K, Maisto S, Bryant K, Justice A. Adjusting alcohol quantity for mean consumption and intoxication threshold improves prediction of nonadherence in HIV patients and HIV-negative controls. Alcohol Clin Exp Res 2008 Sep;32(9):1645-1651 [FREE Full text] [doi: 10.1111/j.1530-0277.2008.00732.x] [Medline: 18616666]
15. Conigliaro J, Justice AC, Gordon AJ, Bryant K, VACS AlcoholBehavior Change Research Group. Role of alcohol in determining human immunodeficiency virus (HIV)-relevant outcomes: A conceptual model to guide the implementation of evidence-based interventions into practice. Med Care 2006 Aug;44(8 Suppl 2):S1-S6. [doi: 10.1097/01.mlr.0000223659.36369.cf] [Medline: 16849963]
16. Bryant KJ. Expanding research on the role of alcohol consumption and related risks in the prevention and treatment of HIV/AIDS. Subst Use Misuse 2006;41(10-12):1465-1507. [doi: 10.1080/10826080600846250] [Medline: 17002990]

XSL•FO
RenderX

17. Galvan FH, Bing EG, Fleishman JA, London AS, Caetano R, Burnam MA, et al. The prevalence of alcohol consumption and heavy drinking among people with HIV in the United States: results from the HIV Cost and Services Utilization Study. J Stud Alcohol 2002 Mar;63(2):179-186. [doi: 10.15288/jsa.2002.63.179] [Medline: 12033694]
18. Samet JH, Horton NJ, Meli S, Freedberg KA, Palepu A. Alcohol consumption and antiretroviral adherence among HIV-infected persons with alcohol problems. Alcohol Clin Exp Res 2004 Apr;28(4):572-577. [doi: 10.1097/01.alc.0000122103.74491.78] [Medline: 15100608]
19. Santos G, Jin H, Raymond HF. Pervasive Heavy Alcohol Use and Correlates of Increasing Levels of Binge Drinking among Men Who Have Sex with Men, San Francisco, 2011. J Urban Health 2015 Aug;92(4):687-700 [FREE Full text] [doi: 10.1007/s11524-015-9958-z] [Medline: 25895622]
20. Hess KL, Chavez PR, Kanny D, DiNenno E, Lansky A, Paz-Bailey G, NHBS Study Group. Binge drinking and risky sexual behavior among HIV-negative and unknown HIV status men who have sex with men, 20 US cities. Drug Alcohol Depend 2015 Feb 01;147:46-52 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.12.013] [Medline: 25555622]
21. Institute of Medicine. The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding. Washington DC: The National Academies Press; 2011.
22. James S, Herman J, Rankin S, Keisling M, Mottet L, Anafi M. The Report of the 2015 US Transgender Survey. National Center for Transgender Equality; Washington DC. URL: https://www.transequality.org/sites/default/files/docs/USTS-Full-Report-FINAL.PDF [accessed 2020-06-01]
23. Wardell J, Shuper P, Rourke S, Hendershot C. Stigma, Coping, and Alcohol Use Severity Among People Living With HIV: A Prospective Analysis of Bidirectional and Mediated Associations. Ann Behav Med 2018 Aug 16;52(9):762-772 [FREE Full text] [doi: 10.1093/abm/kax050] [Medline: 30124756]
24. Batchelder AW, Safren S, Mitchell AD, Ivardic I, O'Cleirigh C. Mental health in 2020 for men who have sex with men in the United States. Sex Health 2017 Feb;14(1):59-71 [FREE Full text] [doi: 10.1071/SH16083] [Medline: 28055823]
25. Rowe C, Santos G, McFarland W, Wilson EC. Prevalence and correlates of substance use among trans female youth ages 16-24 years in the San Francisco Bay Area. Drug Alcohol Depend 2015 Feb 01;147:160-166 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.11.023] [Medline: 25548025]
26. Gonzalez A, Barinas J, O'Cleirigh C. Substance use: impact on adherence and HIV medical treatment. Curr HIV/AIDS Rep 2011 Dec;8(4):223-234. [doi: 10.1007/s11904-011-0093-5] [Medline: 21858414]
27. Spire B, Duran S, Souville M, Leport C, Raffi F, Moatti J, APROCO cohort study group. Adherence to highly active antiretroviral therapies (HAART) in HIV-infected patients: from a predictive to a dynamic approach. Soc Sci Med 2002 May;54(10):1481-1496. [doi: 10.1016/s0277-9536(01)00125-3] [Medline: 12061483]
28. Turner CM, Coffin P, Santos D, Huffaker S, Matheson T, Euren J, et al. Race/ethnicity, education, and age are associated with engagement in ecological momentary assessment text messaging among substance-using MSM in San Francisco. J Subst Abuse Treat 2017 Apr;75:43-48 [FREE Full text] [doi: 10.1016/j.jsat.2017.01.007] [Medline: 28237053]
29. Yang C, Linas B, Kirk G, Bollinger R, Chang L, Chander G, et al. Feasibility and Acceptability of Smartphone-Based Ecological Momentary Assessment of Alcohol Use Among African American Men Who Have Sex With Men in Baltimore. JMIR Mhealth Uhealth 2015 Jun 17;3(2):e67 [FREE Full text] [doi: 10.2196/mhealth.4344] [Medline: 26085078]
30. Rendina HJ, Ventuneac A, Mustanski B, Grov C, Parsons JT. Prospective Measurement of Daily Health Behaviors: Modeling Temporal Patterns in Missing Data, Sexual Behavior, and Substance Use in an Online Daily Diary Study of Gay and Bisexual Men. AIDS Behav 2016 Aug;20(8):1730-1743 [FREE Full text] [doi: 10.1007/s10461-016-1359-0] [Medline: 26992392]
31. Westergaard RP, Genz A, Panico K, Surkan PJ, Keruly J, Hutton HE, et al. Acceptability of a mobile health intervention to enhance HIV care coordination for patients with substance use disorders. Addict Sci Clin Pract 2017 Apr 26;12(1):11 [FREE Full text] [doi: 10.1186/s13722-017-0076-y] [Medline: 28441962]
32. Hensel DJ, Fortenberry JD, Harezlak J, Craig D. The feasibility of cell phone based electronic diaries for STI/HIV research. BMC Med Res Methodol 2012 Jun 12;12:75 [FREE Full text] [doi: 10.1186/1471-2288-12-75] [Medline: 22691189]
33. Moore RC, Kaufmann CN, Rooney AS, Moore DJ, Eyler LT, Granholm E, et al. Feasibility and Acceptability of Ecological Momentary Assessment of Daily Functioning Among Older Adults with HIV. Am J Geriatr Psychiatry 2017 Aug;25(8):829-840 [FREE Full text] [doi: 10.1016/j.jagp.2016.11.019] [Medline: 28065496]
34. Cherenack EM, Wilson PA, Kreuzman AM, Price GN, Adolescent Medicine Trials Network for HIV/AIDS Interventions. The Feasibility and Acceptability of Using Technology-Based Daily Diaries with HIV-Infected Young Men Who have Sex with Men: A Comparison of Internet and Voice Modalities. AIDS Behav 2016 Aug;20(8):1744-1753 [FREE Full text] [doi: 10.1007/s10461-016-1302-4] [Medline: 26837629]
35. Kingdon MJ, Storholm ED, Halkitis PN, Jones DC, Moeller RW, Siconolfi D, et al. Targeting HIV prevention messaging to a new generation of gay, bisexual, and other young men who have sex with men. J Health Commun 2013;18(3):325-342. [doi: 10.1080/10810730.2012.727953] [Medline: 23320963]
36. Kubicek K, Carpineto J, McDavitt B, Weiss G, Iverson EF, Au C, et al. Integrating professional and folk models of HIV risk: YMSM's perceptions of high-risk sex. AIDS Educ Prev 2008 Jun;20(3):220-238 [FREE Full text] [doi: 10.1521/aeap.2008.20.3.220] [Medline: 18558819]

XSL•FO
RenderX

37. Kubicek K, Carpineto J, McDavitt B, Weiss G, Kipke MD. Use and perceptions of the internet for sexual information and partners: a study of young men who have sex with men. Arch Sex Behav 2011 Aug;40(4):803-816 [FREE Full text] [doi: 10.1007/s10508-010-9666-4] [Medline: 20809373]
38. Turner CM, Arayasirikul S, Trujillo D, Lê V, Wilson EC. Social Inequity and Structural Barriers to Completion of Ecological Momentary Assessments for Young Men Who Have Sex With Men and Trans Women Living With HIV in San Francisco. JMIR Mhealth Uhealth 2019 May 08;7(5):e13241 [FREE Full text] [doi: 10.2196/13241] [Medline: 31066714]
39. Parsons JT, Rosof E, Mustanski B. The temporal relationship between alcohol consumption and HIV-medication adherence: a multilevel model of direct and moderating effects. Health Psychol 2008 Sep;27(5):628-637 [FREE Full text] [doi: 10.1037/a0012664] [Medline: 18823189]
40. Barai N, Monroe A, Lesko C, Lau B, Hutton H, Yang C, et al. The Association Between Changes in Alcohol Use and Changes in Antiretroviral Therapy Adherence and Viral Suppression Among Women Living with HIV. AIDS Behav 2017 Jul;21(7):1836-1845 [FREE Full text] [doi: 10.1007/s10461-016-1580-x] [Medline: 27752873]
41. Arayasirikul S, Trujillo D, Turner CM, Le V, Wilson EC. Implementing a Digital HIV Care Navigation Intervention (Health eNav): Protocol for a Feasibility Study. JMIR Res Protoc 2019 Nov 08;8(11):e16406 [FREE Full text] [doi: 10.2196/16406] [Medline: 31702561]
42. mSurvey. URL: https://msurvey.co/ [accessed 2019-04-18]
43. Hendershot CS, Stoner SA, Pantalone DW, Simoni JM. Alcohol use and antiretroviral adherence: review and meta-analysis. J Acquir Immune Defic Syndr 2009 Oct 01;52(2):180-202 [FREE Full text] [doi: 10.1097/QAI.0b013e3181b18b6e] [Medline: 19668086]
44. Ji L, Chow S, Schermerhorn AC, Jacobson NC, Cummings EM. Handling Missing Data in the Modeling of Intensive Longitudinal Data. Struct Equ Modeling 2018;25(5):715-736 [FREE Full text] [doi: 10.1080/10705511.2017.1417046] [Medline: 31303745]
45. Young R, Johnson DR. Handling Missing Values in Longitudinal Panel Data With Multiple Imputation. J Marriage Fam 2015 Feb;77(1):277-294 [FREE Full text] [doi: 10.1111/jomf.12144] [Medline: 26113748]

Abbreviations

ART: antiretroviral therapy
CASI: computer-administered self-interviewing
CDC: Centers for Disease Control and Prevention
EMA: ecological momentary assessment
GED: General Educational Development
MSM: men who have sex with men
TW: trans women
UNAIDS: Joint United Nations Programme on HIV/AIDS

Edited by G Eysenbach; submitted 21.07.20; peer-reviewed by L Chang, C Latkin; comments to author 11.08.20; revised version received 16.09.20; accepted 16.09.20; published 15.10.20.

Please cite as:
Turner CM, Trujillo D, Le V, Wilson EC, Arayasirikul S
Event-Level Association Between Daily Alcohol Use and Same-Day Nonadherence to Antiretroviral Therapy Among Young Men Who Have Sex With Men and Trans Women Living With HIV: Intensive Longitudinal Study
JMIR Mhealth Uhealth 2020;8(10):e22733
URL: http://mhealth.jmir.org/2020/10/e22733/
doi:10.2196/22733
PMID:33055070

©Caitlin Marie Turner, Dillon Trujillo, Victory Le, Erin C Wilson, Sean Arayasirikul. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 15.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Preventing and Addressing the Stress Reactions of Health Care Workers Caring for Patients With COVID-19: Development of a Digital Platform (Be + Against COVID)

José Joaquín Mira[1,2,3,4*], PhD; María Asunción Vicente[4,5*], PhD; Adriana Lopez-Pineda[1,3*], PhD; Irene Carrillo[1,3,4*], PhD; Mercedes Guilabert[1,4*], PhD; César Fernández[4,5*], PhD; Virtudes Pérez-Jover[1,4*], PhD; Jimmy Martin Delgado[3], PhD; Pastora Pérez-Pérez[6*], PhD; Angel Cobos Vargas[7*], MSc; María Pilar Astier-Peña[8*], PhD; Olga Beatriz Martínez-García[7*], MSc; Bárbara Marco-Gómez[9*], MSc; Cristina Abad Bouzán[9*], MSc

[1]Health Psychology Department, Miguel Hernández University, Elche, Spain
[2]Alicante-Sant Joan Health District, Alicante, Spain
[3]The Foundation for the Promotion of Health and Biomedical Research of Valencia Region, Alicante, Spain
[4]Prometeo/2017/173 Excellence Group, Generalitat Valenciana, Valencia, Spain
[5]Telematics Engineering Area, Miguel Hernández University, Elche, Spain
[6]Andalusian Agency for Health Care Quality, Seville, Spain
[7]Hospital Universitario Clínico San Cecilio, Granada, Spain
[8]Centro de Salud La Jota, Zaragoza, Spain
[9]Zona de salud de Calatayud, Zaragoza, Spain
[*]these authors contributed equally

Corresponding Author:
César Fernández, PhD
Telematics Engineering Area
Miguel Hernández University
Edificio Innova
Elche, 03202
Spain
Phone: 34 966658423
Email: c.fernandez@umh.es

Abstract

Background: COVID-19 became a major public health concern in March 2020. Due to the high rate of hospitalizations for COVID-19 in a short time, health care workers and other involved staff are subjected to a large workload and high emotional distress.

Objective: The objective of this study is to develop a digital tool to provide support resources that might prevent and consider acute stress reactions in health care workers and other support staff due to the COVID-19 pandemic.

Methods: The contents of the digital platform were created through an evidence-based review and consensus conference. The website was built using the Google Blogger tool. The Android version of the app was developed in the Java and XML languages using Android Studio version 3.6, and the iOS version was developed in the Swift language using Xcode version 11.5. The app was evaluated externally by the Andalusian Agency for Healthcare Quality.

Results: We detected the needs and pressing situations of frontline health care workers, and then, we proposed a serial of recommendations and support resources to address them. These resources were redesigned using the feedback received. A website in three different languages (Spanish, English, and Portuguese) and a mobile app were developed with these contents, and the AppSaludable Quality Seal was granted to the app. A specific self-report scale to measure acute stress and additional tools were included to support the health care workforce. This instrument has been used in several Latin American countries and has been adapted considering cultural differences. The resources section of the website was the most visited with 18,516 out of 68,913 (26.9%) visits, and the “Self-Report Acute Stress Scale” was the most visited resource with 6468 out of 18,516 (34.9%) visits.

XSL•FO
RenderX

Conclusions: The *Be + against COVID* platform (website and app) was developed and launched to offer a pool of recommendations and support resources, which were specifically designed to protect the psychological well-being and the work morale of health care workers. This is an original initiative different from the usual psychological assistance hotlines.

(JMIR Mhealth Uhealth 2020;8(10):e21692) doi:10.2196/21692

KEYWORDS

COVID-19; pandemic; internet; social media; mobile app; psychosocial; support system; health personnel; app

Introduction

The World Health Organization (WHO) officially declared the COVID-19 outbreak a pandemic on March 11, 2020 [1]. This infectious disease was first identified in China in December 2019 and is caused by the novel coronavirus, designated SARS-CoV-2. As of June 1, 2020, a total of 5,891,182 confirmed cases had been reported worldwide. In April 2020, Spain became the second country in the world with the most reported COVID-19 cases, and as of June 1, 2020, it is the fifth country with 239,429 confirmed cases [2]. Moreover, Spain is the country with the highest rate of infections among health care workers (HCWs), representing 20.4% of all recorded cases [3].

Due to the fast spread of COVID-19 infections, the number of confirmed cases increased rapidly worldwide, and currently, it is a major public health concern [4]. As many patients with COVID-19 require hospitalization, many hospitals are saturated, and therefore, a lot of HCWs and other involved staff are being subjected to extraordinary workloads and high emotional stress [5]. Previous evidence has shown that adverse psychological reactions such as anxiety, stress, and fear were observed among health care professionals during and after other infectious disease outbreaks [6,7]. Currently, scientific studies are reporting mental health symptoms among HCWs treating patients with COVID-19 [8-10]. Cai et al [8] found that HCWs at the frontline of the COVID-19 epidemic in China were worried about the risks of infection and protective measures, resulting in psychological distress. A cross-sectional study in China showed that a significant proportion of frontline HCWs experienced anxiety, depression, insomnia symptoms, and psychological distress during the current COVID-19 pandemic [9].

To protect the psychological well-being and to prevent moral injuries of HCWs, it is necessary to better control infectious diseases [11], ensure correct functioning of the health system, and ensure that patients are receiving appropriate medical attention. Xiang et al [12] appealed for mental health care for the 2019 novel coronavirus outbreak. Health authorities must plan actions immediately to limit the impact that caring for patients with COVID-19 has on HCWs [9,10]. For the time being, evidence on how to deal with psychological distress and other mental health symptoms of HCWs who are involved in the current health crisis are not available [7], and the extent of intervention programs to assist second victims in health systems is limited [9]. However, special interventions to help frontline HCWs have been implemented in countries like China. Thus, the objective of this study is to develop a digital platform to provide support resources that might prevent acute stress reactions in HCWs and other support staff due to the COVID-19 outbreak.

Methods

The initial step to create the platform was the development of the content and the support resources. Information about the most common concerns for HCWs during the COVID-19 outbreak and how to address them were collected through an evidence-based review and consensus conference from March 14-28, 2020. First, we carried out an evidence-based review about previous pandemics, the psychological impact of the COVID-19 outbreak on HCWs in China, and mental health care measures. Second, an online consensus conference was conducted by the principal investigator with physicians from different specialties, nurses, clinical psychologists, and responsible persons for quality and patient safety.

All members of the consensus group had experience implementing programs to support second victims.

Through the exchange of previous experiences and existing intervention results, we first identified the groups of HCWs most vulnerable to professional overload during the COVID-19 outbreak. Second, we classified the most common problem situations for HCWs. Third, we proposed possible actions to address these concerns and prevent stress symptoms.

A series of support resources were specifically designed according to the detected needs and proposed actions.

Once the platform content was defined, a website and a mobile app were developed. Due to its simplicity and efficiency in SEO positioning, the Google Blogger [13] tool was used to build the website. The blogger theme used was responsive to facilitate the usability of the web from mobile devices. The website was published in three different language versions (Spanish, English, and Brazilian Portuguese) to reach nations with common cultural characteristics. Websites were included in subdomains of the main web address of the Miguel Hernandez University, and they use a university server to store the documents available.

Concerning the app, the two most used mobile platforms were addressed: Android and iOS. According to recent reports [14], these platforms account for 99.4% of world mobile operating system market share (Android 74.14%, iOS 25.26%), so almost all mobile phone users would be able to download the app. Regarding software architecture, it was decided to develop a native app for each platform instead of creating a hybrid app valid for Android and iOS due to the better performance of native apps. As a consequence, even though the aspect and usability of both apps (Android and iOS) was the same, their internal design and the libraries used were specific for each app.

XSL•FO
RenderX

The Android version of the app was developed in Java and XML languages using Android Studio version 3.6 (Google). The iOS version was developed in the Swift language using Xcode version 11.5 (Apple Inc). Both software development tools were installed on a 2017 iMac computer running macOS 10.15.4. To assure that the app met a set of quality criteria to be used in the health context, it was evaluated externally by the Andalusian Agency for Healthcare Quality through their certification program "Safety and Quality Strategy in Mobile Health Apps" [15]. This agency grants the AppSaludable Quality Seal if the app meets 31 requirements. These requirements are structured in four blocks: design and appropriateness, quality and safety of information, provision of services, and confidentiality and privacy. Fulfilling the requirements of an external and independent agency assures that the app quality and safety standards are adequate and that it can be recommended for health management. With regards to flexibility and scalability, both apps were designed with an easy to maintain structure so that future changes and additional functionalities could be carried out simply. In both apps (Android and iOS), the architecture allowed two ways of updating: (1) uploading new versions to Google Play and App Store markets and (2) refreshing the documents shown in the app or adding new documents, which could be done instantly without the need for uploading new versions and, thus, is faster and easier for the users.

Web traffic and visitor's preferences of the website's Spanish version as well as the app discharges were analyzed 2 months after the launch. The English and Portuguese versions of the website were launched later than the Spanish version; therefore, we considered that a longer time was needed to analyze the traffic data of these versions. Google Analytics software was used to collect the following indicators: number of visitors, number of visits, map of visitors by country, type of devices used for access, and page views. Google Analytics does not provide any personally identifiable data.

Results

Definition of the Digital Platform Content

Initially, the discussion group identified the most vulnerable group of HCWs to professional overload: HCWs at the frontline of COVID-19 who are in direct contact with patients with possible or confirmed infections and support staff of health care departments or units, such as the emergency department, primary health care, home hospitalization, critical care, reanimation unit, internal medicine, pneumology, and infectious diseases departments.

Taking into consideration the literature review on the SARS-CoV-2 pandemic and the opinions of the discussion group participants and their own experiences, the most common pressing situations for HCWs were identified and classified into five categories (see Textbox 1).

XSL•FO
RenderX

Textbox 1. Description of pressing situations for health care workers identified by category.

Organization, human resources, and materials

- To be temporarily working in health care settings for which no appropriate training has been provided
  - New recruitment specific for the crisis or transfers to more complex health care facilities and higher biological risk subjected to health care needs
- To receive instructions, sometimes contradictory, on the control of risks and procedures (with no clear assignment of tasks)
  - Inconsistencies in the chain of command and individual proposals regarding the use of personal protective equipment, such as the use of face masks
  - Shortage of material for offering the appropriate care to all patients and reducing the risk of infection
- Reduction in the number of human resources due to professionals leaving from risk exposures
  - The working time is extended, the frequency of shifts increased, and the periods of physical and mental rest reduced.
  - Tasks for which specific training has not been received are assumed or are carried out after express training. This may cause insecurity.
- Dissolution of stable work teams
  - Incorporation of new professionals, which changes the dynamics of the work group
  - Overloading of more experienced professionals
- Patients with other pathologies cannot receive the care they had been receiving.
  - Over-the-phone care is provided, which increases the risk of adverse events in many cases due to the omission of actions and low probability of detecting it.
- A different perspective from residents
  - They are not under the wing of the consultants but placed in a different situation regarding clinical attitude.
- Precrisis conflicts between team members
  - Previous conflicts may surface now because of the task distribution in extreme situations.

Environmental stressors and other stressors linked to crises situations

- To work in an environment of particular biological risk
  - This risk can also affect other patients, colleagues, and family members.
- To express substitutions for the loss of colleagues in isolation at home or resulting in cases with COVID-19
- To be overflown for periods that are increasing
  - Wanting to not appear weak or incapable to provide an answer all the time
- No clear horizon of “how long this is going to last”

Human factors

- Helplessness when witnessing reckless behaviors from patients and people who accompany them (usually due to unawareness) and mistakes among professionals (from tiredness, stress, other)
  - Involuntary errors are possible (which may lead to adverse events) and can be made by other professionals during patient care.
  - To feel helpless, irritable, and doubting your own ability, which leads to other risk situations for the patients
- Powerlessness and disaffection at seeing how patients who are afraid of being sick with COVID-19 have to be alone; unaccompanied; and, in some cases, die in solitude. This situation is having a profound impact on professionals.

Fear or panic reactions

- Fear, and occasionally panic, when finding out a colleague is under passive surveillance or kept in isolation at home
- Fear of infecting a family member or a close acquaintance
- To minimize symptoms that may indicate contagion by pressure to not leave services uncovered.

XSL•FO
RenderX

Critical decision making on health care issues

- To be obliged to make patient triage and other decisions reserved for major catastrophes that imply relevant ethical matters
- To face the decision to prioritize levels of attention, generating a new organizational situation unknown until now

Regarding the design of interventions to prevent and handle the identified problems or pressing situations, the literature review on the experiences in Wuhan and Hunan (China) hospitals was considered. Besides, we reviewed other methods for the recovery of the second victim, which were applied in different conditions but can also be applied in this context. Finally, considering the aforementioned situations (Textbox 1), the discussion group proposed possible strategies and actions to approach the most severe demands and needs. Multimedia Appendix 1 shows the detected needs and resources to face each problem situation of HCWs during the COVID-19 outbreak. For each need, many recommendations were offered and 19 support resources were created as intervention proposals. Infographics were designed to complement some of these resources (see Multimedia Appendix 2) [16]. It is noteworthy that the nineth resource, named "self-report," differs from the rest, and it consists of a COVID-19 acute stress scale for the HCWs. This is an ad hoc 10-question test to assess whether the HCWs were overwhelmed by the situation they were experiencing [17].

Digital Platform Design and Development

A digital platform was developed including the aforementioned content and support resources. The name of this platform is *Be + against COVID*, and it is composed of a website in three languages (Spanish, English, and Brazilian Portuguese) and a mobile app (for Android and iOS). The site is freely accessible to anyone with an internet connection, and the app is freely available to download from the official stores.

Website Description

The three language versions share a similar content structure, although the Spanish version is the one that contains more information due to the periodic publication of news in its blog section (so-called "Positive News" [18]). Information on the website was organized following a basic structure, that is the main menu with seven items: (1) Home, (2) Pressing situations, (3) Resources, (4) Self-Report, (5) Download, (6) App, and (7) Activities. Figure 1 shows the home page of the *Be + against COVID* website. The content of each item is described in Textbox 2. The website contents were presented as text, documents (PDF format), infographics (PDF format), and videos (image format).

XSL•FO
RenderX

Figure 1. Home page and main menu from Be + against COVID website (computer view). ENG: English; ESP: Spanish; PT_BR: Brazilian Portuguese.

XSL•FO
RenderX

Textbox 2. Be + against COVID website content.

- Home
  - Introduction
  - Purpose of this website
  - Methods
  - Project Team
  - References
- Pressing situations
  - Human factors
  - Environmental stressors
  - Fear or panic reactions
  - Organization human resources and material
  - Decision making on health care issues
- Resources
  - App Be + against COVID
  - Resources in short
  - Resource (R)1. Information on achievements and actions taken
  - R2. Involve professionals in audio-visual messages to broadcast information on guidelines (eg, safe removal of personal protective equipment)
  - R3. Daily news on the situation of the center
  - R4. Homogeneous structure of corporate messages
  - R5. A liaison for the coordination with contract employees
  - R6. Identify and refute unfounded rumors and incorrect information
  - R7. Briefings at the beginning of every shift with a particular focus on new hires
  - R8. Home isolation professionals to act as distant trainers and tutors for new hires
  - R9. Self-report measure of acute stress
  - R10. Awareness on the need to face affective response and accept support
  - R11. Mental health hotline for health professionals and support by specialized personnel
  - R12. Momentary work recovery (short breaks)
  - R13. Defusing (face-to-face or remotely): Get rid of all emotional overload before the end of the workday to avoid taking it home and recover strength for the next shift
  - R14. Referral to individual counseling to help overcome acute stress reactions
  - R15. Set up rest areas for the recovery of the professionals before the end of the shift
  - R16. Maintain contact and inform about the situation in the center. Long-distance institutional accompaniment. Facilitate reincorporation
  - R17. Become aware of the actions expected to be carried out by middle managers: Responsible leadership
  - R18. Promote informative leadership, transparency, realism, and positive messages
  - R19. Make a plan to deal with the volume of delayed health care activities: Alleviate the foreseen impact on health and support professionals of this physical and mental overload
- Self Report
- Download
  - Documents
  - Videos focused on how to cope with distress
- App
  - App information

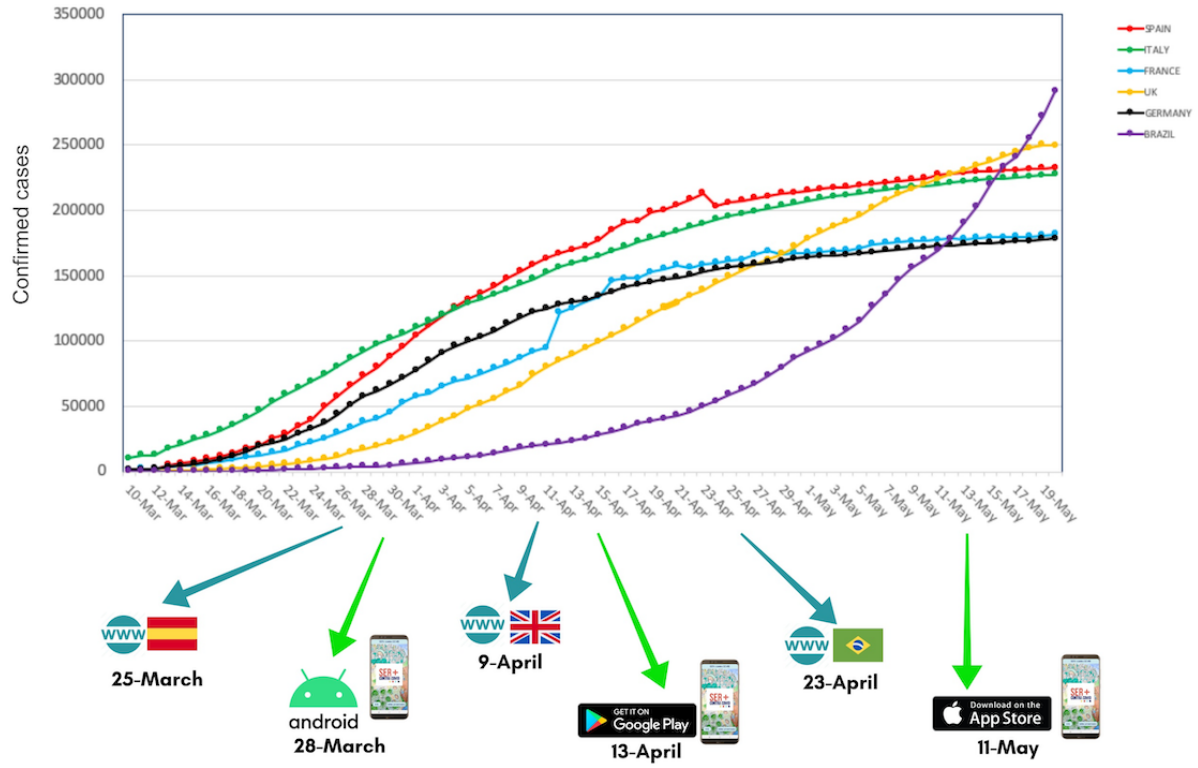
XSL•FO
RenderX

- Download from Google Play
- Download from Appstore

- Activities

The Spanish language version of the website [18] was the first to be published on March 16, 2020, because of the origin of the research team, mostly of Spanish nationality. However, it was not until March 25, 2020, that all the resources were available. The English version [19] was available from April 9, 2020, and the Brazilian Portuguese version [20] was published on April 23, 2020. Figure 2 shows the publication of each part of the platform chronologically during the most critical phase of the SARS-CoV-2 pandemic in Europe.

Figure 2. Be + against COVID platform development during the most critical phase of the SARS-CoV-2 pandemic.



Mobile App Description

To comply with the highest quality standards, the app development process included several tests, reviews, and evaluations that modified the initial design. The first app version was developed from the requirements specified by the project team (members of the consensus group). The app was then tested internally by project members. Due to the size of the project team (with more than 20 hospitals or research centers involved), multiple comments and suggestions were received, some of them conflicting. Table 1 summarizes the main conflicting comments for the app design and the final decisions that were made. This process step resulted in a second version of the app. This version was evaluated by the Andalusian Agency for Healthcare Quality. The accreditation agency proposed several improvements that required an app redesign.

This new version of the app was finally granted the AppSaludable (HealthyApp) Quality Seal on April 21, 2020 [21] (see Figure 3 [15]).

XSL•FO
RenderX

Table 1. Main conflicting suggestions during app development process.

Topic	Option 1	Option 2	Decision	Reason
Hybrid app vs native app	Hybrid to minimize developing effort	Native to improve performance in each platform (Android/iOS)	Native	Voted decision among developers
User registration	Yes to ensure reaching the correct target audience and gather statistical data	No to reinforce users' confidence in privacy	No	Voted decision
Interface design	Colorful and optimistic to show positiveness	Neutral to show respect during a difficult situation for health professionals	Colorful and optimistic	Voted decision
External, nonproprietary documents and videos	Include them (citing sources) to offer as much information and resources as possible	Include only proprietary material to improve uniformity of style and content.	Only proprietary material included	Voted decision

Figure 3. The evaluation process to obtain the AppSaludable Quality Seal.

Before being published in the Google and Apple stores, the app was reviewed by both companies. During the COVID-19 pandemic, both Apple and Google increased the depth of their reviews of all apps related to COVID-19. The result was a long review process with both companies, which also required small app redesigns. All redesigns were made for both platforms to keep Android and iOS versions as similar as possible. The app first appeared in Google's Play Store on April 13, 2020, and in Apple's App Store on May 11, 2020. As of May 31, 2020, the latest Android version is 1.8 and the latest iOS version is 1.1. Although version numbers are different, functionalities, contents, and aesthetics are the same in both apps. Figure 4 shows the complete development process of the mobile app.

The app consisted of a short presentation, which introduces the app's purpose and provides access to an app information section. Once the introduction is complete, the app is structured into three main modules. The first module, "Advices and recommendations," shows a list of resources that can be useful for HCWs caring for patients infected with COVID-19. These resources include posters, documents, and videos. The app was designed so that the resources were updated whenever a new version was available on the server, but at the same time, they could be accessed offline to account for possible situations where access to the internet may be limited. The second module, "Self-assessment on the ability to cope the COVID-19 crisis," presented a 10-question test that tried to assess whether the HCWs were overwhelmed by the situation they were experiencing (resource 9 of the website). Based on the test results, the app proposed a series of guidelines or recommendations, which included in-app advice and web resources.

This tool underwent a process of cross-cultural and linguistic adaptation and is being currently used in studies in Argentina, Brazil, Colombia, Chile, Ecuador, and Spain, assessing the acute stress of the health care workforce during the outbreak. The third module, "Visit website," redirected the users to the project website, where all the resources and information of the project were available. Figure 5 shows the app structure.

XSL•FO
RenderX

Figure 4. App design, test, evaluation, and review process.

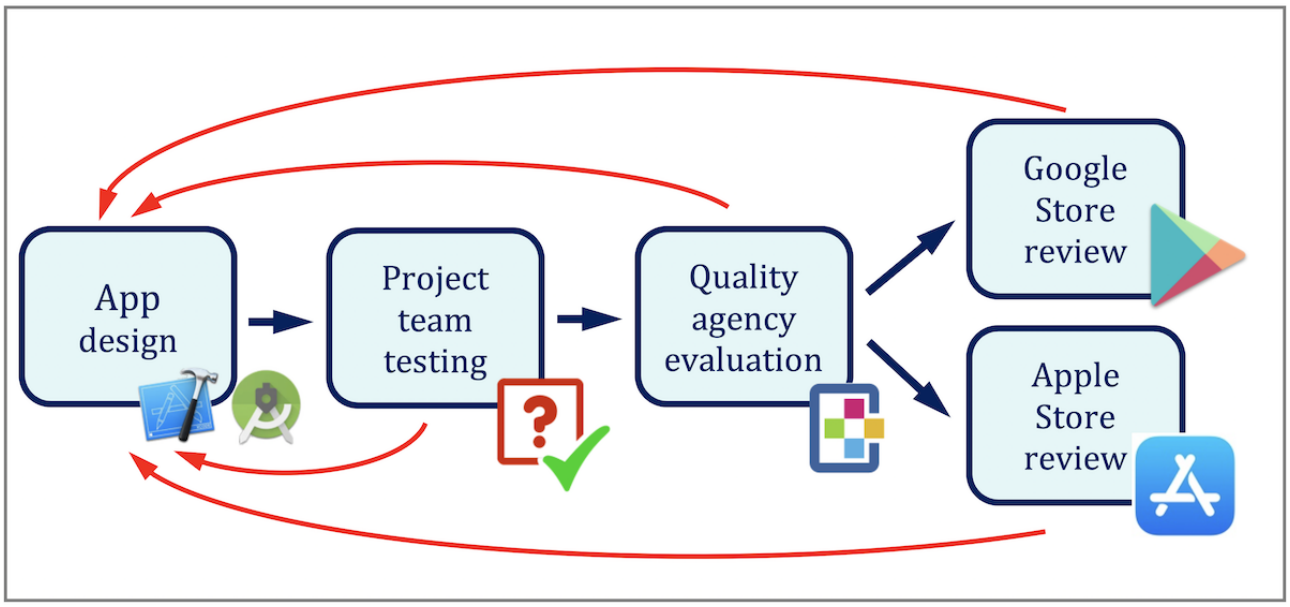
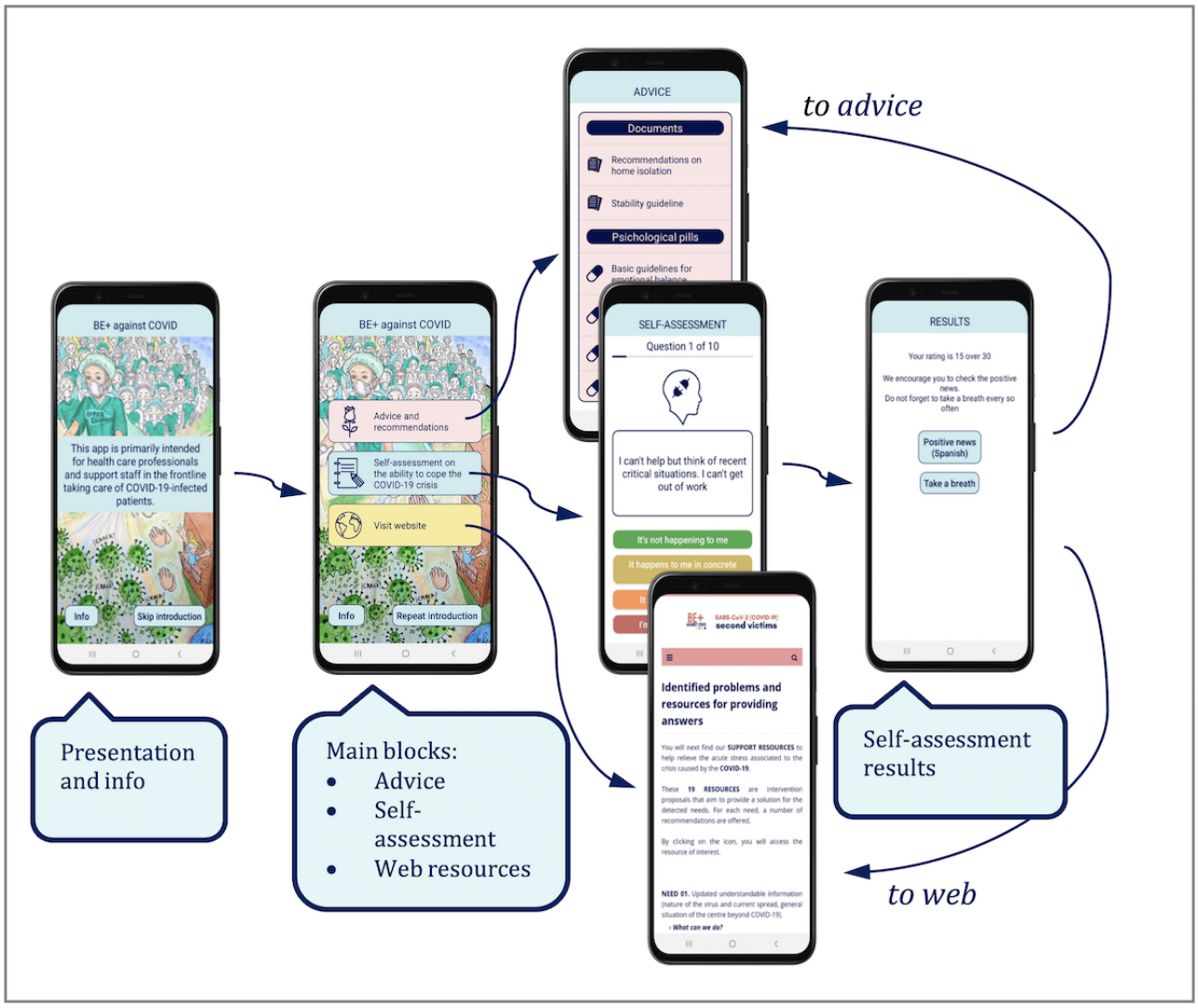


Figure 5. App structure explained by screenshots.



XSL•FO
RenderX

Diffusion Through Social Networks

The *Be + against COVID* platform was made known to the HCWs directly through the members of the research group (see Multimedia Appendix 3), their organizations, and through the main social networks of the members and organizations involved (mainly through Facebook and Twitter). The platform's twitter account (@second_victims) was used to further disseminate the contents of the web and the various updates and publications of the app.

No general diffusion means were used to ensure that the web and app only reached our target audience. Therefore, except for residual cases, the majority of web visits and app downloads, which will be detailed in the next sections, came from health care professionals.

Web Analytics

Spanish Website Stats

From March 25, 2020 (launch of Spanish version website), to May 31, 2020, more than 7500 users visited the website in almost 70,000 different sessions. Of the 68,913 sessions, the website was accessed mostly via mobile phones with 55.8% (n=38,453) of the sessions, followed by 41.9% (n=28,875) via desktop and 2.3% (n=1585) via tablet. Users from 64 different countries accessed the website but most of the visits (n=50,058, 72.6%) were made from Spain (Table 2).

Table 2. List of the countries that visited the Spanish version of the website most during the most critical phase of the SARS-CoV-2 pandemic (from March 25 to May 31, 2020).

Country	Users, n	Visits, n
Spain	5612	50,058
Argentina	765	7312
Chile	284	2666
United States	214	930
Colombia	141	1499
Mexico	96	1399
Brazil	91	1262
Others	493	3787

Table 3 shows the traffic distribution to the website in the same period, and the most viewed pages of the 68,913 sessions were the home page, with 43.2% (n=29,763) of the sessions, and resources, with 26.9% (n=18,516) of the sessions. The most viewed resource was the self-report questionnaire (resource 9; see Table 4).

Table 3. Traffic distribution to the Spanish version of the website (percentage of visits to webpages from March 25 to May 31, 2020).

Category	Web contents	Visits (N=68,913), n (%)
Home	Introduction, purpose, methods, project team, and references	29,763 (43.2)
Resources	Resources in short and the 19 detailed resources	18,516 (26.9)
Other	Download, Documents, Videos, Activities, HCWs[a] agreements, Tell us your experience, etc	7003 (10.2)
App	App information, manual, and links to the official stores	5695 (8.3)
Problems	Pressing situations, human factors, environmental stressors, fear or panic reactions, organization, human resources, material, and decision making	4067 (5.9)
Positive news	Posts on positive news related to the SARS-CoV-2 pandemic	3869 (5.6)

[a]HCW: health care worker.

XSL•FO
RenderX

Table 4. Visits to resources pages (Spanish version of the website) from March 25 to May 31, 2020.

Resources' contents	Visits (N=18,516), n (%)
All resources in short	3062 (16.5)
R[a]1 Achievements	766 (4.1)
R2 Informative videos	683 (3.7)
R3 Daily News	528 (2.9)
R4 homogenous messages	376 (2.0)
R5 Contract employees	304 (1.6)
R6 Fake News	248 (1.3)
R7 Briefings	624 (3.4)
R8 Long distance trainers	399 (2.2)
R9 Self-report	6468 (34.9)
R10 Emotional coping	1112 (6.0)
R11 Hotline support	459 (2.5)
R12 Recovery Time	736 (4.0)
R13 Defusing	785 (4.2)
R14 Referral to a professional	327 (1.8)
R15 Resting areas	413 (2.2)
R16 Contact during confinement	248 (1.3)
R17 Responsible leadership	254 (1.4)
R18 Informative leadership	241 (1.3)
R19 Post crisis	485 (2.6)

[a]R: resource.

App Stats

Figure 6 shows the number of daily app downloads from Google Play (Android version) and the App Store (iOS version). The Android version was first published on Google Play on April 1, 2020, while the iOS version was first published on May 11, 2020. There were 472 total app downloads during this period (Android: n=361; iOS: n=111). App downloads from our server, which started on March 28, 2020 (before the publication in the official app stores), were not included in these figures.

Regarding the self-assessment test, a total of 388 tests were completed from the app as of May 31, 2020. This number increased according to the intensity of the outbreak and particularly in territories where incidence and mortality of COVID-19 was higher.

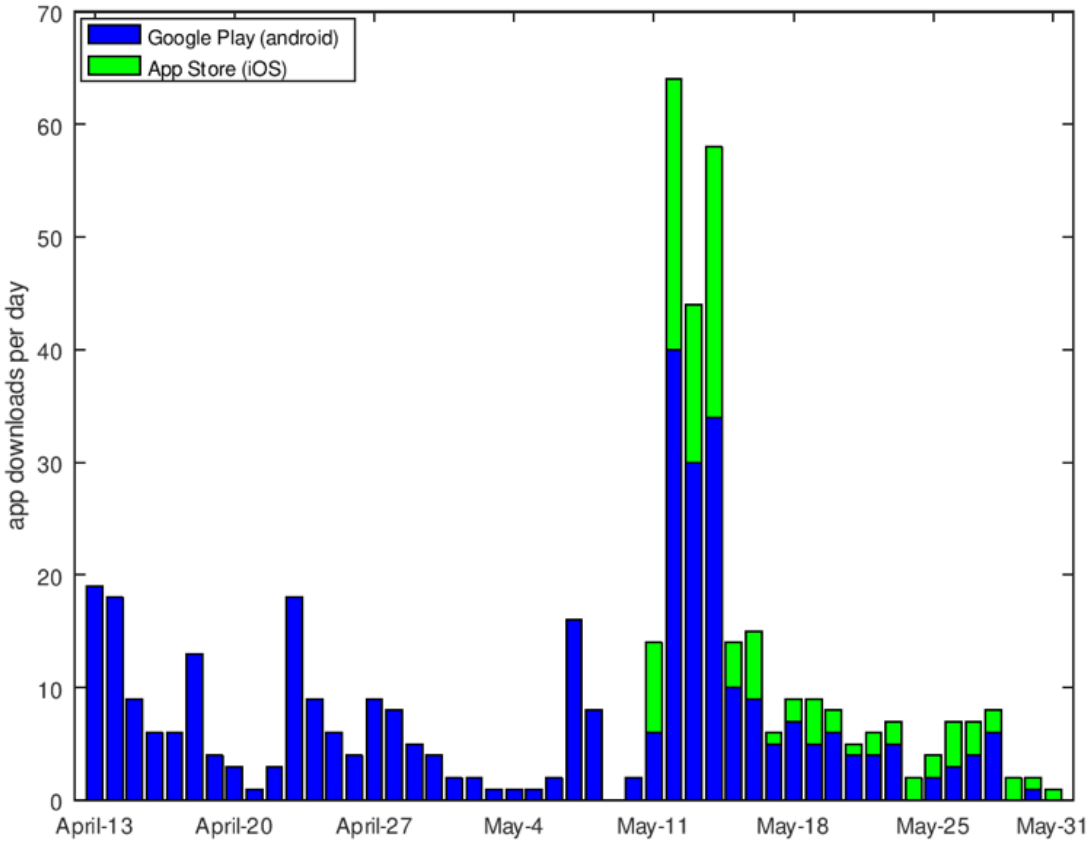
A simple statistical analysis was carried out using the data gathered from app users who took the self-assessment test (n=388). After showing the test results, the app offers the users different choices depending on the results. Most users (n=204, 52.5%) opted for browsing the list of psychological pills and reading about one of these pills, followed by 21.2% (n=82) of users who decided to have a look at the positive news and 18.6% (n=72) of users who opted for reading instructions on how to take a breath. The remaining 7.7% (n=30) of users chose to access the resources for overcoming acute stress reactions.

A more in-depth analysis showed the user engagement capability of the app. Using device identifiers to detect repeated user accesses to the app self-assessment test, we found that most of the 388 users (n=246, 63.4%) carried out the test only once, but 36.6% (n=142) of users took the survey more than once on different dates. More in detail, 14.7% (n=57) of the users took the test twice, 11.1% (n=43) three times, 2.8% (n=11) four times, and 8.0% (n=31) five or more times. Concerning the time spent from the first and last self-assessment carried out by the same user (long-term engagement), 2.1% (n=8) of the users took the test on different dates separated by at least 2 months, 6.3% (n=24) at least 1 month, 13.4% (n=51) at least 15 days, and 16.2% (n=63) at least 1 week.

Given that the app was published just a few weeks ago, the number of results is still small; however, the data collected offers interesting information on user behavior.

XSL•FO
RenderX

Figure 6. Daily app downloads from official app stores.



Discussion

Principal Results

This paper shows the development and the start-up of a project that arose from the urgent need to prevent and address the pressing situation that HCWs were and are facing in the world. Since Spain has been one of the first and most affected countries by the COVID-19 outbreak worldwide, this project can serve as an experience for others. The *Be + against COVID* platform (website and mobile app) was developed in three language versions to guide on possible actions that can be put into practice in health care centers to help confront the emotional impact on frontline HCWs caused by the current pandemic. Since there was no background about similar initiatives, the experience of China, previously published initiatives to facilitate the recovery of the second victim, and the interventions already implemented in some hospitals in Spain were considered to plan prevention and approach actions. During the first 2 months from the launch of the Spanish language version of the website, users from different countries accessed the website, but visitors were mostly from Spain. The most interesting and useful part of the website for HCWs was the resources section composed by intervention proposals that aim to provide a solution for the detected needs, and the "self-report" was the most visited resource. Currently, the Spanish language version of the website is being updated with content related to the recovery phase (postcrisis).

These resources were designed to help address the emotional consequences of caring for patients with COVID-19 in the context of the outbreak. This situation was characterized by a lack of foresight, demand for an immediate response from health professionals, and a shortage of resources. The focus was on providing resources through this platform for individual use.

Professionals from different disciplines (psychologists, psychiatrists, occupational health professionals, clinicians, and nurses) were involved in the design and redesign of these resources to support frontline health professionals in the care of patients with COVID-19, gathering experience in many hospitals and health care centers. They have assessed the adequacy of resources to the needs experienced by the professionals and have suggested improvements. In addition, Latin American professionals have been involved in the transcultural and linguistic adaptation of the resources and the self-report scale. Given that the expansion of the outbreak has been different according to each territory, the proposal of resources to face acute stress must be wide, varied, and flexible. The data confirm that acute stress increases in those territories with the greatest number of cases and highest mortality. In this case, the acceptability of the proposal is greatest [22]. In other countries, some hospitals provide psychological assistance hotlines to HCWs and carry out prevention programs. To offer a pool of recommendations and support resources through a free website is a different and original initiative. To the best of our knowledge, there is no similar experience in Europe or Latin

XSL•FO
RenderX

America. We consider that this paper may be useful to develop a web-based tool with a similar objective or in a similar context. The Brazilian research group Enfermagem da Escola Paulista de Enfermagem da Universidade Federal de São Paulo (EPE-UNIFESP) translated our resources into Portuguese, and they were incorporated into its institutional website [23].

Limitations

Given the current situation, conditions may differ across time and context, and therefore, our proposal of interventions is limited to the time and context of this study. These proposed actions can change due to the urgency and demand of the situation. Additionally, the response of HCWs to the COVID-19 crisis might be different between countries, and therefore, the problem situations and resources may be applicable in health care systems of countries with similar sociocultural characteristics. The urgency of the situation did not allow us to analyze the effectiveness of the proposed interventions. Future research should investigate it.

Comparison With Prior Work

Psychological assistance has been provided through telephone-, internet-, and application-based interventions to the health care community by some local or national institutions in response to the COVID-19 outbreak [9]. Several international organizations such as the WHO, the United Nations, and the International Red Cross Society have published recommendations that team leaders can consider to reduce stress and the psychological distress of frontline HCWs. Social support, clear communication, flexible working hours, and psychological help are some of these recommendations [24]. The State Council of China developed a psychological intervention plan, which consists of setting up nationwide psychological assistance hotlines to help during the outbreak [11,25]. However, some workers did not recognize any problems and refused any psychological help [11]. The Canadian Medical Association also offers some resources for supporting HCWs that focus on psychological assistance provided by volunteer psychologists [26]. On the other hand, Blake et al [27] developed and evaluated a digital learning package that includes evidence-based guidance, support, and signposting regarding psychological well-being for frontline HCWs from the United Kingdom. This online support package is free and comprised of 88 slides, and its content is similar to that of the digital platform developed by our team. Similar to our digital platform, this electronic package describes the actions that team leaders can take to support HCWs during the COVID-19 outbreak and offers some support resources (infographic and video formats) to protect psychological well-being [27]. The evaluation showed that it is useful and appropriate for any UK health care professional. In contrast to our website, which was published in three different language versions, Blake et al [27] seemed to be focused on workers from the United Kingdom to develop this package. Moreover, nowadays, a mobile app can be more practical for some users. Nevertheless, the fidelity, engagement, usability, and practicality of our digital platform should be evaluated in future research.

Extrapolation to Other Contexts

Although there are differences between the health systems of the different countries and it is evident that the spread of the SARS-CoV-2 pandemic is different in each continent and country, most of the problems identified, and for which these resources have been devised, are repeated in the literature [9,28-30], and it is verified in the experience described by health professionals in Brazil, Peru, Argentina, Colombia, Mexico, Ecuador, Italy, Portugal, and Central Europe.

The extrapolation of resources to mitigate acute stress among frontline health professionals in the care of patients with COVID-19, therefore, could be carried out in both public and private centers.

The development and redefinition of resources to support the health care workforce and address new outbreaks must take into account differences in organizational culture [31-33], psychological safety [34], and availability of financial and skilled human resources among countries.

Although similarities have been observed in the sources and acute stress responses of these professionals, there is a need to adapt resources to the context as, for example, has been done in Brazil, adapting the resources originally designed for the Spanish National Healthcare System. Future research studies may consider these differences when assessing the applicability of the resources designed.

Upcoming Developments

Preparing for a spike in new COVID-19 cases from health care organizations and professionals, particularly on the frontline of care for patients with COVID-19, should be prioritized by health planners and managers, academics, and health care professionals. Systems have more resources and better organization and information to reduce clinical uncertainty, but templates need to be recovered physically and emotionally. The new developments of this platform are incorporating proposals to strengthen the resilience of organizations and professionals.

Conclusions

A digital platform (website and app) was developed and launched to provide support resources that might prevent and approach the emotional impact of HCWs and other support staff due to the current COVID-19 outbreak.

Acknowledgments

We would like to express our gratitude to the following Spanish institutions for their help to spread the *Be + against COVID* platform to the HCWs: Atención Primaria Calatayud, Atención Primaria Zaragoza, the Foundation for the Promotion of Health and Biomedical Research of Valencia Region–Generalitat Valenciana, Grup Sagessa, Hospital General Universitari d'Elda, Hospital Reina Sofía (Murcia), Hospital Universitario 12 de Octubre (Madrid), Hospital Universitario Clínico San Cecilio

XSL•FO
RenderX

(Granada), Hospital Universitario Fundación Alcorcón, Hospital Virgen del Castillo (Yecla), Miguel Hernández University, Osansubidea, Servicio de Medicina Preventiva y Salud Pública – Hospital Universitario Ramón y Cajal Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS; Madrid), and Servicio de Salud de Castilla La Mancha (Toledo) y Sociedad Española de Calidad Asistencial. In addition, we would like to thank IRYCIS for their essential help in publishing the app in the App store. We also want to thank the researchers of the EPE-UNIFESP group for sharing with us the translation of the platform into Brazilian Portuguese.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Identified problem situations of health care workers during the COVID-19 outbreak, generated needs, and available resources for an appropriate response.
[DOC File , 243 KB - mhealth_v8i10e21692_app1.doc]

Multimedia Appendix 2
Infographics of some of the resources from the Be + against COVID platform.
[PDF File (Adobe PDF File), 1337 KB - mhealth_v8i10e21692_app2.pdf]

Multimedia Appendix 3
Project team.
[DOCX File , 14 KB - mhealth_v8i10e21692_app3.docx]

References

1. WHO announces COVID-19 outbreak a pandemic. World Health Organization/Europe. 2020 Mar 12. URL: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic [accessed 2020-04-13]
2. Actualización nº 122. Enfermedad por el coronavirus (COVID-19). Ministerio de Sanidad. 2020 May 31. URL: https://www.mscbs.gob.es/en/profesionales/saludPublica/ccayes/alertasActual/nCov-China/documentos/Actualizacion_122_COVID-19.pdf [accessed 2020-06-01]
3. Rapid risk assessment: coronavirus disease 2019 (COVID-19) in the EU/EEA and the UK– ninth update. European Centre for Disease Prevention and Control. 2020 Apr 23. URL: https://www.ecdc.europa.eu/en/publications-data/rapid-risk-assessment-coronavirus-disease-2019-covid-19-pandemic-ninth-update
4. Coronavirus disease (COVID-2019) situation reports. World Health Organization. 2019. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/ [accessed 2020-04-13]
5. Liu N, Zhang F, Wei C, Jia Y, Shang Z, Sun L, et al. Prevalence and predictors of PTSS during COVID-19 outbreak in China hardest-hit areas: gender differences matter. Psychiatry Res 2020 May;287:112921 [FREE Full text] [doi: 10.1016/j.psychres.2020.112921] [Medline: 32240896]
6. Lee SM, Kang WS, Cho A, Kim T, Park JK. Psychological impact of the 2015 MERS outbreak on hospital workers and quarantined hemodialysis patients. Compr Psychiatry 2018 Nov;87:123-127 [FREE Full text] [doi: 10.1016/j.comppsych.2018.10.003] [Medline: 30343247]
7. Styra R, Hawryluck L, Robinson S, Kasapinovic S, Fones C, Gold WL. Impact on health care workers employed in high-risk areas during the Toronto SARS outbreak. J Psychosom Res 2008 Feb;64(2):177-183 [FREE Full text] [doi: 10.1016/j.jpsychores.2007.07.015] [Medline: 18222131]
8. Cai H, Tu B, Ma J, Chen L, Fu L, Jiang Y, et al. Psychological impact and coping strategies of frontline medical staff in Hunan between January and March 2020 during the outbreak of coronavirus disease 2019 (COVID-19) in Hubei, China. Med Sci Monit 2020 Apr 15;26:e924171 [FREE Full text] [doi: 10.12659/MSM.924171] [Medline: 32291383]
9. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. JAMA Netw Open 2020 Mar 02;3(3):e203976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]
10. Bohlken J, Schömig F, Lemke MR, Pumberger M, Riedel-Heller SG. [COVID-19 pandemic: stress experience of healthcare workers - a short current review]. Psychiatr Prax 2020 May;47(4):190-197 [FREE Full text] [doi: 10.1055/a-1159-5551] [Medline: 32340048]
11. Chen Q, Liang M, Li Y, Guo J, Fei D, Wang L, et al. Mental health care for medical staff in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):e15-e16 [FREE Full text] [doi: 10.1016/S2215-0366(20)30078-X] [Medline: 32085839]

XSL•FO
RenderX

12. Xiang Y, Yang Y, Li W, Zhang L, Zhang Q, Cheung T, et al. Timely mental health care for the 2019 novel coronavirus outbreak is urgently needed. Lancet Psychiatry 2020 Mar;7(3):228-229 [FREE Full text] [doi: 10.1016/S2215-0366(20)30046-8] [Medline: 32032543]
13. Blogger. URL: https://www.blogger.com/about/ [accessed 2020-05-22]
14. Mobile operating system market share worldwide. StatCounter Global Stats. URL: https://gs.statcounter.com/os-market-share/mobile/
15. Safety and Quality Strategy in Mobile Health Apps. URL: http://www.calidadappsalud.com/en/ [accessed 2020-05-22]
16. Mira J, Torijano M, Abad C, Sánchez A, Lahera M, Cobos A. Addressing acute stress among professionals caring for COVID-19 patientsxperience in Spain (March-April 2020). Bull World Health Organ. Forthcoming 2020.
17. Mira JJ, Cobos A, Martínez-García O, Bueno-Domínguez MJ, Astier-Peña MP, Pérez-Pérez P, et al. The Acute Stress Scale in healthcare professionals caring for patients with COVID-19. Validation study. Int J Integr Care 2020. [doi: 10.21203/rs.3.rs-39710/v1]
18. Segundas víctimas del SARS-CoV-2 (COVID-19). URL: https://segundasvictimascovid19.umh.es/p/inicio.html [accessed 2020-05-22]
19. SARS-CoV-2 (COVID-19) Second Victims. URL: https://secondvictimscovid19.umh.es/p/home.html [accessed 2020-05-22]
20. Segundas vitimas SARS-CoV-2 (COVID-19). URL: https://segundasvitimascovid19.umh.es/p/apresentacao.html [accessed 2020-05-22]
21. Ser+ contra COVID. Distintivo AppSaludable. 2020 Apr 21. URL: http://www.calidadappsalud.com/distintivo/info?app=SerMasContraCovid [accessed 2020-05-22]
22. Li X, Yu H, Bian G, Hu Z, Liu X, Zhou Q, et al. Prevalence, risk factors, and clinical correlates of insomnia in volunteer and at home medical staff during the COVID-19. Brain Behav Immun 2020 Jul;87:140-141 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.008] [Medline: 32380272]
23. Projetos de Extensão DASSE. Departamento de Administração e serviços de saude e enfermagem. URL: https://sp.unifesp.br/epe/dasse/projeto-de-extensao [accessed 2020-06-01]
24. Petzold MB, Plag J, Ströhle A. [Dealing with psychological distress by healthcare professionals during the COVID-19 pandemia]. Nervenarzt 2020 May;91(5):417-421 [FREE Full text] [doi: 10.1007/s00115-020-00905-0] [Medline: 32221635]
25. A notification to set up nationwide psychological assistance hotlines against the 2019-nCoV outbreak. The State Council of China. 2020 Feb 02. URL: http://www.gov.cn/xinwen/2020-02/02/content_5473937.htm [accessed 2020-06-03]
26. Wu PE, Styra R, Gold WL. Mitigating the psychological effects of COVID-19 on health care workers. CMAJ 2020 Apr 27;192(17):E459-E460. [doi: 10.1503/cmaj.200519] [Medline: 32295761]
27. Blake H, Bermingham F, Johnson G, Tabner A. Mitigating the psychological impact of COVID-19 on healthcare workers: a digital learning package. Int J Environ Res Public Health 2020 Apr 26;17(9) [FREE Full text] [doi: 10.3390/ijerph17092997] [Medline: 32357424]
28. Tabah A, Ramanan M, Laupland KB, Buetti N, Cortegiani A, Mellinghoff J, PPE-SAFE contributors. Personal protective equipment and intensive care unit healthcare worker safety in the COVID-19 era (PPE-SAFE): an international survey. J Crit Care 2020 Oct;59:70-75 [FREE Full text] [doi: 10.1016/j.jcrc.2020.06.005] [Medline: 32570052]
29. Kisely S, Warren N, McMahon L, Dalais C, Henry I, Siskind D. Occurrence, prevention, and management of the psychological effects of emerging virus outbreaks on healthcare workers: rapid review and meta-analysis. BMJ 2020 May 05;369:m1642. [doi: 10.1136/bmj.m1642] [Medline: 32371466]
30. Shanafelt T, Ripp J, Trockel M. Understanding and addressing sources of anxiety among health care professionals during the COVID-19 pandemic. JAMA 2020 Jun 02;323(21):2133-2134. [doi: 10.1001/jama.2020.5893] [Medline: 32259193]
31. Tamburri LM. Creating healthy work environments for second victims of adverse events. AACN Adv Crit Care 2017;28(4):366-374. [doi: 10.4037/aacnacc2017996] [Medline: 29212644]
32. Quillivan RR, Burlison JD, Browne EK, Scott SD, Hoffman JM. Patient safety culture and the second victim phenomenon: connecting culture to staff distress in nurses. Jt Comm J Qual Patient Saf 2016 Aug;42(8):377-386 [FREE Full text] [doi: 10.1016/s1553-7250(16)42053-2] [Medline: 27456420]
33. White RM, Delacroix R. Second victim phenomenon: is 'just culture' a reality? An integrative review. Appl Nurs Res 2020 Jun 19:151319. [doi: 10.1016/j.apnr.2020.151319] [Medline: 32868148]
34. Zhao F, Ahmed F, Faraz NA. Caring for the caregiver during COVID-19 outbreak: does inclusive leadership improve psychological safety and curb psychological distress? A cross-sectional study. Int J Nurs Stud 2020 Jul 30;110:103725 [FREE Full text] [doi: 10.1016/j.ijnurstu.2020.103725] [Medline: 32810720]

Abbreviations

EPE-UNIFESP: Enfermagem da Escola Paulista de Enfermagem da Universidade Federal de São Paulo
HCW: health care worker
IRYCIS: Instituto Ramón y Cajal de Investigación Sanitaria
SECA: Sociedad Española de Calidad Asistencial
WHO: World Health Organization

XSL•FO
RenderX

Edited by G Eysenbach, L Buis; submitted 22.06.20; peer-reviewed by B Knezevic, D Chao, A Teles, M Rauws; comments to author 15.07.20; revised version received 30.07.20; accepted 09.09.20; published 05.10.20.

Please cite as:
Mira JJ, Vicente MA, Lopez-Pineda A, Carrillo I, Guilabert M, Fernández C, Pérez-Jover V, Martin Delgado J, Pérez-Pérez P, Cobos Vargas A, Astier-Peña MP, Martínez-García OB, Marco-Gómez B, Abad Bouzán C
Preventing and Addressing the Stress Reactions of Health Care Workers Caring for Patients With COVID-19: Development of a Digital Platform (Be + Against COVID)
JMIR Mhealth Uhealth 2020;8(10):e21692
URL: https://mhealth.jmir.org/2020/10/e21692
doi:10.2196/21692
PMID:32936769

©José Joaquín Mira, María Asunción Vicente, Adriana Lopez-Pineda, Irene Carrillo, Mercedes Guilabert, César Fernández, Virtudes Pérez-Jover, Jimmy Martin Delgado, Pastora Pérez-Pérez, Angel Cobos Vargas, María Pilar Astier-Peña, Olga Beatriz Martínez-García, Bárbara Marco-Gómez, Cristina Abad Bouzán. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 05.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

COVID-19 Contact Tracing Apps: Predicted Uptake in the Netherlands Based on a Discrete Choice Experiment

Marcel Jonker[1,2], PhD; Esther de Bekker-Grob[1,2], PhD; Jorien Veldwijk[1,2], PhD; Lucas Goossens[1,2], PhD; Sterre Bour[1], MSc; Maureen Rutten-Van Mölken[1,2,3], PhD

[1]Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, Netherlands

[2]Erasmus Choice Modelling Centre, Erasmus University Rotterdam, Rotterdam, Netherlands

[3]Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, Netherlands

Corresponding Author:
Marcel Jonker, PhD
Erasmus School of Health Policy & Management
Erasmus University Rotterdam
Rotterdam
Netherlands
Phone: 31 10 408 8555
Email: marcel@mfjonker.com

Abstract

Background: Smartphone-based contact tracing apps can contribute to reducing COVID-19 transmission rates and thereby support countries emerging from lockdowns as restrictions are gradually eased.

Objective: The primary objective of our study is to determine the potential uptake of a contact tracing app in the Dutch population, depending on the characteristics of the app.

Methods: A discrete choice experiment was conducted in a nationally representative sample of 900 Dutch respondents. Simulated maximum likelihood methods were used to estimate population average and individual-level preferences using a mixed logit model specification. Individual-level uptake probabilities were calculated based on the individual-level preference estimates and subsequently aggregated into the sample as well as subgroup-specific contact tracing app adoption rates.

Results: The predicted app adoption rates ranged from 59.3% to 65.7% for the worst and best possible contact tracing app, respectively. The most realistic contact tracing app had a predicted adoption of 64.1%. The predicted adoption rates strongly varied by age group. For example, the adoption rates of the most realistic app ranged from 45.6% to 79.4% for people in the oldest and youngest age groups (ie, ≥75 years vs 15-34 years), respectively. Educational attainment, the presence of serious underlying health conditions, and the respondents' stance on COVID-19 infection risks were also correlated with the predicted adoption rates but to a lesser extent.

Conclusions: A secure and privacy-respecting contact tracing app with the most realistic characteristics can obtain an adoption rate as high as 64% in the Netherlands. This exceeds the target uptake of 60% that has been formulated by the Dutch government. The main challenge will be to increase the uptake among older adults, who are least inclined to install and use a COVID-19 contact tracing app.

(JMIR Mhealth Uhealth 2020;8(10):e20741) doi:10.2196/20741

KEYWORDS

COVID-19; discrete choice experiment; contact tracing; participatory epidemiology; participatory surveillance; app; uptake; prediction; smartphone; transmission; privacy; mobile phone

Introduction

The COVID-19 pandemic has formed an unprecedented public health, societal, and economic crisis. Given that no vaccine is available yet and that treatment options are limited, prevention is crucial. In an effort to stop the spread of the virus, societies have been locked down to varying degrees with social distancing; stay-at-home measures; and closures of schools, universities, and business. These policies are socially painful and economically costly.

Targeting the quarantine measures, whether enforced or voluntary, could ease the social and economic impact of these

XSL•FO
RenderX

policies. Such measures require that people who are infected be quickly identified and isolated, and that their recent contacts are quarantined for the duration of the disease incubation period. Essentially, when contacts of people who are infected can be traced and quarantined at a sufficient speed, unaffected people can continue to live their lives in a more or less normal fashion. According to the World Health Organization, "when systematically applied, contact tracing will break the chains of transmission of an infectious disease and is thus an essential public health tool" [1]. There is evidence that contact tracing was effective in previous pandemics, and several model studies point to benefits in the COVID-19 crisis as well [2,3].

The main disadvantage of contact tracing is that it is labor intensive and time-consuming, and can only be effective when conducted fast. When people develop symptoms that indicate a COVID-19 infection and subsequently test positive for COVID-19, a high proportion of their contacts must be warned and quarantined quickly to avoid further infections. Manual contact tracing by public health authorities may not be able to achieve this [4]. The problem could be exacerbated if public health authorities are unable to recruit and train sufficient staff for the task.

For these reasons, the use of digital methods (ie, smartphone-based contact tracing apps) have been proposed to facilitate and accelerate contact tracing. The basic idea is that app users who have been in close proximity to someone who turned out to have been infected with COVID-19 receive a warning and are asked to self-quarantine. Such an app could theoretically replace a week's work of manual contact tracing (per infected person) with an almost instantaneous notification once an infection has been ascertained [5].

COVID-19 contact tracing apps can only be successful in limiting the spread of the virus if a sufficient number of people are willing to download and use them. This becomes increasingly important especially when social distancing measures are relaxed and manual tracing is unable to act comprehensively and fast [5]. Based on modelling studies, the Dutch government has formulated the aim that 60% of the population should use the app [6]. It is not known whether this level of uptake can be achieved. In Singapore, it took 1 month before 20% of the population had started using the app [7], whereas in the Isle of Wight more than 40% had downloaded the contact tracing app within 10 days [8].

Dutch authorities have explicitly stated that the app will only be used for contact tracing and not to monitor or enforce self-quarantine, or to provide access to public places. However, the exact specifications of the Dutch contact tracing app have not been established yet, but they may have an impact on people's willingness to use it. Therefore, the main aim of this study is to estimate the future uptake of a smartphone-based contact tracing app in the Dutch population and the extent that this could be optimized by changing the specifications of the app. The secondary aim is to describe differences in expected uptake between subgroups. Both aims are addressed using a discrete choice experiment (DCE).

Methods

In a DCE, preferences for a product such as a COVID-19 contact tracing app are established by decomposing the product into separate characteristics (referred to as *attributes*) and different specifications of these characteristics (referred to as *attribute levels*) [9,10]. For example, the attribute "financial incentive" comprises the levels "€0," "€5," and "€10" per month (€1 = US$1.18) (see Textbox 1). The relative importance of the attribute levels is then empirically established by asking respondents to make trade-offs in a series of choice tasks. Within each choice task there are two or more products to choose from, and respondents are repeatedly asked to indicate which option they prefer. Statistical regressions are subsequently used to derive numerical values for the relative attractiveness of the attributes and its levels, using methods that have a solid foundation in random utility theory [11].

The selected attributes and attribute levels in this DCE study (see Textbox 1) reflected the Dutch context, in the sense that the Dutch government has already issued the development of a COVID-19 contact tracing app to alert users when someone they were recently near becomes infected. Currently, the launch of the app is foreseen to support a further lifting of the lockdown. However, the launch is contingent upon the app meeting the European General Data Protection Regulation privacy and safety regulations (ie, from the outset) [12]. Consequently, attributes that describe different safety and privacy levels of the app were not included in the DCE and instead were described as being part of the context of the DCE (ie, held constant across all choice tasks). The latter accommodated additional focus on the type of warnings, testing, control over the communication of a positive test result, and the potential financial incentive for app users in the DCE.

XSL•FO
RenderX

Textbox 1. Included attributes and levels.

Group size ("App users can attend…")

1. Activities up to 3 people
2. Activities up to 10 people
3. Activities up to 30 people
4. Activities up to 100 people

Warning type

1. That you were close to a person who was infected in the last 2 weeks
2. At which date and time you were close to a person who was infected

Who is warned

1. Only you
2. You and the local health authorities (GGD), but only with your consent
3. You and automatically the local health authorities (GGD)

Testing after a warning

1. Only when someone has symptoms
2. Everyone will be tested

Who can upload test results

1. Only you
2. This is done automatically by the government and/or local health authorities (GGD)

Financial incentive (per month)

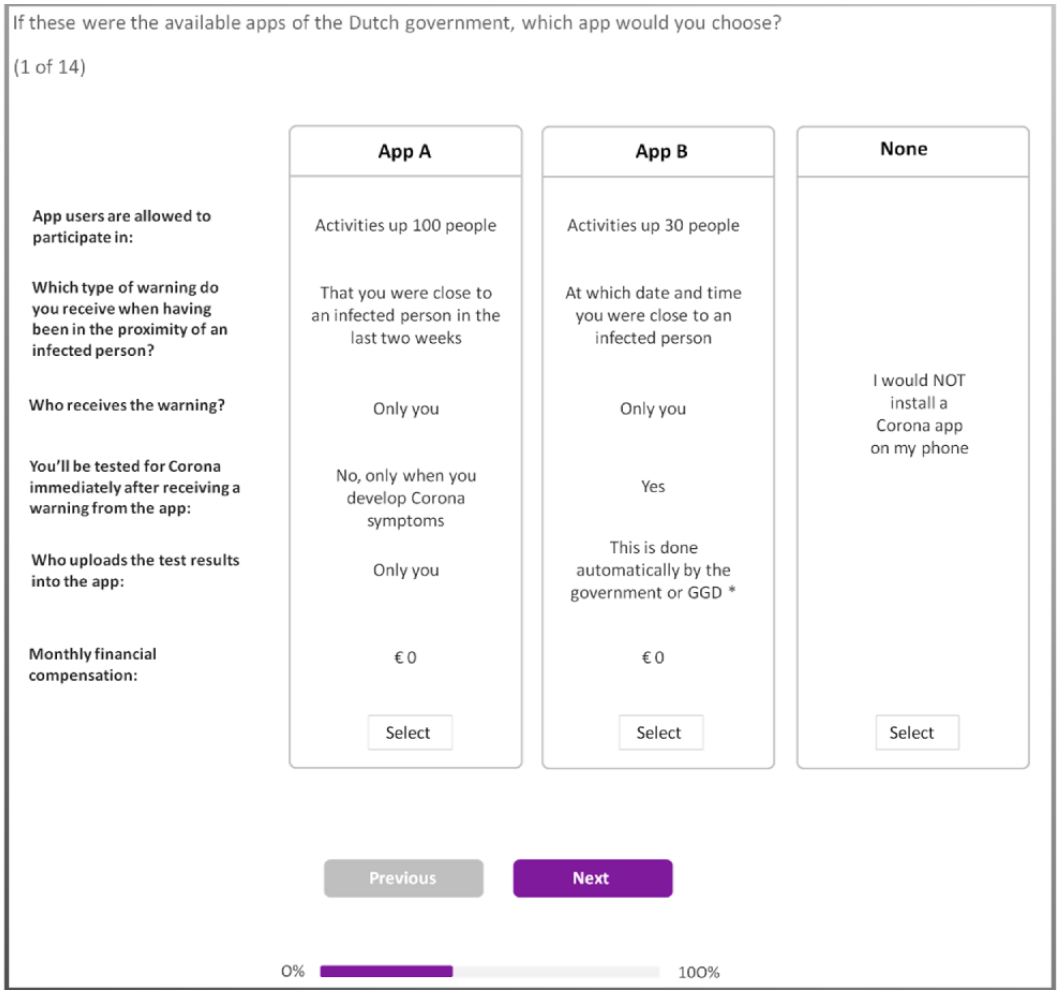
1. €0
2. €5
3. €10

Once the attributes and attribute levels were defined, an initial version of the survey instrument was created using Sawtooth Software (Sawtooth Software Inc) and administered in a Dutch online panel managed by Dynata, a commercial survey sample provider. The instrument included a DCE based on a nearly orthogonal design with 300 design versions and 13 choice tasks per version. Orthogonality minimizes the correlations between attribute levels in the choice tasks and ensures statistical identification of the preference parameters. To assess the stability of the respondents' preferences, the 14th choice task in the DCE design was a duplication of the fifth choice task [13,14]. Figure 1 provides a choice task example. As shown, respondents were able to choose between two apps and an opt out. Moreover, a small amount of attribute level overlap was used to reduce the cognitive burden of the survey and improve respondents' attribute attendance [15]. In addition to the DCE tasks, several background, warm-up, attitudinal, and survey evaluation questions were included in the survey.

The survey instrument was pilot tested using a sample of 238 online respondents. Based on the feedback from survey participants and an evaluation of the estimated preferences, the survey instrument was revised to improve understandability and to reduce the cognitive burden of the survey. The introduction and description of the attributes were enhanced, an additional warm-up question was added, supplementary debriefing questions were included, and the levels of one of the DCE attributes (ie, the group size) were revised to better reflect the observed nonlinear preference structure. This implied that the initial data collection became incompatible with the subsequent data collection and was excluded from the final analysis.

XSL•FO
RenderX

Figure 1. Example discrete choice task. Note: translated; original in Dutch. *GGD = local health authorities.



The revised and improved instrument (available upon request) was pilot-tested using a second pilot sample of 260 participants. Based on the feedback of the respondents and our evaluation of the estimated preferences, no further changes were required. Consequently, another 640 participants were obtained to achieve an overall survey sample of 900 respondents, which was sufficient based on formal sample size calculations as well as commonly used rules of thumb [16]. The overall sample of 900 respondents was designed to be nationally representative in terms of sex, age, and educational attainment of the Dutch general population 15 years and older. All data were collected in week 16 of 2020.

Once data collection was completed, the survey satisfaction and cognitive debriefing questions were summarized by averaging the 7-point Likert scores. The survey's dropout rates were directly observed, and completion timings were calculated as the cumulative time spent on the pages of the questionnaire, maximized at 5 minutes per page to correct for respondents taking a break in between survey questions. The stability of respondents' preferences was assessed based on the percentage of respondents with an identical choice in repeated choice task.

Descriptive statistics were used to describe respondent characteristics for the entire sample and between subgroups that were defined based on the observed choice behavior in the DCE. More specifically, a comparison was made between respondents who always chose the app, sometimes chose the app, and never chose the app and thus always choose the opt-out option in the DCE. We examined whether the choice behavior differed by sex, age group, highest level of education, general health, chronic conditions or reduced resistance, whether someone experienced COVID-19 symptoms, and whether respondents actively used health apps. We also compared the three subgroups in terms of attitude toward the six DCE evaluation questions, the maximum group size they identified to feel comfortable with, and nine general statements related to COVID-19 tracing apps based on the Health Belief Model [17,18] that were included in the survey instrument.

Simulated maximum likelihood methods were used to estimate population- and individual-level preferences using a mixed logit (MIXL) model [19]. Such a model uses the observed choices as the dependent variable and the characteristics (ie, attribute levels) of the COVID-19 contact tracing apps shown to respondents as explanatory variables. The estimations were conducted using Stata 15 (StataCorp) with the simulated maximum likelihood calculated using 2000 Halton draws to ascertain stable coefficients and with a full variance-covariance matrix estimation aimed at accommodating potential nonzero correlations between the random parameters.

Respondents who did not choose the COVID-19 contact tracing app in any of the 14 choice tasks were excluded from the MIXL estimations and assigned to a separate (latent) class. For these respondents, it was impossible to ascertain whether they would

XSL•FO
RenderX

theoretically be willing to consider installing a contact tracing app (which would merely imply a very positive opt-out parameter) or, alternatively, had lexicographic preferences and would never consider installing any contact tracing app. In the choice predictions these respondents were treated as an exogenous class with zero willingness to install and use the app. This avoided a spillover effect and upwards biased opt-out parameters of the respondents who did choose the contact tracing app in at least one of the choice tasks.

The subsequent uptake calculations were performed for three different contact tracing apps (ie, the best, worst, and most realistic app) and were based on the respondents' individual-level preference coefficients. The best and worst contact tracing apps were defined by the MIXL estimates. The most realistic app was defined based on all publicly available information at the time of publication. For each respondent, the predicted uptake probability was calculated using the standard logit rule, after which the predicted sample adoption rate was calculated as the mean of the individual-level uptake probabilities. This is more reliable than calculations directly based on the MIXL sample mean parameters because it takes respondent heterogeneity appropriately into account. For each of the apps, the adoption rate was calculated for the entire sample of 900 respondents and for several subsamples such as different educational backgrounds (low, medium, or high) and age groups (15-34, 35-54, 55-74, ≥75 years).

Results

Study Population

From the total 986 panel members who started the survey and were found eligible to participate (due to quota restrictions), 900 (91.3%) completed the questionnaire, resulting in 86 dropouts (8.7%). The resulting sample was representative for the Dutch population with respect to age, sex, and education level. Of the 900 respondents, 39% (n=351) were 55 years or older, 442 (49%) respondents were male, and one-third had a lower education level (Table 1). Approximately 70% of the respondents reported that they were in good health, and 625 (69%) respondents did not have a chronic disease or a compromised immune system. Almost 25% of the respondents reported that they experienced COVID-19 symptoms during the last 2 months, and 1.4% (n=14) of respondents tested positive for COVID-19. Almost all respondents owned and used a smartphone, smartwatch, or tablet (n=827, 91.9%), and 47.60% (n=428) of respondents already used health-related apps on their mobile device. The majority of the respondents indicated that the survey was (very) interesting (n=645, 72%) and (very) clear (n=738, 82%). There were 48 (5%) respondents that found the survey (very) unclear (Table 2).

XSL•FO
RenderX

Table 1. Respondents' sociodemographic characteristics for the total sample and stratified by respondents who always, sometimes, or never chose the COVID-19 app in the discrete choice experiment.[a]

Demographics	Total (N=900), n (%)	Always (n=460), n (%)	Sometimes (n=214), n (%)	Never (n=226), n (%)
Gender				
Male	442 (49.1)	215 (48.6)	114 (25.8)	113 (25.6)
Female	458 (50.9)	245 (53.5)	100 (21.8)	113 (24.7)
Age group (years)				
15-34	268 (29.8)	168 (62.7)	76 (28.4)	24 (9.0)
35-54	281 (31.2)	131 (46.6)	74 (26.3)	76 (27.0)
55-74	265 (29.4)	124 (46.8)	59 (22.3)	82 (30.9)
≥75	86 (9.6)	37 (43.0)	5 (5.8)	44 (51.2)
Education level				
Low	274 (30.4)	134 (48.9)	51 (18.6)	89 (32.5)
Medium	342 (38.0)	187 (54.7)	79 (23.1)	76 (22.2)
High	284 (31.5)	139 (48.9)	84 (29.6)	61 (21.5)
Geographical region				
Heavily impacted[b]	612 (68.0)	304 (49.7)	149 (24.3)	159 (26.0)
Mildly impacted	288 (32.0)	156 (54.2)	65 (22.6)	67 (23.3)
Self-perceived general health				
Good or very good	635 (70.6)	322 (50.7)	159 (25.0)	154 (24.3)
Fair	232 (25.8)	119 (51.3)	49 (21.1)	64 (27.6)
Bad or very bad	33 (3.7)	19 (57.6)	6 (18.2)	8 (24.2)
Health issues				
Lung disease	112 (12.4)	72 (64.3)	19 (17.0)	21 (18.8)
Heart disease	79 (8.8)	40 (50.6)	17 (21.5)	22 (27.8)
Diabetes	87 (9.7)	44 (50.6)	13 (14.9)	30 (34.5)
Kidney disease	11 (1.2)	7 (63.6)	2 (18.2)	2 (18.2)
Compromised immune system	61 (6.8)	38 (62.3)	13 (21.3)	10 (16.4)
Self-reported COVID-19 symptoms during last 2 months				
Yes	218 (24.2)	131 (60.1)	53 (24.3)	34 (15.6)
No	645 (71.7)	312 (48.4)	152 (23.6)	181 (28.1)
Tested for COVID-19 infection				
Yes, positive test	14 (1.6)	8 (57.1)	6 (42.9)	0 (0)
Yes, negative test	25 (2.8)	15 (60.0)	9 (36.0)	1 (4.0)
No	855 (95.0)	434 (50.8)	197 (23.0)	224 (26.2)
Owns and uses smartphone/smartwatch or tablet	827 (91.9)	446 (53.9)	203 (24.5)	178 (21.5)
Uses health apps on smartphone/smartwatch or tablet	428 (47.6)	272 (63.6)	100 (23.4)	56 (13.1)
Feels comfortable around				
Groups of 3 people	292 (32.4)	108 (37.0)	69 (23.6)	115 (39.4)
Groups of 10 people	312 (34.7)	196 (62.8)	66 (21.2)	50 (16.0)
Groups of 30 people	165 (18.3)	107 (64.8)	36 (21.8)	22 (13.3)
Groups of 100 people	79 (8.8)	34 (43.0)	31 (39.2)	14 (17.7)

XSL•FO
RenderX

Demographics	Total (N=900), n (%)	Always (n=460), n (%)	Sometimes (n=214), n (%)	Never (n=226), n (%)
Groups of 1000 people	52 (5.8)	15 (28.8)	12 (23.1)	25 (48.1)

[a]The percentages in column 2 add up to 100% vertically, whereas the percentages in column 3-5 add up to 100% horizontally.

[b]Heavily impacted regions are Noord-Brabant, Limburg, Zuid-Holland, Noord-Holland, and Gelderland.

XSL•FO
RenderX

Table 2. Respondents' attitude toward COVID-19 and evaluation of the survey for the total sample and stratified by respondents who sometimes, always, or never preferred to use the COVID-19 app.[a]

Attitudinal statements[b]	Total (N=900), n (%)	Always (n=460), n (%)	Sometimes (n=214), n (%)	Never (n=226), n (%)
I find a contact tracing app to be useful				
Agree	414 (46.0)	309 (67.2)	87 (40.7)	18 (8.0)
Disagree	191 (21.2)	30 (6.5)	33 (15.4)	128 (56.6)
I worry about the security of a contact tracing app				
Agree	447 (49.7)	174 (37.8)	125 (58.4)	148 (65.5)
Disagree	198 (22.0)	140 (30.4)	38 (17.8)	20 (8.8)
I object to using a contact tracing app				
Agree	268 (29.8)	51 (11.1)	63 (29.4)	154 (68.1)
Disagree	356 (39.6)	274 (59.6)	68 (31.8)	14 (6.2)
I think it is very serious if I get infected with COVID-19				
Agree	581 (64.6)	322 (70.0)	134 (62.6)	125 (55.3)
Disagree	91 (10.1)	38 (8.3)	30 (14.0)	23 (10.2)
I think I would get seriously ill if I get infected with COVID-19				
Agree	568 (63.1)	311 (67.6)	120 (56.1)	137 (60.6)
Disagree	119 (13.2)	60 (13.0)	38 (17.8)	21 (9.3)
I think I have a high chance of getting infected with COVID-19				
Agree	198 (22.0)	123 (26.7)	46 (21.5)	29 (12.8)
Disagree	231 (25.7)	115 (25.0)	59 (27.6)	57 (25.2)
I think I have a high chance of getting seriously ill when infected with COVID-19				
Agree	374 (41.6)	214 (46.5)	78 (36.4)	82 (36.3)
Disagree	196 (21.8)	99 (21.5)	59 (27.6)	38 (16.8)
I think a COVID-19 app is a good way to control and fight COVID-19				
Agree	416 (46.2)	312 (67.8)	84 (39.3)	20 (8.8)
Disagree	222 (24.7)	41 (8.9)	52 (24.3)	129 (57.1)
I would use a contact tracing app if it becomes available				
Agree	388 (43.1)	307 (66.7)	72 (33.6)	9 (4.0)
Disagree	250 (27.8)	30 (6.5)	56 (26.2)	164 (72.6)
Survey evaluation				
The choice questions were clear				
Agree	738 (82.0)	385 (83.7)	172 (80.4)	181 (80.1)
Disagree	48 (5.3)	23 (5.0)	16 (7.5)	9 (4.0)
The choice questions were interesting				
Agree	645 (71.7)	373 (81.1)	161 (75.2)	111 (49.1)
Disagree	70 (7.8)	23 (5.0)	17 (7.9)	30 (13.3)
I could easily recognize the differences between the apps in the choice questions				
Agree	667 (74.1)	360 (78.3)	158 (73.8)	149 (65.9)
Disagree	76 (8.4)	39 (8.5)	21 (9.8)	16 (7.1)
I could easily choose between the apps in the choice questions				
Agree	633 (70.3)	332 (72.2)	154 (72.0)	147 (65.0)
Disagree	101 (11.2)	59 (12.8)	27 (12.6)	15 (6.6)

XSL•FO
RenderX

Attitudinal statements[b]	Total (N=900), n (%)	Always (n=460), n (%)	Sometimes (n=214), n (%)	Never (n=226), n (%)
I could easily have answered more choice questions				
Agree	617 (68.6)	331 (72.0)	152 (71.0)	134 (59.3)
Disagree	50 (5.6)	25 (5.4)	13 (6.1)	12 (5.3)
There were too many choice questions				
Agree	174 (19.3)	90 (19.6)	40 (18.7)	44 (19.5)
Disagree	469 (52.1)	259 (56.3)	115 (53.7)	95 (42.0)

[a]The percentages in columns 2-5 add up to 100% vertically, but in columns 3-5, 100% is the amount of people in that specific group.

[b]Reported on respondents who completely agreed or agreed and who completely disagreed or disagreed; percentages do not count up to 100%, as respondents who answered *neutral* were not included in this table.

DCE Results

All of the contact tracing app attributes influenced respondents' preferences (Table 3). On average, respondents preferred a COVID-19 contact tracing app that offers them additional benefits in terms of a small financial reward of €5 or €10 a month, being allowed to meet with groups of up to 10 or 30 people, and being tested if they were near a person who was infected. On average, respondents wanted to remain in charge of their own data by giving explicit permission to share the alert with the local health authorities (GGD) and entering a positive COVID-19 test result into the app themselves. They preferred alerts that are specific with respect to date and time.

Table 3. Mixed logit estimation results.

Attributes	Population mean (SE)	95% CI	Population SD (SE)	95% CI
No app	–3.44 (0.32)	–4.07 to –2.80	3.97 (0.28)	3.43 to 4.51
Group size				
3 people (reference)	0	N/A[a]	0	N/A
10 people	0.56 (0.09)	0.39 to 0.74	1.27 (0.14)	1.01 to 1.55
30 people	0.45 (0.10)	0.25 to 0.65	1.77 (0.13)	1.51 to 2.03
100 people	0.04 (0.12)	–0.20 to 0.28	2.50 (0.17)	12.17 to 2.84
Warning type				
Limited information (reference)	0	N/A	0	N/A
Detailed information	0.23 (0.06)	0.10 to 0.36	1.14 (0.09)	0.97 to 1.31
Who is warned				
Only you (reference)	0	N/A	0	N/A
You and automatically the local health authorities	0.01 (0.07)	–0.12 to 0.15	1.02 (0.09)	0.83 to 1.20
You and local health authorities after your consent	0.28 (0.07)	0.14 to 0.42	0.94 (0.09)	0.74 to 1.13
Testing after a warning				
Only when someone has symptoms (reference)	0	N/A	0	N/A
Everyone	0.40 (0.09)	0.23 to 0.57	1.96 (0.10)	1.77 to 2.15
Who can upload test results				
You (reference)	0	N/A	0	N/A
Local health authorities	0.05 (0.08)	–0.11 to 0.20	1.70 (0.08)	1.53 to 1.87
Financial incentive (€ per month)				
0 (reference)	0	N/A	0	N/A
5	0.85 (0.11)	0.62 to 1.07	2.44 (0.13)	2.18 to 2.70
10	1.29 (0.16)	0.97 to 1.60	3.70 (0.18)	3.35 to 4.05

[a]N/A: not applicable.

XSL•FO
RenderX

The attribute levels had the expected sign (showing theoretical validity), and 84% (760/900) of the respondents showed consistency in their choices (ie, they opted for the same alternative in the fifth and 14th DCE choice tasks). The utility pattern for the attribute *group size* was hyperbolic (ie, respondents preferred an app that allowed them to meet with 10 or 30 individuals instead of 3 individuals but were less positive about meeting with 100 individuals; Table 3). Furthermore, on average and relative to the other attributes, financial incentive was the most important attribute, while the attribute describing who enters into the app that the app user has tested positive was the least important attribute. However, the standard deviations of the alternative specific constant (ie, random intercept) and all attribute (levels) indicated a wide variation in preferences among respondents.

COVID-19 Contact Tracing App Uptake

Over half of the respondents (460/900, 51%) chose a contact tracing app in all choice tasks, which means that they preferred a contact tracing app with the least preferred specifications over no contact tracing app at all (see Table 1). About 25% (226/900) of the respondents had strict preferences against a contact tracing app (ie, they chose the opt-out alternative in all 14 DCE tasks) and could not be persuaded to choose a contact tracing app, not even the app with the most preferred specifications. The choices of the remaining 24% (214/900) of the respondents depended on the specifications of the app.

Assuming that the most realistic COVID-19 contact tracing app, given the situation in the Netherlands at the time of writing, is defined by an app that allows the app user to meet with 30 individuals at the same time, warns the app user that they were close to a person who was infected in the last 2 weeks, warns only the app user, allows the app user to undergo a COVID-19 test only after they has COVID-19 symptoms, is updated by the app user that they tested positive for COVID-19, and does not give the app user a financial incentive, the predicted adoption rate of the most realistic app was 64.1% (Table 4). One-way changes in our app's attribute levels had a relatively small impact on the predicted adoption rate (Figure 2). Changing from a contact tracing app with the least preferred to the most preferred attribute levels, the estimated adoption rate of the contact tracing app for the Dutch population increased from 59.3% to 65.7%. It should be noted that such changes do not perfectly correlate with the MIXL mean preference parameters; the degree of preference heterogeneity is an equally important determinant.

There are important sociodemographic differences in predicted adoption rates. Survey respondents aged between 15 and 34 years were more likely to use a contact tracing app than people 75 years or older. Survey respondents younger than 35 years were also more sensitive to the specifications of the app. When comparing the contact tracing app with the least preferred to the app with the most preferred specifications, the adoption rates increased from 72.4% to 81.7% for people younger than 35 years and decreased from 46.4% to 45.6% for people 75 years or older. The predicted adoption rates also differed by educational attainment. Survey respondents with lower levels of education were less likely to install the app and less sensitive to the specifications of the app. When comparing the least and most preferred contact tracing app, the adoption rates increased from 55.4% to 59.1% for the lowest educated respondents and from 59.4% to 67.8% for the highest educated respondents. Furthermore, as general health worsened, the proportion of respondents that always preferred a contact tracing app increased. That proportion was also higher among respondents with a lung disease, a kidney disease, and a compromised immune system compared to respondents without health problems.

We also observed important attitudinal differences in adoption. Respondents who indicated feeling safe in large groups (up to 1000 people), considered the chance of being infected with COVID-19 to be small, and did not think they would become seriously ill when infected by COVID-19 were more likely to reject the app irrespective of its specifications. That also holds for respondents who were more worried about the security of the app.

Besides the attributes included in this study, frequently mentioned reasons that favor the use of a COVID-19 contact tracing app were prevention (being able to control the virus), uncertainty reduction (ie, clarity and security), and more freedom. Frequently mentioned barriers were related to privacy concerns, safety concerns (data leaks), not owning a smartphone, potentially required out-of-pocket costs, and a low expected adoption rate in the society.

XSL•FO
RenderX

Table 4. Predicted COVID-19 contact tracing app adoption rates (%), stratified by age and education level.

Apps	All (n=900), %	15-34 years (n=268), %	35-54 years (n=281), %	55-74 years (n=265), %	≥75 years (n=86), %	Low educ[a] (n=274), %	Medium educ (n=342), %	High educ (n=284), %
Most preferred app[b]	65.7	81.7	61.4	60.4	46.0	59.1	69.2	67.8
Most realistic app[c]	64.1	79.4	60.4	58.4	45.6	59.3	67.1	65.0
Least preferred app[d]	59.3	72.4	55.4	54.3	46.4	55.4	62.3	59.4

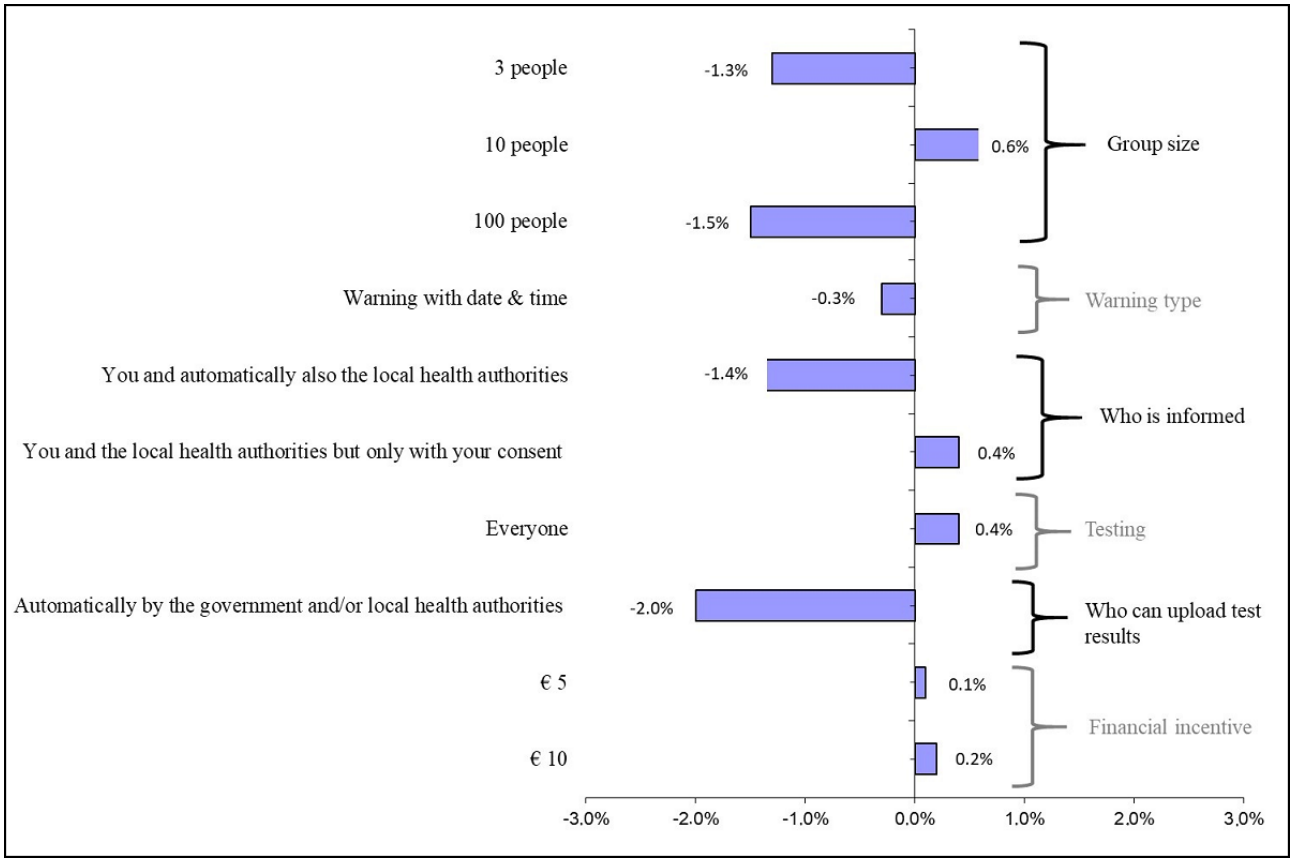
[a]educ: education.

[b]Specifications of the most preferred COVID-19 contact tracing app were the app user is allowed to meet with 10 individuals at the same time, warns the app user that they were close to a person who was infected in the last 2 weeks, warns the app user and the local health authorities (GGD) after permission, allows the app user to undergo a COVID-19 test, is updated by the app user that they tested positive for COVID-19, and does give the app user a financial incentive of €10 per month.

[c]Specifications of the most realistic app were allows the app user to meet with 30 individuals at the same time, warns the app user that they were close to a person who was infected in the last 2 weeks, warns the app user, allows the app user to undergo a COVID-19 test only after they have COVID-19 symptoms, is updated by the app user that they tested positive for COVID-19, and does not give the app user a financial incentive.

[d]Specifications of the least preferred app were allows the app user to meet with 3 individuals at the same time, warns the app user the date and time that they were close to a person who was infected, warns the local health authorities (GGD), allows the app user to undergo a COVID-19 test only after they have COVID-19 symptoms, is updated by the local health authorities (GGD) that the app user tested positive for COVID-19, and does not give the app user a financial incentive.

Figure 2. Univariate marginal estimates for increase in predicted adoption rate; attributes level changes vs base case. Note: The base case is the most realistic COVID-19 contact tracing app that allows the user to meet with 30 individuals, warns the user that they were close to a person who was infected in the last 2 weeks, warns the user and the local health authorities (GGD), allows the user to undergo a COVID-19 test only after they have COVID-19 symptoms, is updated by the local health authorities (GGD) that the user tested positive for COVID-19, and does not give the user a financial incentive. This base case is indicated as zero change in the probability of the x-axis.



Discussion

Main Findings

Our study suggests that an adoption rate as high as 66% can be achieved for a contact tracing COVID-19 app in the Netherlands. However, there is wide variation in preferences. Over half of the respondents always chose to use an app, about 25% of the respondents could never be persuaded to choose an app, and the choice of the remaining 24% of the respondents depended on the specifications of the app. Changing the specifications from the least to the most preferred increased the predicted adoption rate from 59% to 66% in the entire sample. In general,

XSL•FO
RenderX

app users prefer an app that offers them additional benefits in terms of being allowed to meet in groups of up to 10 and 30 people, and being tested immediately after the alert that they were near a person who was infected. App users want to remain in charge of their own data by giving explicit permission to share the alert with the public health authorities and entering a positive test result into the app themselves. They prefer alerts that are specific with respect to date and time. A small financial reward of €5 or €10 a month is appreciated.

Policy Context and Implications

The presented results should be viewed in the context of the discussions about a COVID-19 app in the Netherlands up until mid-April, when the data were collected. In the Netherlands, the peak in the number of patients with COVID-19 in hospital intensive care units was reached in the first week of April, and the curve was at the beginning of a decline, which was not yet clear at that time. Test capacity was limited and only available for individuals with severe symptoms and hospital staff.

In mid-April, the Dutch government organized a 2-day long *appathon* to review and test 7 different candidate apps that were selected from a long list of 660 proposed apps. The appathon was broadcasted on the internet. It turned out that the candidate apps all had privacy and security issues. Consequently, none of the apps in the appathon were selected and the Ministry of Health initiated the development of a new app, which would, from the outset, be designed with strict privacy and security in mind.

This governmental decision confirms that we took the right choice context for our study, namely, that the proposed app would meet the required privacy and data security issues instead of asking respondents to trade off privacy and security for benefits or specifications of the app. The Dutch authorities also made it clear that it would not adopt a contact tracing app that stored location data and that contact data should not be stored for longer than 2-3 weeks, which concurs with our decision not to include location data and length of data storage in the trade-offs either. In the literature, smartphone apps that seem to meet these conditions have been presented [12,20], and almost all new apps that appear in the continuously updated database of the Massachusetts Institute of Technology, which captures details of every significant automated contact tracing effort around the world, are based on the relatively secure and privacy respecting Bluetooth application programming interface, as introduced by Google and Apple [21].

Our finding that the adoption rate of the most realistic app was 34% points higher for respondents aged 15-34 years than for respondents 75 years or older may have policy implications. It suggests the need for a tailored communication strategy to maximize the uptake of the contact tracing app. Our data indicated that older adults felt less comfortable in larger groups and were more anxious about getting infected and getting seriously ill when infected, which is logical given the higher prevalence of health problems among older adults and their greater susceptibility to COVID-19. If this indicates that older adults would feel insufficiently protected by a contact tracing app, this may have contributed to the lower adoption rate among older adults. A tailored communication strategy should address these concerns and convince older adults of the necessity to share data to control a virus that largely spreads asymptomatically even if the app does not provide individual protection. Because people younger than 35 years were more sensitive to the specifications of the app they can be tempted to adopt the app by communicating the benefits to the app user, such as being allowed to meet in larger groups, immediately getting tested after contact with a person who was infected with COVID-19, and perhaps a financial reward.

Appropriately addressing the observed attitudinal differences toward adopting the app is another challenge for policy makers. Perhaps the group that feels safe in large groups of app users, thinks the chance is small they will get the virus, and does not think that they will become seriously ill if they are infected by the virus represents a group that downplays the seriousness of the situation. Education that is specifically tailored to these attitudes might be necessary.

Comparison With Other Studies

The context that respondents were offered in this study (ie, that the app would comply with privacy and security legislations) is likely to have contributed to the high adoption rates. Nevertheless, the only other choice-based study about COVID-19 apps published so far has reported even higher adoption rates, despite the fact that they did include attributes like using the app to enforce self-isolation, anonymity, length of data storage, and responsibility for the app project [22]. In this UK-wide study, the app with recommended specifications had a 73.5% adoption rate compared with 64.1% in our study. One of the possible reasons could be that the study in the United Kingdom was done earlier, when the infection peak had not yet been reached and people felt more insecure.

There have been several other surveys about COVID-19 apps, but these were not choice based and did not ask respondents to trade-off positive versus negative characteristics of an app, as is done in a DCE. In a large international survey conducted in France, Germany, Italy, the United Kingdom, and the United States, strong support for a contact tracing app was found regardless of the respondents' country or background characteristics [23].

Limitations

The study was conducted in a representative sample of the Dutch population with respect to age, gender, and education. Nevertheless, we should acknowledge that respondents were members of an internet panel of a market research organization, which makes them more likely to have a positive attitude toward internet and digital devices in general and thus more likely to adopt an app. This is related to our finding that people who already use health apps were more likely to prefer the contact tracing app than people who do not use health apps. However, the impact of using an internet panel is probably limited, as 88% of Dutch citizens owns a smartphone and over one-third has a health app installed on it [24].

It is obvious that the adoption rates in our study are based on stated preferences, which might differ from revealed preferences. First, although stated preferences may accurately reflect an individual's intention to use an app, they may not accurately

XSL•FO
RenderX

predict real-world use of an app [25]. There are few external validation studies of DCEs, but there are cases such as influenza vaccination and colorectal cancer screening in which over 90% of choices were correctly predicted at an individual level [25,26]. However, one may argue that there are less privacy and security issues involved in these cases. Second, the presented analyses do not take dynamics into account and thus only predict the potential uptake of a contact tracing app and not the time it takes for the predicted uptake to be achieved. The latter likely depends on the attractiveness of the app but also on external factors, including the amount of effort from local health authorities and the government to promote the contact tracing app using public health campaigns. Third, the achieved adoption rates of the contact tracing app will likely depend on the timing of its launch. If people still recognize the seriousness of the COVID-19 pandemic and the necessity of a contact tracing app as they did in our study (based on mid-April 2020 data collection), they may be willing to cooperate and share personal data more easily than if they view the crises as being defeated. With COVID-19 restrictions currently being eased, it seems conceivable that respondent preferences could change accordingly.

Conclusion

Based on the presented results, with predicted app adoption rates ranging from 59% to 66%, we conclude that it is possible for a secure and privacy-respecting COVID-19 contact tracing app to reach a high adoption rate. Taking account of the preferred specifications of the app will contribute to a more widespread adoption. The main challenge will be to increase the adoption rate among older adults (≥75 years of age), since even the app with the most preferred characteristics had a 36%-point lower adoption rate compared to respondents 35 years and younger.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Cherries checklist.
[PDF File (Adobe PDF File), 223 KB - mhealth_v8i10e20741_app1.pdf]

References

1. Contact tracing in the context of COVID-19: interim guidance, 10 May 2020. World Health Organization. 2020 May 10. URL: https://apps.who.int/iris/handle/10665/332049
2. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Global Health 2020 Apr;8(4):e488-e496. [doi: 10.1016/s2214-109x(20)30074-7]
3. Ngonghala CN, Iboi E, Eikenberry S, Scotch M, MacIntyre CR, Bonds MH, et al. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. Math Biosci 2020 Jul;325:108364 [FREE Full text] [doi: 10.1016/j.mbs.2020.108364] [Medline: 32360770]
4. Cheng H, Jian S, Liu D, Ng T, Huang W, Lin H, Taiwan COVID-19 Outbreak Investigation Team. Contact tracing assessment of COVID-19 transmission dynamics in Taiwan and risk at different exposure periods before and after symptom onset. JAMA Intern Med 2020 May 01 [FREE Full text] [doi: 10.1001/jamainternmed.2020.2020] [Medline: 32356867]
5. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491) [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
6. De Jonge H. Handelingen. Tweede Kamer. 2020 Apr 08. URL: https://www.tweedekamer.nl/
7. TraceTogether. 2020. URL: https://www.tracetogether.gov.sg/
8. Cellan-Jones R. Coronavirus: Isle of Wight contact-tracing app trial - a mixed verdict so far. BBC News. 2020 May 18. URL: https://www.bbc.com/news/technology-52709568
9. Hensher DA, Rose JM, Greene WH. Applied Choice Analysis: A Primer. New York: Cambridge University Press; 2005.
10. Ryan M. Discrete choice experiments in health care. BMJ 2004 Feb 14;328(7436):360-361 [FREE Full text] [doi: 10.1136/bmj.328.7436.360] [Medline: 14962852]
11. McFadden D. Conditional logit analysis of qualitative choice behavior. University of California at Berkeley 1973.
12. Abeler J, Bäcker M, Buermeyer U, Zillessen H. COVID-19 contact tracing and data protection can go together. JMIR mHealth uHealth 2020 Apr 20;8(4):e19359 [FREE Full text] [doi: 10.2196/19359] [Medline: 32294052]
13. Johnson FR, Yang J, Reed SD. The internal validity of discrete choice experiment data: a testing tool for quantitative assessments. Value Health 2019 Feb;22(2):157-160 [FREE Full text] [doi: 10.1016/j.jval.2018.07.876] [Medline: 30711059]
14. Tervonen T, Schmidt-Ott T, Marsh K, Bridges JF, Quaife M, Janssen E. Assessing rationality in discrete choice experiments in health: an investigation into the use of dominance tests. Value Health 2018 Oct;21(10):1192-1197. [doi: 10.1016/j.jval.2018.04.1822] [Medline: 30314620]
15. Jonker MF, Donkers B, de Bekker-Grob E, Stolk EA. Attribute level overlap (and color coding) can reduce task complexity, improve choice consistency, and decrease the dropout rate in discrete choice experiments. Health Econ 2019 Mar;28(3):350-363 [FREE Full text] [doi: 10.1002/hec.3846] [Medline: 30565338]

XSL•FO
RenderX

16. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. Patient 2015 Oct;8(5):373-384 [FREE Full text] [doi: 10.1007/s40271-015-0118-z] [Medline: 25726010]
17. Becker MH. The health belief model and personal health behavior. Health Education Monogr 1974;2:324-473.
18. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
19. Train K. Discrete Choice Methods with Simulation. New York: Cambridge University Press; 2009.
20. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-peer contact tracing: development of a privacy-preserving smartphone app. JMIR mHealth uHealth 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: 10.2196/18936] [Medline: 32240973]
21. O'Neill P, Ryan-Mosley T, Johnson B. A flood of coronavirus apps are tracking us. Now it's time to keep track of them. MIT Technology Review. 2020 May 07. URL: https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/
22. Wiertz C, Banerjee A, Acar OA, Ghosh A. Predicted adoption rates of contact tracing app configurations - insights from a choice-based conjoint study with a representative sample of the UK population. SSRN J 2020 Apr 30 [FREE Full text] [doi: 10.2139/ssrn.3589199]
23. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: cross-country survey evidence. SSRN J 2020. [doi: 10.2139/ssrn.3590505]
24. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: a source of new digital inequalities? Inf Soc 2018 Apr 26;34(3):183-193. [doi: 10.1080/01972243.2018.1438550]
25. de Bekker-Grob EW, Swait JD, Kassahun HT, Bliemer MC, Jonker MF, Veldwijk J, et al. Are healthcare choices predictable? The impact of discrete choice experiment designs and models. Value Health 2019 Sep;22(9):1050-1062 [FREE Full text] [doi: 10.1016/j.jval.2019.04.1924] [Medline: 31511182]
26. de Bekker-Grob E, Donkers B, Bliemer M, Veldwijk J, Swait J. Can healthcare choice be predicted using stated preference data? Soc Sci Med 2020 Feb;246:112736 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112736] [Medline: 31887626]

Abbreviations

DCE: discrete choice experiment
MIXL: mixed logit

Edited by G Eysenbach; submitted 27.05.20; peer-reviewed by T Yasaka, J Li; comments to author 20.06.20; revised version received 30.06.20; accepted 11.08.20; published 09.10.20.

Please cite as:
Jonker M, de Bekker-Grob E, Veldwijk J, Goossens L, Bour S, Rutten-Van Mölken M
COVID-19 Contact Tracing Apps: Predicted Uptake in the Netherlands Based on a Discrete Choice Experiment
JMIR Mhealth Uhealth 2020;8(10):e20741
URL: https://mhealth.jmir.org/2020/10/e20741
doi:10.2196/20741
PMID:32795998

©Marcel Jonker, Esther de Bekker-Grob, Jorien Veldwijk, Lucas Goossens, Sterre Bour, Maureen Rutten-Van Mölken. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 09.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Syndromic Surveillance Insights from a Symptom Assessment App Before and During COVID-19 Measures in Germany and the United Kingdom: Results From Repeated Cross-Sectional Analyses

Alicia Mehl[1*], BA, MSc; Francois Bergey[1*], MSc; Caoimhe Cawley[1], BSc, MSc, PhD; Andreas Gilsdorf[1], MD, Dr med

Department of Epidemiology & Public Health, Ada Health GmbH, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Alicia Mehl, BA, MSc
Department of Epidemiology & Public Health
Ada Health GmbH
Karl-Liebknecht-Str 1
Berlin, 10178
Germany
Phone: 49 30 403 67 390
Email: alicia.mehl@ada.com

Abstract

Background: Unprecedented lockdown measures have been introduced in countries worldwide to mitigate the spread and consequences of COVID-19. Although attention has been focused on the effects of these measures on epidemiological indicators relating directly to the infection, there is increased recognition of their broader health implications. However, assessing these implications in real time is a challenge, due to the limitations of existing syndromic surveillance data and tools.

Objective: The aim of this study is to explore the added value of mobile phone app–based symptom assessment tools as real-time health insight providers to inform public health policy makers.

Methods: A comparative and descriptive analysis of the proportion of all self-reported symptoms entered by users during an assessment within the Ada app in Germany and the United Kingdom was conducted between two periods, namely before and after the implementation of “Phase One” COVID-19 measures. Additional analyses were performed to explore the association between symptom trends and seasonality, and symptom trends and weather. Differences in the proportion of unique symptoms between the periods were analyzed using a Pearson chi-square test and reported as log2 fold changes.

Results: Overall, 48,300-54,900 symptomatic users reported 140,500-170,400 symptoms during the Baseline and Measures periods in Germany. Overall, 34,200-37,400 symptomatic users in the United Kingdom reported 112,100-131,900 symptoms during the Baseline and Measures periods. The majority of symptomatic users were female (Germany: 68,600/103,200, 66.52%; United Kingdom: 51,200/71,600, 72.74%). The majority were aged 10-29 years (Germany: 68,500/100,000, 68.45%; United Kingdom: 50,900/68,800, 73.91%), and about one-quarter were aged 30-59 years (Germany: 26,200/100,000, 26.15%; United Kingdom: 14,900/68,800, 21.65%). Overall, 103 symptoms were reported either more or less frequently (with statistically significant differences) during the Measures period as compared to the Baseline period, and 34 of these were reported in both countries. The following mental health symptoms (log2 fold change, *P* value) were reported less often during the Measures period: *inability to manage constant stress and demands at work* (–1.07, *P*<.001), *memory difficulty* (–0.56, *P*<.001), *depressed mood* (–0.42, *P*<.001), and *impaired concentration* (–0.46, *P*<.001). *Diminished sense of taste* (2.26, *P*<.001) and *hyposmia* (2.20, *P*<.001) were reported more frequently during the Measures period. None of the 34 symptoms were found to be different between the same dates in 2019. In total, 14 of the 34 symptoms had statistically significant associations with weather variables.

Conclusions: Symptom assessment apps have an important role to play in facilitating improved understanding of the implications of public health policies such as COVID-19 lockdown measures. Not only do they provide the means to complement and cross-validate hypotheses based on data collected through more traditional channels, they can also generate novel insights through a real-time syndromic surveillance system.

XSL•FO
RenderX

(JMIR Mhealth Uhealth 2020;8(10):e21364) doi:10.2196/21364

KEYWORDS

epidemiology; participatory epidemiology; participatory surveillance; COVID-19 symptom assessment apps; symptom checker apps; syndromic surveillance; COVID-19 measures; COVID-19 lockdown; digital public health; health effects of COVID-19 measures, infoveillance

Introduction

Background

Since its emergence at the end of 2019, COVID-19 has had an enormous and wide-ranging impact, with millions of confirmed cases and hundreds of thousands of deaths reported worldwide [1]. Governments have introduced a series of measures ranging from work and school closures to social distancing and lockdowns to mitigate the spread and consequences of infection. The measures have been unprecedented on many levels: in how disruptive they are to daily life, by the proportions of populations affected, in the duration of implementation, and by their global reach. As a result, the daily lives of millions have been upended.

The extent of the countermeasures to the pandemic raises questions as to their impact on public and individual health. Although much of the initial focus in the medical and scientific community has been on understanding the virus and infection itself [2], as well as the effects of various policy measures on epidemiological indicators relating directly to COVID-19 [3,4], there has also been increased recognition of the broader health effects [5].

Some of the broader implications concern the direct and immediate consequences of lockdown. For instance, as people are distressed due to social isolation and the economic fallout of the crisis, an upsurge in the incidence and severity of mental health problems has been predicted [5-11]. Increased handwashing (a primary recommendation to reduce transmission of COVID-19) is expected to result in increased skin irritation and dermatitis [12,13]. Other implications are related to more indirect factors such as delays and cancellations of surgeries and nonurgent treatments for patients with cancer and other diseases [14-16], interruptions in drug and commodities supply chains [17,18], and drops in vaccination rates for vaccine-preventable diseases [19,20]. The consequences of these interruptions to medical services during this pandemic will likely create a higher morbidity, but the impact may not be visible for several years to come [17].

Assessing the health implications of lockdown policies in real-time is a challenge. Apart from the time lag, evidence cited in support of purported health effects is often anecdotal or based on surveys conducted among medical professionals. The latter provide important insights into changes in health symptoms identified at the point of contact, but may not account for the effects of social distancing guidelines on health-seeking behavior. Traditional syndromic surveillance data [21] provides valuable information for a defined set of indicators over time, but is unable to distinguish whether the trends reflect changes in disease incidence or in the uptake of health services. More significant efforts toward data collection among the general public are currently being undertaken [22]. However, these are missing baseline data from before the emergence of COVID-19 that would facilitate a clearer understanding of the impact. In general, surveys can provide snapshots of a highly dynamic situation, but tend to be restricted in scope, meaning that changes in health beyond predefined indicators risk being overlooked.

In the past years, a growing number of studies have shown the benefits of mobile phone app–based symptom assessment tools for improved health outcomes (eg, in lowering the barrier to seek help) [23-29]. If the literature on symptom assessment tools has primarily focused on individual health benefits (eg, early detection of rare diseases [30]), recent research has also demonstrated their potential contribution to public health. Specifically with regard to COVID-19, symptom tracker apps have helped flag anosmia (loss of the sense of smell) as a potentially relevant symptom of infection [31,32], and have challenged common understandings of presenting symptoms of COVID-19 [33].

Goal of This Study

In the present study, we add to this body of literature by exploring the added value of symptom assessment tools for the analysis of broader health implications resulting from public health interventions such as lockdown measures. We were particularly interested in symptoms not related to COVID-19 that could be potentially overlooked as an unintended consequence of the measures. A tool like Ada can provide real-time insights for policy makers, informing their decisions as they aim for the right balance of such measures. This study is based on self-reported symptom data from Ada, a digital symptom assessment app that uses a probabilistic reasoning engine that collects demographic information, symptoms, and medical history to suggest possible conditions and then guides individuals to the most appropriate care. The Ada app is available in seven languages and has over ten million users. It is described in Gilbert et al [34] in further detail. Using an inductive approach and focusing on the immediate impact of COVID-19 control measures, we compared symptom data reported by users in Germany and the United Kingdom before and during the first phase of COVID-19 lockdown measures to identify changes and continuities in the incidence of self-reported symptoms.

The aim of this study was to explore the potential of the Ada symptom assessment app to generate real-time health insights to inform public health policy makers.

Methods

Study Focus

This analysis is a comparative descriptive study of self-reported symptoms entered by users in an Ada assessment completed during the “Phase One” COVID-19 interventions (the Measures

XSL•FO
RenderX

period) compared to a Baseline period in Germany and the United Kingdom. The Measures periods started when all five major nonpharmaceutical interventions described in Flaxman et al [4] were implemented in the respective countries. Specifically, these interventions are school closures, case-based measures (strong recommendation of self-isolation when showing COVID-19–like symptoms), banning of public events, encouragement of social distancing, and lockdowns. The Measures period lasted until April 22, 2020, when data was extracted. The end of the Baseline period was defined as the day before any of the five interventions were implemented in the respective countries, and the length of the Baseline period was equal to the Measures period in the same country. The period between the Baseline and Measures periods is excluded from the analysis due to the partial implementation of "Phase One" COVID-19 measures. This is an exploratory analysis of all symptoms (not only COVID-19–related) reported by symptomatic users during the defined periods. In this study, we consider a symptomatic user as one who completed at least one assessment during the periods of analysis, whether the assessment was self-reported or reported by someone on their behalf (ie, a legal guardian if under 16 years of age). We considered symptom patterns to be trends if in both countries, the proportion that a specific symptom was reported (out of all reported symptoms) was significantly different between the two periods.

Additional Analyses

Ad hoc analyses were later performed to test identified trends for symptoms against selected potential confounders, such as seasonality and weather. To explore the potential impact of seasonality on trends, the results were compared to the results of the same analysis conducted for the same dates in 2019. To explore the impact of weather on trends, associations between the monthly proportion of reported symptoms and weather variables (average monthly temperature [°C], monthly precipitation [mm], and monthly hours of sunshine) were investigated during the period from January 2019 to March 2020. The weather analysis was restricted to Germany.

Participants

All assessments completed by Ada users in Germany and the United Kingdom during the Baseline and Measures periods were included in the analysis. We analyzed pseudonymized health data for public health purposes, according to the European General Data Protection Regulation (GDPR), and users are duly informed of such use of their data (information available at any time in Ada's Privacy Policy). Additionally, users maintain their right to object to such processing for reasons arising from their particular situation, as required by the GDPR.

Variables

An Ada assessment consists of several different parts: (1) the user enters an unlimited number of symptoms, (2) the user is asked about other potential symptoms they could have, and (3) an assessment result is provided, with conditions that could potentially explain the reported symptoms and adequate triage. This analysis only includes symptoms that are self-reported by a user in the first part of the assessment: that is, responses to the initial question "Let's start with the symptom that's troubling you the most," followed by "Do you have any other symptoms?" Upon entering free text, the user is then given a range of medically-curated options (based on linguistic relevance) to select from.

The variable of interest, $S_{i,k}$, representing the proportion a symptom i is self-reported during the period k is defined as

Where $s_{i,j,k}$ equals 1 if the user j self-reported the symptom i during the period k, and equals 0 otherwise; and $x_{j,k}$ equals 1 if the user j has completed one assessment during the period k, and equals 0 otherwise.

The age variable was grouped using the following categories (years): 0-9, 10-19, 20-29, 30-39, 40-59 and ≥60.

Weather data (average monthly temperature [°C], monthly precipitation [mm], and monthly hours of sunshine) for Germany was extracted from the Deutscher Wetter Dienst database for the period from January 2019 to March 2020 [35].

For ease of reporting and to aid interpretation, symptoms that were reported in significantly different proportions during the Baseline and Measures periods were grouped using the International Classification of Diseases, Version 2019 (ICD-10) of the World Health Organization [36]. ICD-10 R subgroups, named "Symptoms, signs and abnormal clinical and laboratory symptoms, not elsewhere classified" were used when possible. Similar categories were grouped together later, as presented in Table 1.

XSL•FO
RenderX

Table 1. List of groups based on the International Classification of Diseases, Version 2019 classification.

Group name	International Classification of Diseases, Version 2019 classification
Circulatory and respiratory systems	R00-R09, J
Digestive system and abdomen	R10-R19
Skin and subcutaneous tissue	R20-R23, L
Musculoskeletal	M
Genitourinary system	R30-R39, N
Cognition, perception, emotional state, and behavior	R40-R46, F
Speech and voice	R47-R49
General symptoms and signs	R50-R69
Nervous system	G
Eye and adnexa	H

Bias

As Ada's medical model and databases are continuously updated, we defined the Baseline period to be as close as possible to the Measures period to limit the impact of these changes on the data. All modeled symptoms that were added, deleted, significantly modified, or significantly affected by the modification of any other symptom from the first day of the Baseline period until the end of the Measures period were removed from this study.

Statistical Methods

Sex and age groups of symptomatic users were reported as percentages and tested for differences between the periods with a Pearson chi-square test. Differences in the proportion of symptoms between the periods were reported as log2 fold changes and were analyzed with a Pearson chi-square test. A log2 fold change of 0.5 means that the proportion of that reported symptom was 1.41 times as large during the Measures period compared to the Baseline. A log2 fold change of 1 is interpreted as being twice as large during the Measures period compared to the Baseline, and a log2 fold change of 2 is four times as large. Conversely, a log2 fold change of –1 means that the proportion of the reported symptom was twice as large during the Baseline period compared to the Measures period. In general, log2 fold change calculations are helpful in understanding relative differences in the proportions of users reporting each symptom between the two periods, but do not reflect how common reporting of that symptom was overall. Associations between weather variables and the proportion of symptoms were tested based on the Spearman Rank correlation coefficient.

When required, *P* values were adjusted for multiple testing using the false discovery rate method. *P* values ≤.05 were considered statistically significant. Statistical analyses and figures were executed using R (Version 3.6.1; R Foundation for Statistical Computing).

The analysis was done using exact numbers, but results representing user numbers are presented rounded to the closest hundred, to ensure a fully anonymized presentation of the results.

Results

Principal Results

An overview of the Baseline and Measures periods in Germany and the United Kingdom (numbers of Ada users, numbers of symptoms reported) are shown in Table 2.

Table 2. Key parameters.

Key parameters	Germany		United Kingdom	
	Baseline	Measures	Baseline	Measures
Dates of period (MM/DD/YY)	02/03/20-03/05/20	03/22/20-04/22/20	02/11/20-03/11/20	03/24/20-04/22/20
Number of days per period	32	32	30	30
Total number of users[a]	467,000	483,100	488,800	501,300
Number of symptomatic users[a]	54,900	48,300	37,400	34,200
Number of reported symptoms by symptomatic users[a]	170,400	140,500	131,900	112,100

[a]Numbers were rounded to the nearest hundred.

Demographic characteristics of users are shown in Table 3. During both the Baseline and Measures periods, the majority of symptomatic users in both countries were female (Germany Baseline: 36,300/54,900, 66.19%; Germany Measures: 32,300/48,300, 66.90%; United Kingdom Baseline: 26,600/37,400, 71.17%; United Kingdom Measures:

XSL•FO
RenderX

24,600/34,200, 71.94%). The majority were aged 10-29 years (Germany Baseline: 37,000/53,200, 69.51%; Germany Measures: 31,400/46,800, 67.13%; United Kingdom Baseline: 27,200/35,800, 75.76%; United Kingdom Measures: 23,700/32,900, 71.89%). Those aged 30-59 years represented roughly one-quarter of symptomatic users (Germany Baseline: 13,300/53,200, 24.94%; Germany Measures: 12,900/46,800, 27.53%; United Kingdom Baseline: 7,100/35,800, 19.92%; United Kingdom Measures: 7,700/32,900, 23.54%). The number of symptomatic users in the Baseline period (Germany: 54,900; United Kingdom: 37,400) was slightly higher than in the Measures period (Germany: 48,300; United Kingdom: 34,200).

Table 3. Demographic characteristics of the study population.

Demographic characteristics	Germany			United Kingdom		
	Baseline, n[a] (%)	Measures, n[a] (%)	*P* value[b]	Baseline, n[a] (%)	Measures, n[a] (%)	*P* value[b]
Sex						
Female	36,300 (66.2)	32,300 (66.9)	.02	26,600 (72.2)	24,600 (72.9)	.02
Age[c] (years)						
0-9	1300 (2.4)	700 (1.6)	<.001	900 (2.4)	700 (2.0)	<.001
10-19	17,100 (32.1)	13,700 (29.2)	<.001	16,200 (45.1)	13,600 (41.2)	<.001
20-29	20,000 (37.5)	17,800 (37.9)	.28	11,000 (30.7)	10,100 (30.7)	.95
30-39	6400 (12.0)	6000 (12.9)	<.001	3100 (8.7)	3300 (9.9)	<.001
40-59	6900 (13.0)	6800 (14.6)	<.001	4000 (11.2)	4500 (13.6)	<.001
≥60	1700 (3.2)	1800 (3.9)	<.001	700 (2.1)	900 (2.7)	<.001

[a]Numbers were rounded to the nearest hundred.

[b]Values were obtained from Pearson chi-square tests for differences between Baseline and Measures data.

[c]Users who did not report a birth year were excluded from the analysis.

In total, 21 symptoms were excluded from the analysis as they had been added to or deleted from the medical model during either the Baseline or Measures period. In addition, three other symptoms were excluded from the analysis as there were significant changes in the associated terms (text entered by users to match the description of a symptom to the term used in the model), which could affect the number of times a symptom is reported. A list of these symptoms is included in Table 1 of Multimedia Appendix 1.

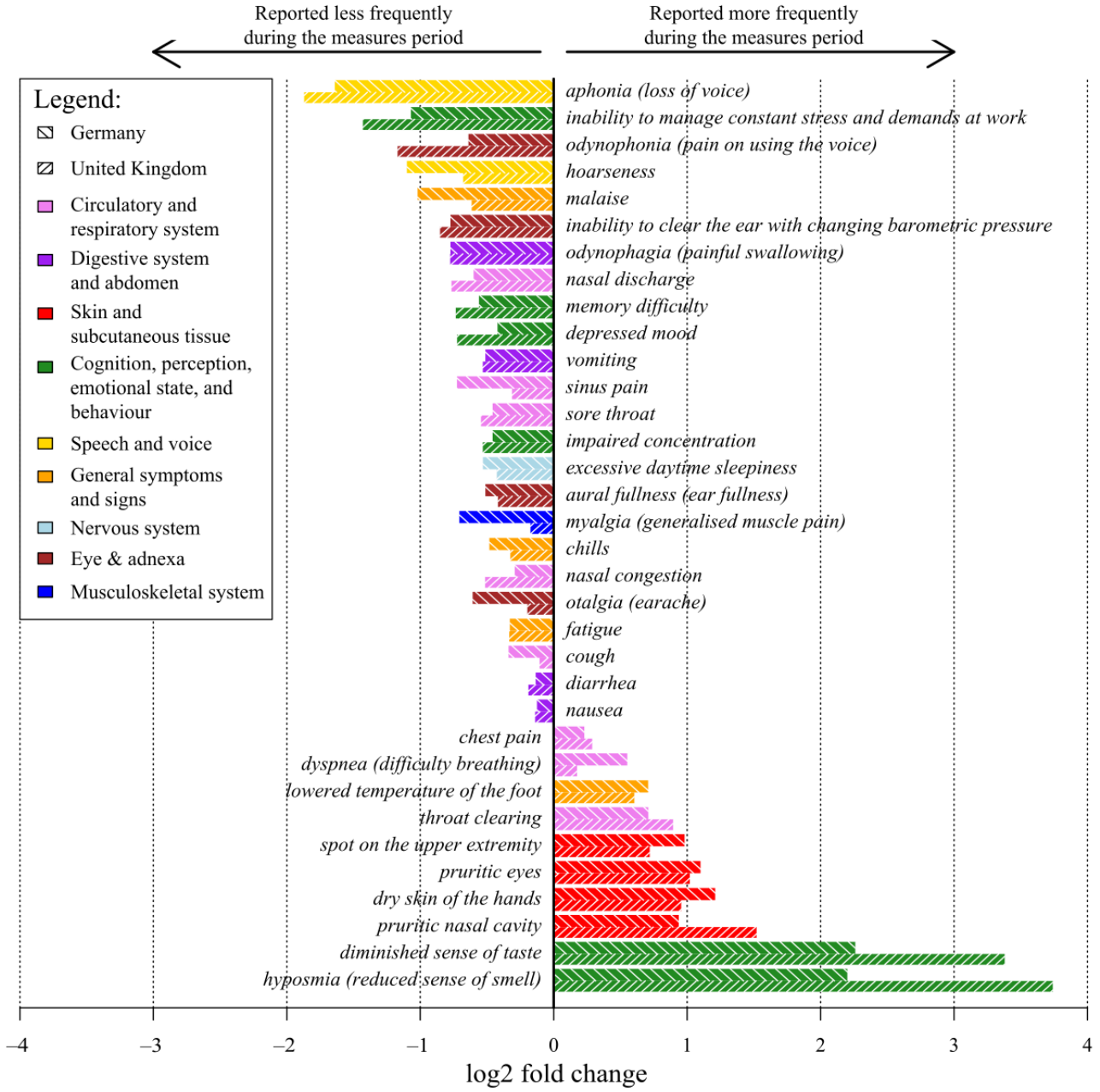
Main Analyses

Out of 1328 symptoms investigated in Germany and 1294 symptoms investigated in the United Kingdom, 103 symptoms were reported either more or less frequently, in either country, during the Measures period as compared to the Baseline period. The complete results can be found in Tables 2 and 3 of Multimedia Appendix 1.

Figure 1 presents the 34 symptoms that showed a statistically significant difference in both countries. Overall, 24 symptoms were reported less often and 10 were reported more often during the Measures period than during the Baseline period.

XSL•FO
RenderX

Figure 1. Relative difference (log2 fold change values) in the proportions of Ada users' reported symptoms with statistically significant differences between the baseline and COVID-19 measures periods in Germany and the United Kingdom.



Out of the 34 significant symptoms, all skin- and tissue-related symptoms (*pruritic nasal cavity*, *spot on the upper extremity*, *dry skin of the hands*, and *pruritic eyes*) were reported more frequently during the Measures period. In contrast, all speech and voice symptoms (*aphonia* and *hoarseness*), all eye and adnexa symptoms (*odynophonia*, *inability to clear the ear with changing barometric pressure*, *aural fullness*, and *otalgia*), all digestive system and abdomen symptoms (*odynophagia*, *diarrhea*, and *nausea*), and all musculoskeletal and nervous systems symptoms (*excessive daytime sleepiness* and *myalgia*) were reported less frequently during the Measures period.

In the cognition, perception, emotional state, and behavior symptoms group, two perception symptoms (*diminished sense of taste* and *hyposmia*) were reported more frequently during the Measures period, whereas the mental health symptoms (*inability to manage constant stress and demands at work*, *memory difficulty*, *depressed mood*, and *impaired concentration*) were reported less frequently during the Measures period. Out of the circulatory and respiratory symptoms, three were reported more frequently during the Measures period (*throat clearing*, *dyspnea*, and *chest pain*) and four were reported less frequently (*nasal discharge*, *sinus pain*, *sore throat*, and *cough*). One general symptom was reported more frequently (*lowered temperature of the foot*) and three (*malaise*, *chills*, and *fatigue*) were reported less frequently during the Measures period.

Additional Analyses

Out of the 34 symptoms found to be different between the Baseline and Measures periods in both Germany and the United Kingdom in 2020, none were found to be different between the same periods in 2019 in both countries. However, looking at the countries separately, in Germany, the data shows that 8 of the 34 significant symptoms were also reported less frequently

XSL•FO
RenderX

during the Measures period in 2019: *cough*, *chills*, *nasal discharge*, *myalgia*, *sore throat*, *malaise*, *fatigue*, and *sinus pain*. The complete results can be found in Tables 4 and 5 of Multimedia Appendix 1.

The monthly weather report in February 2020 (which corresponds to most of the Baseline period) differed from that of April 2020 (which corresponds to most of the Measures period) in Germany. The average temperature increased by 5.1 °C (from 5.3 °C to 10.4 °C), monthly precipitation decreased by 107.8 mm (from 124.1 mm to 16.3 mm), and the number of hours of sunshine per month increased by 228.5 hours (from 63.9 to 292.4 hours). In total, 14 of the 34 significant symptoms had statistically significant associations with weather. Increased temperature was positively associated with reporting *spot on the upper extremity* and negatively associated with *chest pain*, *lowered temperature of the foot*, *odynophagia*, *malaise*, *myalgia*, *cough*, *otalgia*, *chills*, *vomiting*, *sinus pain*, and *nasal congestion*. Increased hours of sunshine were positively associated with *pruritic eyes*, *spot on the upper extremity*, and *pruritic nasal cavity*, and negatively associated with *lowered temperature of the foot* and *sinus pain*. The complete results can be found in Table 6 and Figure 1 of Multimedia Appendix 1.

Discussion

Principal Findings

The results presented above show significant differences in the frequency and proportion of self-reported symptoms in Ada assessments before and after the implementation of measures aimed at reducing the transmission of COVID-19. Importantly, the same differences were found in both Germany and the United Kingdom, despite the divergent trajectories of these countries in the lead-up to the implementation of lockdown policies [37], as well as other national differences such as those relating to health systems [38]. Furthermore, these differences were not found during the same time periods in 2019, suggesting that the lockdown measures could have contributed to the results.

Many of the observed differences were to be expected. The reduced frequency of reported respiratory symptoms (*nasal discharge*, *sore throat*, *cough*, *sinus pain*, *nasal congestion*, *hoarseness*, *odynophonia*, *aphonia*) and influenza-like illness (ie, *malaise*, *fatigue*, *chills*, *myalgia*) following the measures is understandable as the cold and flu season has also waned exceptionally rapidly during this period [39], likely facilitated by reduced contact resulting from lockdown. To support this interpretation, *cough*, *chills*, *nasal discharge*, *myalgia*, *sore throat*, *malaise*, *fatigue*, and *sinus pain* were also reported less frequently during the Measures period in 2019 in Germany, suggesting that seasonal changes are reflected in the data. The increased reporting of dry hands following the measures is consistent with more frequent handwashing during the Measures period, as expected by dermatologists [12,13]. Increased reports of *pruritic eyes* and *pruritic nasal cavity* following the measures could be a consequence of seasonal hay fever (known to be worse in spring than during winter months due to increased pollen in the air [40]), as these symptoms were found to be associated with increased sunshine (presumably when people spend more time outdoors).

The reduction of gastrointestinal symptoms could be associated with the closing of preschool/day care settings (as they are known to contribute to the spread of these diseases [41]), restaurant closures, or improved hand hygiene. The decrease in ear problems could be a result of the sharp reduction in air travel, resulting in fewer people experiencing pressure adjustment problems called “airplane ear” [42], or related to the end of the cold and flu season. More research is needed to explore these hypotheses.

We also observed an increase in the reporting of *hyposmia*, *diminished sense of taste*, and *dyspnea* during the Measures period. These are less frequent but also typical COVID-19 symptoms and were increasingly recognized in the general public [31,43,44]. They were not found to be associated with seasonality or weather. This increase could be related to COVID-19 infections, or an artefact resulting from increased awareness of these symptoms due to media coverage.

A more surprising result is that *depressed mood*, *inability to manage constant stress and demands at work*, *impaired concentration*, *memory difficulty*, and *excessive daytime sleepiness* were reported in a lower proportion during the Measures period. This is not only contrary to conventional wisdom, but also runs counter to what was observed during previous infectious disease outbreaks (such as severe acute respiratory syndrome [SARS], Middle East respiratory syndrome [MERS], and Ebola [45,46]), as well as to the literature reporting on the effects of COVID-19 on mental health [6-11,47,48]. Despite the fact that the temperature was warmer during the Measures period and there was more sunshine than usual, the analysis on weather data for Germany did not show a significant impact on the changed mental health symptoms. Our findings are supported by a growing body of evidence from ongoing studies [49] and recently published research [50], as well as anecdotal evidence reported in various mainstream media reports [51-53] that suggest that, at least in the short term, the mental health effects of the COVID-19 measures may not be as negative as expected. One factor may be that the stress of everyday life and work/study is reduced during lockdown and when working or studying from home. In addition, the reduction of *excessive daytime sleepiness* may be due to the fact that people may sleep better or more during lockdown as there are fewer opportunities to go out and socialize in the evenings. Ongoing studies into the mental health effects of COVID-19 and its countermeasures may shed further light on these questions.

Limitations

It is important to interpret the study results taking into consideration the characteristics of the study population and the normal use case for Ada. Due to the specific age and sex distribution of users (predominantly young and female), the results are not generalizable to other population groups, especially the elderly. Furthermore, this analysis was limited to Germany and the United Kingdom (due to sufficient user numbers), and represents a two-month snapshot of reported symptoms. Users who know they have a disease might not use Ada if their symptoms deteriorate, as the cause is already known.

XSL•FO
RenderX

In addition, this is an analysis of patient-reported symptoms, which are not validated. The impact of user acquisition strategies is not known. However, the similarity in trends observed across two countries (and for respiratory symptoms, over the same period of time in 2019) adds weight to our findings.

Despite these limitations, the analysis presented in this paper has a number of unique strengths. First, the analysis was conducted using a large, existing data set that updates in real time and covers over 1400 unique self-reported symptoms since November 2016, allowing the monitoring of changes in trends over time. Second, the data is user-driven as a user self-reports their symptoms during an assessment on their own initiative. This allows for identification of changes that would not be detected in traditional studies focusing on specific and/or predefined areas. Third, the large number of symptoms presented in the results that are consistent with expectations, the observation of seasonal differences, and the observation that the results were similar in both Germany and the United Kingdom indicate that the Ada data is reliable. Fourth, the app covers a large range of symptoms that are not captured by traditional surveillance systems. For example, during the COVID-19 pandemic, we were able to compare the occurrence of *hyposmia* and *diminished sense of taste* during the outbreak with data from previous years, as those symptoms were already covered by the app before the pandemic brought them into focus. Although the clinical soundness of Ada's model at the level of individual diagnostics has already been demonstrated in other studies [54,55], the presented results build confidence that the data collected through the Ada app can also detect health changes in a population in real time.

Future Research

Future research can build on these strengths, focusing on the reasons for some of the detected changes and expanding the analysis to more countries. Of particular interest are countries from the Global South and low- and middle-income countries, given the comparative paucity of up-to-date health data in these countries and the differentials in the burden of disease. In addition, investigating changes in trends over time as the implementation of the COVID-19 measures changes according to the reality in different countries (ie, as individuals return to work) will offer meaningful insights into the effects of policy changes.

Conclusions

Our findings suggest that symptom assessment tools might have a role to play in improving understanding of the implications of public health measures. In this analysis, we have shown an innovative use of an existing data set that would enable policy makers to inform and monitor public health measures with a real-time, low-resource syndromic surveillance system that is relevant both during the COVID-19 pandemic and in the future.

Acknowledgments

Stephen Gilbert (Director of Clinical Evaluation), Vincent Jörres (Communications & Public Affairs Lead), Vanessa Lemarié (Director of Business Development & Program Owner of Rare Disease Initiative in Business Development), Nicola Vona (Director of Reasoning Technology in Product), and Paul Wicks (Scientific Advisor) reviewed the manuscript and provided feedback. This study was funded by Ada Health GmbH.

Authors' Contributions

AM and FB contributed to the planning of this study (study conception, protocol development), which was revised by CC and AG. AM and FB contributed to the data collection and analysis. AM, FB, CC, and AG contributed to the reporting (report writing). All authors contributed comments on drafts of the report. AM is the guarantor for this work. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

AM, FB, CC, and AG are employees or company directors of Ada Health GmbH.

Multimedia Appendix 1
Supplementary material.
[PDF File , 378 KB - mhealth_v8i10e21364_app1.pdf]

References

1. COVID-19 Map. Johns Hopkins Coronavirus Resource Center. URL: https://coronavirus.jhu.edu/map.html [accessed 2020-05-11]
2. EPPI-Mapper. URL: http://eppi.ioe.ac.uk/COVID19_MAP/covid_map_v10.html [accessed 2020-05-13]
3. Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. The Lancet Public Health 2020 May;5(5):e261-e270. [doi: 10.1016/s2468-2667(20)30073-6]
4. Flaxman S, Mishra S, Gandy A, Unwin HJT, Coupland H, Mellan TA, et al. Estimating the number of infections and the impact of nonpharmaceutical interventions on COVID-19 in 11 European countries. 2020 Mar 30. URL: https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-30-COVID19-Report-13.pdf [accessed 2020-10-02]

XSL•FO
RenderX

5. Douglas M, Katikireddi SV, Taulbut M, McKee M, McCartney G. Mitigating the wider health effects of covid-19 pandemic response. BMJ 2020 Apr 27;369:m1557. [doi: 10.1136/bmj.m1557] [Medline: 32341002]
6. United Nations Policy Brief: COVID-19 and the Need for Action on Mental Health Internet. 2020 May. URL: https://www.un.org/sites/un2.un.org/files/un_policy_brief-covid_and_mental_health_final.pdf [accessed 2020-10-02]
7. Torales J, O'Higgins M, Castaldelli-Maia JM, Ventriglio A. The outbreak of COVID-19 coronavirus and its impact on global mental health. Int J Soc Psychiatry 2020 Jun 31;66(4):317-320. [doi: 10.1177/0020764020915212] [Medline: 32233719]
8. Shigemura J, Ursano R, Morganstein J, Kurosawa M, Benedek D. Public responses to the novel 2019 coronavirus (2019-nCoV) in Japan: Mental health consequences and target populations. Psychiatry Clin Neurosci 2020 Apr;74(4):281-282 [FREE Full text] [doi: 10.1111/pcn.12988] [Medline: 32034840]
9. Moccia L, Janiri D, Pepe M, Dattoli L, Molinaro M, De Martin V, et al. Affective temperament, attachment style, and the psychological impact of the COVID-19 outbreak: an early report on the Italian general population. Brain Behav Immun 2020 Jul;87:75-79 [FREE Full text] [doi: 10.1016/j.bbi.2020.04.048] [Medline: 32325098]
10. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
11. Galea S, Merchant RM, Lurie N. The Mental Health Consequences of COVID-19 and Physical Distancing: The Need for Prevention and Early Intervention. JAMA Intern Med 2020 Jun 01;180(6):817-818. [doi: 10.1001/jamainternmed.2020.1562] [Medline: 32275292]
12. Beiu C, Mihai M, Popa L, Cima L, Popescu M. Frequent Hand Washing for COVID-19 Prevention Can Cause Hand Dermatitis: Management Tips. Cureus 2020 Apr 02;12(4):e7506 [FREE Full text] [doi: 10.7759/cureus.7506] [Medline: 32373409]
13. MacGibeny MA, Wassef C. Preventing adverse cutaneous reactions from amplified hygiene practices during the COVID-19 pandemic: how dermatologists can help through anticipatory guidance. Arch Dermatol Res 2020 May 09 [FREE Full text] [doi: 10.1007/s00403-020-02086-x] [Medline: 32388643]
14. The Lancet Oncology. COVID-19: global consequences for oncology. Lancet Oncol 2020 Apr;21(4):467 [FREE Full text] [doi: 10.1016/S1470-2045(20)30175-3] [Medline: 32240603]
15. Tan KK, Lau J. Cessation of cancer screening: An unseen cost of the COVID-19 pandemic? Eur J Surg Oncol 2020 May 11 [FREE Full text] [doi: 10.1016/j.ejso.2020.05.004] [Medline: 32409112]
16. Laupacis A. Working together to contain and manage COVID-19. CMAJ 2020 Mar 30;192(13):E340-E341. [doi: 10.1503/cmaj.200428] [Medline: 32184199]
17. Hogan AB, Jewell B, Sherrard-Smith E, Vesga JF, Watson OJ, Whittaker C, et al. Potential impact of the COVID-19 pandemic on HIV, tuberculosis, and malaria in low-income and middle-income countries: a modelling study. The Lancet Global Health 2020 Sep;8(9):e1132-e1141. [doi: 10.1016/S2214-109X(20)30288-6]
18. Riley T, Sully E, Ahmed Z, Biddlecom A. Estimates of the Potential Impact of the COVID-19 Pandemic on Sexual and Reproductive Health In Low- and Middle-Income Countries. Int Perspect Sex Reprod Health 2020 Apr 16;46:73-76. [doi: 10.1363/46e9020] [Medline: 32343244]
19. Measles Rubella Initiat. More than 117 million children at risk of missing out on measles vaccines, as COVID-19 surges. URL: https://measlesrubellainitiative.org/measles-news/more-than-117-million-children-at-risk-of-missing-out-on-measles-vaccines-as-covid-19-surges/ [accessed 2020-05-11]
20. Hoffman J. Vaccine Rates Drop Dangerously as Parents Avoid Doctor's Visits. The New York Times. 2020 Apr 23. URL: https://www.nytimes.com/2020/04/23/health/coronavirus-measles-vaccines.html [accessed 2020-05-11]
21. Emergency Department Syndromic Surveillance System: England. Public Health England Real-time Syndromic Surveillance Team. 2020 Apr. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/882234/EDSSSBulletin2020wk17.pdf [accessed 2020-10-02]
22. Longitudinal studies | COVID-MINDS. URL: https://www.covidminds.org/longitudinalstudies [accessed 2020-05-11]
23. Laranjo L, Dunn A, Tong H, Kocaballi A, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 01;25(9):1248-1258 [FREE Full text] [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
24. Lucas GM, Gratch J, King A, Morency L. It's only a computer: Virtual humans increase willingness to disclose. Computers in Human Behavior 2014 Aug;37:94-100. [doi: 10.1016/j.chb.2014.04.043]
25. Miner AS, Milstein A, Schueller S, Hegde R, Mangurian C, Linos E. Smartphone-Based Conversational Agents and Responses to Questions About Mental Health, Interpersonal Violence, and Physical Health. JAMA Intern Med 2016 May 01;176(5):619-625 [FREE Full text] [doi: 10.1001/jamainternmed.2016.0400] [Medline: 26974260]
26. Steinhubl SR, Topol EJ. Now we're talking: bringing a voice to digital medicine. The Lancet 2018 Aug;392(10148):627. [doi: 10.1016/s0140-6736(18)31803-8]
27. Nobles AL, Leas EC, Caputi TL, Zhu S, Strathdee SA, Ayers JW. Responses to addiction help-seeking from Alexa, Siri, Google Assistant, Cortana, and Bixby intelligent virtual assistants. NPJ Digit Med 2020 Jan 29;3(1):11 [FREE Full text] [doi: 10.1038/s41746-019-0215-9] [Medline: 32025572]

XSL•FO
RenderX

28. Kocaballi AB, Quiroz JC, Rezazadegan D, Berkovsky S, Magrabi F, Coiera E, et al. Responses of Conversational Agents to Health and Lifestyle Prompts: Investigation of Appropriateness and Presentation Structures. J Med Internet Res 2020 Feb 09;22(2):e15823 [FREE Full text] [doi: 10.2196/15823] [Medline: 32039810]
29. Bickmore TW, Trinh H, Olafsson S, O'Leary TK, Asadi R, Rickles NM, et al. Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant. J Med Internet Res 2018 Sep 04;20(9):e11510 [FREE Full text] [doi: 10.2196/11510] [Medline: 30181110]
30. Ronicke S, Hirsch MC, Türk E, Larionov K, Tientcheu D, Wagner AD. Can a decision support system accelerate rare disease diagnosis? Evaluating the potential impact of Ada DX in a retrospective study. Orphanet J Rare Dis 2019 Mar 21;14(1):69 [FREE Full text] [doi: 10.1186/s13023-019-1040-6] [Medline: 30898118]
31. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020 Jul 11;26(7):1037-1040. [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]
32. Denis F, Galmiche S, Dinh A, Fontanet A, Scherpereel A, Benezit F, et al. Epidemiological Observations on the Association Between Anosmia and COVID-19 Infection: Analysis of Data From a Self-Assessment Web Application. J Med Internet Res 2020 Jun 11;22(6):e19855 [FREE Full text] [doi: 10.2196/19855] [Medline: 32496206]
33. Drew DA, Nguyen LH, Steves CJ, Menni C, Freydin M, Varsavsky T, COPE Consortium. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. Science 2020 Jun 19;368(6497):1362-1367 [FREE Full text] [doi: 10.1126/science.abc0473] [Medline: 32371477]
34. Gilbert S, Mehl A, Baluch A, Cawley C, Challiner J, Fraser H, et al. Original research: How accurate are digital symptom assessment apps for suggesting conditions and urgency advice?: a clinical vignettes comparison to GPs. MedRxiv 2020 May 13:1. [doi: 10.1101/2020.05.07.20093872]
35. Wetter und Klima - Deutscher Wetterdienst - Klimaüberwachung - Deutschland - Zeitreihen und Trends. URL: https://www.dwd.de/DE/leistungen/zeitreihen/zeitreihen.html?nn=480164#buehneTop [accessed 2020-05-12]
36. ICD-10 Version 2019. URL: https://icd.who.int/browse10/2019/en [accessed 2020-05-12]
37. Gibney E. Whose coronavirus strategy worked best? Scientists hunt most effective policies. Nature 2020 May;581(7806):15-16 [FREE Full text] [doi: 10.1038/d41586-020-01248-1] [Medline: 32341558]
38. Papanicolas I, Mossialos E, Gundersen A, Woskie L, Jha AK. Performance of UK National Health Service compared with other high income countries: observational study. BMJ 2019 Nov 27;367:l6326. [doi: 10.1136/bmj.l6326] [Medline: 31776110]
39. Robert Koch Institut. Epidemiologisches Bulletin 16/2020. URL: https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2020/Ausgaben/16_20.html [accessed 2020-10-05]
40. Institute for Quality and Efficiency in Health Care (IQWiG). Allergies: Overview. 2020 Apr 23. URL: https://www.ncbi.nlm.nih.gov/books/NBK447112/ [accessed 2020-10-05]
41. Hebbelstrup Jensen B, Jokelainen P, Nielsen ACY, Franck KT, Rejkjær Holm D, Schønning K, et al. Children Attending Day Care Centers are a Year-round Reservoir of Gastrointestinal Viruses. Sci Rep 2019 Mar 01;9(1):3286 [FREE Full text] [doi: 10.1038/s41598-019-40077-9] [Medline: 30824842]
42. Bhattacharya S, Singh A, Marzo RR. "Airplane ear"-A neglected yet preventable problem. AIMS Public Health 2019;6(3):320-325. [doi: 10.3934/publichealth.2019.3.320] [Medline: 31637280]
43. Bénézit F, Le Turnier P, Declerck C, Paillé C, Revest M, Dubée V, RAN COVID Study Group. Utility of hyposmia and hypogeusia for the diagnosis of COVID-19. Lancet Infect Dis 2020 Sep;20(9):1014-1015 [FREE Full text] [doi: 10.1016/S1473-3099(20)30297-8] [Medline: 32304632]
44. Lao WP, Imam SA, Nguyen SA. Anosmia, hyposmia, and dysgeusia as indicators for positive SARS-CoV-2 infection. World J Otorhinolaryngol Head Neck Surg 2020 Apr 17 [FREE Full text] [doi: 10.1016/j.wjorl.2020.04.001] [Medline: 32313712]
45. Sommer IE, Bakker PR. What can psychiatrists learn from SARS and MERS outbreaks? Lancet Psychiatry 2020 Jul;7(7):565-566 [FREE Full text] [doi: 10.1016/S2215-0366(20)30219-4] [Medline: 32437680]
46. Van Bortel T, Basnayake A, Wurie F, Jambai M, Koroma AS, Muana AT, et al. Psychosocial effects of an Ebola outbreak at individual, community and international levels. Bull World Health Organ 2016 Mar 01;94(3):210-214 [FREE Full text] [doi: 10.2471/BLT.15.158543] [Medline: 26966332]
47. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020 Jun;7(6):547-560 [FREE Full text] [doi: 10.1016/S2215-0366(20)30168-1] [Medline: 32304649]
48. van Agteren J, Bartholomaeus J, Fassnacht D, Iasiello M, Ali K, Lo L, et al. Using Internet-Based Psychological Measurement to Capture the Deteriorating Community Mental Health Profile During COVID-19: Observational Study. JMIR Ment Health 2020 Jun 11;7(6):e20696 [FREE Full text] [doi: 10.2196/20696] [Medline: 32490845]
49. Covid Social Study. URL: https://www.covidsocialstudy.org/ [accessed 2020-05-20]
50. Jacobson NC, Lekkas D, Price G, Heinz MV, Song M, O'Malley AJ, et al. Flattening the Mental Health Curve: COVID-19 Stay-at-Home Orders Are Associated With Alterations in Mental Health Search Behavior in the United States. JMIR Ment Health 2020 Jun 01;7(6):e19347 [FREE Full text] [doi: 10.2196/19347] [Medline: 32459186]

XSL•FO
RenderX

51. Farr M. Australians are as happy as before coronavirus lockdown, Guardian Essential poll finds. The Guardian. 2020 Apr 13. URL: https://www.theguardian.com/australia-news/2020/apr/14/australians-are-as-happy-as-before-coronavirus-lockdown-guardian-essential-poll-finds [accessed 2020-05-21]
52. Vu V. Wie geht es Ihnen heute: Geht es Ihnen seit Corona besser? Zeit. 2020 Apr 29. URL: https://www.zeit.de/gesellschaft/zeitgeschehen/2020-04/wie-geht-es-ihnen-heute-warum-besser [accessed 2020-05-21]
53. Jarral F. The lockdown paradox: why some people?s anxiety is improving during the crisis. The Guardian. 2020 Apr 29. URL: https://www.theguardian.com/commentisfree/2020/apr/29/coronavirus-lockdown-anxiety-mental-health [accessed 2020-05-21]
54. Jungmann SM, Klan T, Kuhn S, Jungmann F. Accuracy of a Chatbot (Ada) in the Diagnosis of Mental Disorders: Comparative Case Study With Lay and Expert Users. JMIR Form Res 2019 Oct 29;3(4):e13863 [FREE Full text] [doi: 10.2196/13863] [Medline: 31663858]
55. Nateqi J, Lin S, Krobath H, Gruarin S, Lutz T, Dvorak T, et al. [From symptom to diagnosis-symptom checkers re-evaluated : Are symptom checkers finally sufficient and accurate to use? An update from the ENT perspective]. HNO 2019 May 16;67(5):334-342. [doi: 10.1007/s00106-019-0666-y] [Medline: 30993374]

Abbreviations

GDPR: General Data Protection Regulation
ICD-10: International Classification of Diseases, Version 2019

Edited by G Eysenbach; submitted 16.06.20; peer-reviewed by F Denis; comments to author 10.07.20; revised version received 15.07.20; accepted 15.09.20; published 09.10.20.

Please cite as:
Mehl A, Bergey F, Cawley C, Gilsdorf A
Syndromic Surveillance Insights from a Symptom Assessment App Before and During COVID-19 Measures in Germany and the United Kingdom: Results From Repeated Cross-Sectional Analyses
JMIR Mhealth Uhealth 2020;8(10):e21364
URL: http://mhealth.jmir.org/2020/10/e21364/
doi:10.2196/21364
PMID:32997640

©Alicia Mehl, Francois Bergey, Caoimhe Cawley, Andreas Gilsdorf. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 09.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Smartphone-Enabled, Telehealth-Based Family Conferences in Palliative Care During the COVID-19 Pandemic: Pilot Observational Study

Yu-Rui Wu[1], MD; Tzu-Jung Chou[2], MD; Yi-Jen Wang[2], BP; Jaw-Shiun Tsai[2], MD, PhD; Shao-Yi Cheng[2], MD, MSc, DrPH; Chien-An Yao[2], MD, MSc, PhD; Jen-Kuei Peng[2], MD, MSc, PhD; Wen-Yu Hu[3], RNC, PhD; Tai-Yuan Chiu[2], MD, MSc; Hsien-Liang Huang[2], MD, MSc

[1]Department of Family Medicine, Taitung Christian Hospital, Taitung, Taiwan

[2]Department of Family Medicine, National Taiwan University Hospital, Taipei City, Taiwan

[3]Department of Nursing, National Taiwan University Hospital, Taipei City, Taiwan

Corresponding Author:
Hsien-Liang Huang, MD, MSc
Department of Family Medicine
National Taiwan University Hospital
7 Chung-Shan South Road
Taipei City, 100
Taiwan
Phone: 886 2 23123456 ext 66832
Email: tennishuang@gmail.com

Abstract

Background: In the palliative care setting, infection control measures implemented due to COVID-19 have become barriers to end-of-life care discussions (eg, discharge planning and withdrawal of life-sustaining treatments) between patients, their families, and multidisciplinary medical teams. Strict restrictions in terms of visiting hours and the number of visitors have made it difficult to arrange in-person family conferences. Phone-based telehealth consultations may be a solution, but the lack of nonverbal cues may diminish the clinician-patient relationship. In this context, video-based, smartphone-enabled family conferences have become important.

Objective: We aimed to establish a smartphone-enabled telehealth model for palliative care family conferences. Our model integrates principles from the concept of shared decision making (SDM) and the value, acknowledge, listen, understand, and elicit (VALUE) approach.

Methods: Family conferences comprised three phases designed according to telehealth implementation guidelines—the previsit, during-visit, and postvisit phases. We incorporated the following SDM elements into the model: “team talk,” “option talk,” and “decision talk.” The model has been implemented at a national cancer treatment center in Taiwan since February 2020.

Results: From February to April 2020, 14 telehealth family conferences in the palliative care unit were analyzed. The patients’ mean age was 73 (SD 10.1) years; 6 out of 14 patients (43%) were female and 12 (86%) were married. The primary caregiver joining the conference virtually comprised mostly of spouses and children (n=10, 71%). The majority of participants were terminally ill patients with cancer (n=13, 93%), with the exception of 1 patient with stroke. Consensus on care goals related to discharge planning and withdrawal of life-sustaining treatments was reached in 93% (n=13) of cases during the family conferences. In total, 5 families rated the family conferences as *good* or *very good* (36%), whereas 9 were *neutral* (64%).

Conclusions: Smartphone-enabled telehealth for palliative care family conferences with SDM and VALUE integration demonstrated high satisfaction for families. In most cases, it was effective in reaching consensus on care decisions. The model may be applied to other countries to promote quality in end-of-life care in the midst of the COVID-19 pandemic.

(JMIR Mhealth Uhealth 2020;8(10):e22069) doi:10.2196/22069

KEYWORDS

smartphone; mobile phone; telehealth; family conference; shared decision making; COVID-19; palliative care; end-of-life care

XSL•FO
RenderX

Introduction

Face-to-face communication is indispensable in palliative care settings, but the COVID-19 outbreak may strain this well-established way of delivering end-of-life family conferences. This emerging infectious disease has posed an unprecedented threat, devastating the economy and medical systems of countries worldwide [1]. Physicians and health care facilities are thus confronted with various challenges, including limited medical resources and capacities as well as corresponding restrictions. Restrictions in hospital visits and visiting hours, put in place to reduce the transmission of COVID-19, have become obstacles to arranging family conferences [2,3]. Fewer family members have access to hospitals for meetings with medical teams compared to prepandemic times. This has decreased chances for communication and interaction between clinicians and patients and their families.

Health care professionals face many challenges and ethical dilemmas when caring for terminally ill patients; finding the most appropriate management solution requires communication among patients, caregivers, and multidisciplinary medical teams [4,5]. Discharge planning and withdrawal of life-sustaining treatments are top-ranking ethical dilemmas in palliative care, especially among inpatients [6,7]. Decisions related to place of care at the end of life have become increasingly complicated during the COVID-19 pandemic due to worries about the disease's high infection rate; however, families may encounter difficulties when caring for patients at home. In end-of-life care, it is our responsibility to realize the essence of shared decision making (SDM), which integrates the patient's preferences with the best-known evidence [8,9]. Furthermore, exploring the family's preferred role and desire is one of the key components of high-quality SDM, and thus incorporating the concept of SDM into family conferences is important [10]. Therefore, family conferences are an appropriate venue for engagement between all parties involved; here, patients and families can express their worries, receive integrated information from the medical team, and achieve concordance of caring goals.

The essence and main principle of palliative care is to relieve the physical and psycho-spiritual sufferings of the patient and their family, and to console their feelings. Previous studies have shown that not only terminally ill patients but their families as well suffer from physical and psychological distress, leading to subsequent comorbidity and increasing mortality [11,12]. On the other hand, the five principles of VALUE (value, acknowledge, listen, understand, and elicit) were proposed to improve physician-family communication, originally in intensive care units (ICUs), and has been shown to significantly decrease emotional disorders, such as posttraumatic stress disorder, in families even 3 months after a patient's death [13,14]. However, in the clinician-patient relationship, nonverbal interactions and appearance cues play fundamental roles in successful communication [15,16]. The patient's nonverbal reactions may be helpful to the physician for diagnosis and treatment decisions, and the clinician's nonverbal behaviors are related to patient satisfaction [17]. The lack of nonverbal communication is one of the major drawbacks of telehealth and diminishes the clinician-patient relationship. Under the circumstances, the optimal way to adhere to infectious disease prevention measures and promote physician-patient relationship might be telehealth-based family conferences with the integration of appropriate communication strategies such as VALUE.

In theory, implementing the VALUE approach for family conferences on the basis of SDM should solidify communication between clinicians and families. In this study, we aimed to establish a model of smartphone-enabled telehealth for palliative care family conferences with SDM and VALUE integration in order to (1) increase access to communication with the clinician under the visitor restrictions imposed by the pandemic, (2) reach consensus on care goals, and (3) achieve patient and family satisfaction with telehealth-based conferences.

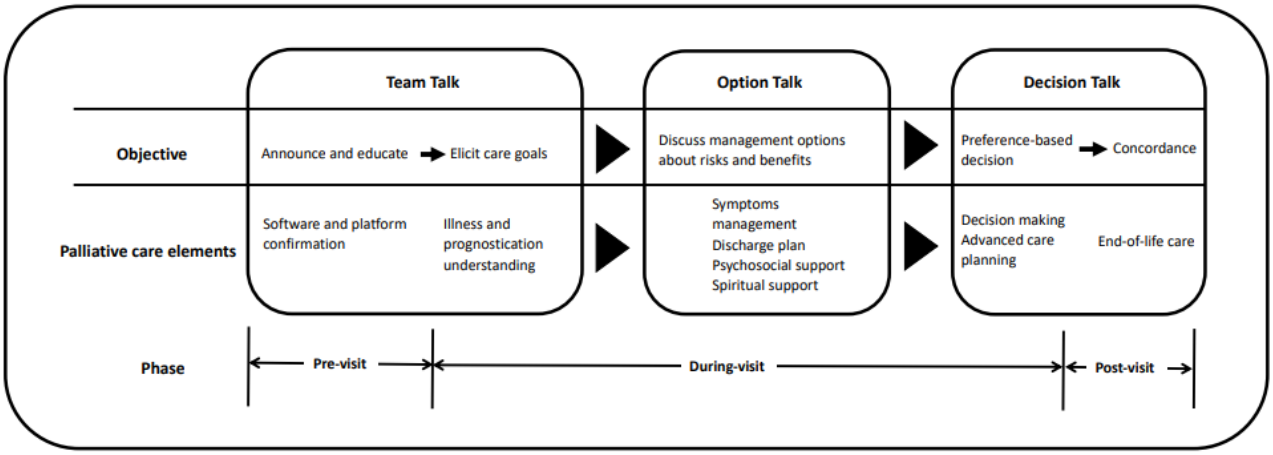
Methods

A Model for Telehealth-Based Family Conferences in Palliative Care

A model for telehealth-based family conferences in palliative care was developed by integrating SDM principles like "team talk," "option talk," and "decision talk" into the interview structure (Figure 1). The telehealth workflow was implemented into family conferences with previsit, during-visit, and postvisit phases based on the American Medical Association's Telehealth Implementation Playbook [18]. The availability of telehealth devices and knowledge of the software platform were important requirements for implementation and participation.

XSL•FO
RenderX

Figure 1. A model for telehealth-based family conferences in palliative care.



The objective of team talk was to assemble the patient, their family, and the multidisciplinary medical team in order to build rapport. In our model, team talk included the previsit phase and the initial part of the during-visit phase. The previsit phase announced the date of the family meeting. Additionally, education on how to use the software and platform on a smartphone was provided to the patient, their family, and the multidisciplinary medical team during this phase. The during-visit phase is marked by the beginning of the telehealth family conference. Moreover, we integrated the palliative care elements of illness and prognostication understanding into the initial stage of this phase. Care goal choices were provided, and the use of the phrase "Shall we describe the options in more detail?" by the physician marked the end of team talk.

Option talk was designed to discuss treatment options, and their associated risks and benefits, with patients and their families. Elements of palliative care such as symptom management and discharge planning were included in option talk. Furthermore, to provide psychosocial and spiritual support, the VALUE approach was emphasized during the reflection and deliberation portion of the telehealth conference. During discussions on the risks and benefits of management options for end-of-life care, such as withdrawal of life-sustaining treatments and discharge planning, possible psychological symptoms in patients like demoralization and distress about dying would induce negative emotions during the conference [19]. The VALUE approach, specifically "value and appreciate what the family member said," "acknowledge the family member's emotions," and "understand who the patient was as a person," is important in communicating the options.

The aim of decision talk was to reach preference-based decisions on care goals. Advanced care planning during end-of-life care like discharge planning and withdrawal of life-sustaining treatments were the main goals in the model. Using the VALUE approach, we elicited informed preferences through the element "understand the patient as a person" to reach the decision. An important step during this phase is contacting every family member on different devices and interfaces to ensure they are ready to make decisions without network connection barriers. The postvisit phase was also included in the decision talk step, and patient and family experiences toward the conference was evaluated.

Implementation of the Telehealth Family Conference Model

The telehealth model was implemented in the palliative care unit of the National Taiwan University Hospital, a national cancer treatment center in Taiwan. The hospital set restrictions on family visits and allowed only one caregiver per patient in the inpatient setting during the COVID-19 pandemic. Telehealth-based family conferences were arranged in order to involve additional family members to communicate critical care decisions. The flowchart shown in Figure 1 was implemented after February 2020 in the palliative care unit. Patients admitted to the palliative care ward were eligible to enroll if they were more than 20 years old and required a meeting concerning care goals with the multidisciplinary palliative care team. Families joining the telehealth conference needed to be capable of using a smartphone.

Before the family conference, the date of the meeting was confirmed with the patient and their family. Discussion on which smartphone software or app to use was also carried out prior to the conference. The platform was chosen according to family members' knowledge and convenience. The interview guide for a video-based family conference was distributed to all medical professionals joining the conference in advance. During the option talk phase, the VALUE approach was integrated into the semistructured meeting. In the decision talk phase, consensus on care goals was achieved, and advanced directives on end-of-life care were documented.

Data Collection and Analysis

The family conference was arranged when a care goal needed to be communicated to the patient, their family, and the medical teams during the study period. The family conference was summarized and uploaded to the electronic health record system of the hospital. Conversations between the patient, their family, and the multidisciplinary team, which were based on the VALUE approach, were documented. Consensus on care goals was achieved if a decision was made regarding discharge planning, including community-based palliative care or continued hospitalization, and withdrawal of life-sustaining treatments, such as antibiotics or blood transfusion, after the telehealth family conference. Feedback was then evaluated

XSL•FO
RenderX

through three questions answered by the patient and their family members together after the conference:

1. Is this your first telehealth family conference?
2. Do you want to use telehealth conferencing again?
3. Do you prefer to talk with your physician face to face?

Family satisfaction toward the meeting was graded on a 5-point scale as follows: 5=very good; 4=good; 3=neutral; 2=bad; and 1=very bad. The analysis was approved by the National Taiwan University Hospital Research Ethics Committee (202004113RINC).

Results

From February to April 2020 during the COVID-19 pandemic, 14 telehealth family conferences in the palliative care unit were analyzed. The characteristics of the patients are presented in Table 1. The patients' mean age was 73 (SD 10.1) years, and 6 (43%) patients were female. Among the 14 patients, 12 (86%) were married, and the primary caregiver joining the conference was most frequently a spouse or children (n=10, 71%). The majority of patients (n=13, 93%) were diagnosed with cancer and were terminally ill; 1 (7%) patient was diagnosed with stroke.

Table 1 also shows the number of family members joining the conference using telehealth devices in addition to the one member permitted at bedside. There were 9 (64%) family conferences with more than 2 family members using telehealth software, and 5 (36%) conferences with one family member joining via telehealth.

Two main themes of the family conference include discharge planning and withdrawal of life-sustaining treatments. Only one family meeting to discuss discharge planning did not reach consensus after the conference whereas 8 (89%) did. In terms of conferences on withdrawal of life-sustaining treatment, 100% (n=5) of families reached concordance. In total, consensus was reached in 13 out of 14 conferences (94%).

There were 12 families (90%) using video conferencing for the first time in the health care context, and 10 families (71%) were willing to use video conferencing again for family meetings. However, 7 families (50%) preferred to communicate with medical teams face to face. Level of satisfaction experienced by the families while using video conferencing is demonstrated in Table 2. In total, 5 out of 14 families rated the family meeting as *good* or *very good* (36%), 9 families provided *neutral* feedback (64%), and there was no negative feedback.

XSL•FO
RenderX

Table 1. Demographic characteristics of patients (N=14).

Patient	Patients
Gender, n (%)	
Male	8 (57)
Female	6 (43)
Age (years), mean (SD)	73 (10.1)
Under 40, n (%)	0 (0)
41-50, n (%)	1 (7)
51-60, n (%)	1 (7)
61-70, n (%)	3 (22)
71-80, n (%)	3 (22)
Over 80, n (%)	6 (43)
Marital status, n (%)	
Married	12 (86)
Single	1 (7)
Separated or divorced	0 (0)
Widowed	1 (7)
Education, n (%)	
Illiterate	2 (14)
Elementary school	6 (43)
Junior high school	3 (21)
High school	0 (0)
Bachelor	2 (14)
Master or PhD	0 (0)
Unknown	1 (7)
Primary caregiver, n (%)	
Spouse	5 (36)
Daughter or son	5 (36)
Sibling	2 (14)
Other	2 (14)
Number of family members capable of using telehealth, n (%)	
1	5 (36)
2	4 (29)
3	2 (14)
4	2 (14)
5	1 (7)
Diagnosis, n (%)	
Cancer	13 (93)
Stroke	1 (7)

XSL•FO
RenderX

Table 2. Family attitudes and satisfaction toward telehealth use in palliative care family conferences.

Variable	Participants, n (%)
Attitudes	
Is this your first time attending a video-based family conference?	
Yes	12 (90)
No	2 (10)
Do you want to use video conferencing again?	
Yes	10 (70)
No	4 (30)
Do you prefer to talk with your doctor face to face?	
Yes	7 (50)
No	7 (80)
Family satisfaction	
How do you feel about this telehealth conference compared to face-to-face conferences?	
Very good	2 (14)
Good	3 (22)
Neutral	9 (64)
Bad	0 (0)
Very bad	0 (0)

The 14 family conferences were divided into two groups that were labeled as the neutral group (n=9, 64%) and the satisfied group (rating: *good* and *very good*; n=5, 36%). Categorical variables in Table 3 demonstrate possible factors that relate to respondents' satisfaction toward telehealth family conferences. A chi-square test did not reveal any statistically significant relationships between the two groups. Further logistic regression analysis also showed no statistically significant associated variables.

XSL•FO
RenderX

Table 3. Univariate analysis (χ^2) comparing the satisfied group (rating: good and very good) to the neutral group.

Variable	Neutral (n=9), n (%)	Satisfied (n=5), n (%)	χ^2	*P* value
Age			0.280	.60
≤65 years	3 (75.0)	1 (25.0)		
>65 years	6 (60.0)	4 (40.0)		
Gender			4.381	.06
Male	7 (87.5)	1 (12.5)		
Female	2 (33.3)	4 (66.7)		
Education			4.563	.21
Less than elementary school	6 (75.0)	2 (25.0)		
High school	2 (66.7)	1 (33.3)		
Bachelor or higher	0 (0.0)	2 (100.0)		
Unknown	1 (100.0)	0 (0.0)		
Marital status			2.385	.30
Married	8 (66.7)	4 (33.3)		
Single	1 (100.0)	0 (0.0)		
Widowed	0 (0.0)	1 (100.0)		
Primary caregiver			4.853	.18
Spouse	4 (80.0)	1 (20.0)		
Daughter or son	0 (0.0)	2 (100.0)		
Sibling	4 (80.0)	1 (20.0)		
Other	1 (50.0)	1 (50.0)		
Number of family members using video conferencing			1.998	.16
1	2 (40.0)	3 (60.0)		
>1	7 (77.8)	2 (22.2)		
Diagnosis			2.385	.30
Cancer	8 (66.7)	4 (33.3)		
Stroke	0 (0.0)	1 (100.0)		

Patients and family members also provided comments about the telehealth model. Some positive comments were as follows:

> *Telehealth is better than a face-to-face meeting since I don't need to go to the hospital. I am really anxious about the current COVID-19 threats.*

> *Thank you for your arrangement on this kind of smartphone-enabled communication. I could join the meeting from my office.*

There were also some suggestions such as:

> *My father is a terminal cancer patient! The hospital should relax the infection control measures and let more family members accompany terminal patients in the emotional moments approaching the end of life!*

> *A big screen like that of a computer would be better than a smartphone-enabled model.*

> *Frequent lags, and the network transmission is not smooth! Interruptions occurred several times during physicians' explanations on the prognosis.*

Discussion

As demonstrated by this study, the telehealth model achieved the aim of enabling more family members to join family conferences under visitor restrictions due to the COVID-19 pandemic. In addition, consensus was achieved on care goals through telehealth communication similar to face-to-face meetings, and high satisfaction toward smartphone-enabled telehealth was seen.

Increases in Communication Through the Telehealth-Based Family Conference During the COVID-19 Pandemic

Under the COVID-19 pandemic, nearly all hospitals implemented visiting restrictions to reduce personal protective equipment usage due to limited resources and the risk of exposure or nosocomial infection; this included restrictions on

XSL•FO
RenderX

visiting hours and the number of visitors [20]. In Taiwan, only one caregiver per patient is permitted to enter the hospital, which increased the difficulty of arranging family conferences. In addition, those who lived abroad were deprived of the opportunity to participate in such events; this also applied to those under home quarantine or home isolation, or following a self-health procedure. Telehealth-based family conferences ameliorated this situation by virtually assembling family members to discuss medical requirements and relieve emotional burdens. With the technology of video software on smartphones, we could now establish clinician-family relationships and facilitate efficient communication with the family even if they are at their workplace or unable to attend in person. Smartphones are so widely used by the general population that it hardly causes any inconvenience to the family and the health care team when used as a venue for conversation. The results of this pilot observational study demonstrated that the number of family members joining the conference increased with the aid of telehealth.

Integration of SDM in the Telehealth Model to Reach Consensus on End-of-Life Care Goals

Previous studies have shown that family conferences played an indispensable role in facilitating communication with the patient and the family, and further optimized the provision of holistic, goal-concordant care [14,21,22]. However, there has been sparse evidence on the effectiveness of telehealth approaches. In addition, terminally ill patients were often confronted with multiple challenges to conquer and decisions to make, including withdrawal of life-sustaining treatments, prognostication awareness, treatment options, and anticipatory bereavement. Appropriate measurements that comply with SDM can help not only patients but also caregivers in further understanding palliative care, and thus enabling smooth communication with medical professionals. Therefore, our study took advantage of the technology trend and incorporated SDM into the core content of our family conferences, with the aim of helping patients and their families to make optimal decisions for end-of-life management. The results of our study showed that concordance on care goals was high even for difficult decisions like discharge planning and withdrawal of life-sustaining treatments. Thus, this model is feasible for adoption in the palliative care ward during situations like the COVID-19 pandemic, and has the potential to influence regulation on health insurance reimbursement.

Integration of the VALUE Approach in Palliative Care Family Conferences to Help Achieve Patient and Family Satisfaction With Telehealth

Compared to the traditional face-to-face model, one of the major concerns of telehealth is the physician-patient relationship. There has been some debate that the virtual setting of the venue might make participants miss visual clues, hence, diminishing the quality and goals of the conference, not to mention the lack of adequate physical contact such as a handshake or an assuring patting action as a vital step to establish rapport [23]. Previous studies have revealed that about 7% of emotional communication takes place verbally, while 22% was expressed by tone of voice and 55% by hand gestures, gaze, and eye contact [22]. Therefore, communicating clinical conditions by talking on the phone to the family seems insufficient in building and maintaining rapport with them due to a lack of nonverbal behaviors [24]. There have been studies showing that patients and physicians could bridge the gap in communication and achieve goals effectively through video consultation; however, application to palliative care settings is lacking, where there is much emphasis on face-to-face conversations [25]. Smartphones with relevant apps have played an important role during the COVID-19 pandemic due to social distancing and strict infection control measures [26]. Multidisciplinary medical teams should practice communicating in the telehealth family medicine context using smartphones during the current stringent period. In our model, we implemented the VALUE approach in family meetings, which enabled and encouraged the family to speak more, and made the discussion more rewarding for both sides. Additionally, the family's emotional cues including (positive) smiles or (negative) frowning were visible on the screen during the meeting. Further studies are needed to explore the influence of these emotional cues on the results of the family conference. The survey on the postvisit phase revealed neutral attitudes or satisfaction toward the current model, and the results indicated that the VALUE approach was suitable for telehealth-based family conferences in palliative care under the current visitor restrictions in hospitals.

Limitations

Here, we acknowledge a number of limitations pertaining to the study. First, the family conference process was not recorded, and the contents were summarized mainly by medical professionals. Therefore, there may be observational bias in the study. Secondly, the wireless network's poor performance during some conferences interfered with communication between physicians and family members; this may influence rapport due to waiting times during these disruptions. The restricted availability of the telehealth software also created obstacles to arranging video-based family meetings. Some families preferred to come to the hospital instead since they were reluctant to install the software or learn how to use it. Lastly, the lack of physical contact was another drawback of telehealth meetings, especially in palliative care settings. Among the various goals of arranging family conferences in the model, discerning when and how to facilitate the family in bidding farewell remained the most difficult topic, which required sensitivity not only through dialogue or facial expressions but also through physical gestures (ie, touching).

Conclusions

During the COVID-19 pandemic, health care professionals must adhere to the restrictions implemented for transmission prevention. A telehealth model for family conferences in palliative care with SDM and VALUE integration demonstrated high satisfaction in family members and was effective in reaching consensus about care decisions. The model may be applied to other countries to promote quality in end-of-life care in the era of COVID-19.

XSL•FO
RenderX

Acknowledgments

We thank the multidisciplinary team in our palliative care ward for offering comprehensive care to patients. Additionally, we would like to express our greatest gratitude to all medical professionals around the world dedicated to combating COVID-19 on the frontline.

Conflicts of Interest

None declared.

References

1. Tanne JH, Hayasaki E, Zastrow M, Pulla P, Smith P, Rada AG. Covid-19: how doctors and healthcare systems are tackling coronavirus worldwide. BMJ 2020 Mar 18;368:m1090. [doi: 10.1136/bmj.m1090] [Medline: 32188598]
2. Calton B, Abedini N, Fratkin M. Telemedicine in the Time of Coronavirus. J Pain Symptom Manage 2020 Jul 31;60(1):e12-e14 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.03.019] [Medline: 32240756]
3. Hsu Y, Liu Y, Lin M, Lee H, Chen T, Chou L, et al. Visiting Policies of Hospice Wards during the COVID-19 Pandemic: An Environmental Scan in Taiwan. Int J Environ Res Public Health 2020 Apr 21;17(8):e [FREE Full text] [doi: 10.3390/ijerph17082857] [Medline: 32326274]
4. Back AL. Patient-Clinician Communication Issues in Palliative Care for Patients With Advanced Cancer. J Clin Oncol 2020 Mar 20;38(9):866-876. [doi: 10.1200/JCO.19.00128] [Medline: 32023153]
5. Hui D, Bruera E. Models of Palliative Care Delivery for Patients With Cancer. J Clin Oncol 2020 Mar 20;38(9):852-865. [doi: 10.1200/JCO.18.02123] [Medline: 32023157]
6. Chiu T, Hu W, Huang H, Yao C, Chen C. Prevailing ethical dilemmas in terminal care for patients with cancer in Taiwan. J Clin Oncol 2009 Aug 20;27(24):3964-3968. [doi: 10.1200/JCO.2008.21.4643] [Medline: 19470919]
7. Huang H, Yao C, Hu W, Cheng S, Hwang S, Chen C, et al. Prevailing Ethical Dilemmas Encountered by Physicians in Terminal Cancer Care Changed After the Enactment of the Natural Death Act: 15 Years' Follow-up Survey. J Pain Symptom Manage 2018 Mar;55(3):843-850. [doi: 10.1016/j.jpainsymman.2017.11.033] [Medline: 29221846]
8. Belanger E. Shared decision-making in palliative care: Research priorities to align care with patients' values. Palliat Med 2017 Jul 16;31(7):585-586. [doi: 10.1177/0269216317713864] [Medline: 28618900]
9. Huang H, Tsai J, Yao C, Cheng S, Hu W, Chiu T. Shared decision making with oncologists and palliative care specialists effectively increases the documentation of the preferences for do not resuscitate and artificial nutrition and hydration in patients with advanced cancer: a model testing study. BMC Palliat Care 2020 Feb 04;19(1):17 [FREE Full text] [doi: 10.1186/s12904-020-0521-7] [Medline: 32019540]
10. Curtis JR, White DB. Practical guidance for evidence-based ICU family conferences. Chest 2008 Oct;134(4):835-843 [FREE Full text] [doi: 10.1378/chest.08-0235] [Medline: 18842916]
11. Rokach A, Matalon R, Safarov A, Bercovitch M. The dying, those who care for them, and how they cope with loneliness. Am J Hosp Palliat Care 2007 Sep 05;24(5):399-407. [doi: 10.1177/1049909107305652] [Medline: 17998407]
12. Song JI, Shin DW, Choi JY, Kang J, Baik YJ, Mo H, et al. Quality of life and mental health in family caregivers of patients with terminal cancer. Support Care Cancer 2011 Oct 10;19(10):1519-1526. [doi: 10.1007/s00520-010-0977-8] [Medline: 21479527]
13. Azoulay E, Chevret S, Leleu G, Pochard F, Barboteu M, Adrie C, et al. Half the families of intensive care unit patients experience inadequate communication with physicians. Crit Care Med 2000 Aug;28(8):3044-3049. [doi: 10.1097/00003246-200008000-00061] [Medline: 10966293]
14. Lautrette A, Darmon M, Megarbane B, Joly LM, Chevret S, Adrie C, et al. A Communication Strategy and Brochure for Relatives of Patients Dying in the ICU. N Engl J Med 2007 Feb;356(5):469-478. [doi: 10.1056/nejmoa063446]
15. Hjelm NM. Benefits and drawbacks of telemedicine. J Telemed Telecare 2005 Jan;11(2):60-70. [doi: 10.1258/1357633053499886] [Medline: 15829049]
16. Mast MS. On the importance of nonverbal communication in the physician-patient interaction. Patient Educ Couns 2007 Aug;67(3):315-318. [doi: 10.1016/j.pec.2007.03.005] [Medline: 17478072]
17. Mast MS, Hall JA, Köckner C, Choi E. Physician gender affects how physician nonverbal behavior is related to patient satisfaction. Med Care 2008 Dec;46(12):1212-1218. [doi: 10.1097/MLR.0b013e31817e1877] [Medline: 19300310]
18. Telehealth Implementation Playbook. American Medical Association. 2020. URL: https://www.ama-assn.org/system/files/2020-04/ama-telehealth-playbook.pdf [accessed 2020-08-20]
19. Rodin G, An E, Shnall J, Malfitano C. Psychological Interventions for Patients With Advanced Disease: Implications for Oncology and Palliative Care. J Clin Oncol 2020 Mar 20;38(9):885-904. [doi: 10.1200/JCO.19.00058] [Medline: 32023159]
20. Mehta AK, Smith TJ. Palliative Care for Patients With Cancer in the COVID-19 Era. JAMA Oncol 2020 May 07. [doi: 10.1001/jamaoncol.2020.1938] [Medline: 32379855]

XSL•FO
RenderX

21. Rhondali W, Dev R, Barbaret C, Chirac A, Font-Truchet C, Vallet F, et al. Family conferences in palliative care: a survey of health care providers in France. J Pain Symptom Manage 2014 Dec;48(6):1117-1124 [FREE Full text] [doi: 10.1016/j.jpainsymman.2014.03.007] [Medline: 24780185]
22. Sabesan S, Allen D, Caldwell P, Loh PK, Mozer R, Komesaroff PA, Royal Australasian College of Physicians Telehealth Working Group. Practical aspects of telehealth: doctor-patient relationship and communication. Intern Med J 2014 Jan;44(1):101-103. [doi: 10.1111/imj.12323] [Medline: 24450527]
23. Hudson P, Quinn K, O'Hanlon B, Aranda S. Family meetings in palliative care: Multidisciplinary clinical practice guidelines. BMC Palliat Care 2008 Aug 19;7:12 [FREE Full text] [doi: 10.1186/1472-684X-7-12] [Medline: 18710576]
24. Schmidt KL, Gentry A, Monin JK, Courtney KL. Demonstration of facial communication of emotion through telehospice videophone contact. Telemed J E Health 2011 Jun;17(5):399-401 [FREE Full text] [doi: 10.1089/tmj.2010.0190] [Medline: 21492032]
25. Menon P, Rabinowitz T, Prelock P, Rose G, Stapleton R. Clinician Perceptions of Telemedicine for Conducting Family Conferences Prior to Transfer to a Tertiary Care Center Intensive Care Unit. Journal of the International Society for Telemedicine and EHealth 2016;4(e20):1-8 [FREE Full text]
26. Meng X, Dai Z, Hang C, Wang Y. Smartphone-enabled wireless otoscope-assisted online telemedicine during the COVID-19 outbreak. Am J Otolaryngol 2020 May;41(3):102476 [FREE Full text] [doi: 10.1016/j.amjoto.2020.102476] [Medline: 32305252]

Abbreviations

ICU: intensive care unit
SDM: shared decision making
VALUE: value, acknowledge, listen, understand, and elicit

Edited by G Eysenbach, L Buis; submitted 10.07.20; peer-reviewed by L Zhou, V Watzlaf, M Ghozali; comments to author 11.09.20; revised version received 21.09.20; accepted 05.10.20; published 28.10.20.

Please cite as:
Wu YR, Chou TJ, Wang YJ, Tsai JS, Cheng SY, Yao CA, Peng JK, Hu WY, Chiu TY, Huang HL
Smartphone-Enabled, Telehealth-Based Family Conferences in Palliative Care During the COVID-19 Pandemic: Pilot Observational Study
JMIR Mhealth Uhealth 2020;8(10):e22069
URL: http://mhealth.jmir.org/2020/10/e22069/
doi:10.2196/22069
PMID:33021483

©Yu-Rui Wu, Tzu-Jung Chou, Yi-Jen Wang, Jaw-Shiun Tsai, Shao-Yi Cheng, Chien-An Yao, Jen-Kuei Peng, Wen-Yu Hu, Tai-Yuan Chiu, Hsien-Liang Huang. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 28.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Performance of Digital Contact Tracing Tools for COVID-19 Response in Singapore: Cross-Sectional Study

Zhilian Huang[1], MPH, PhD; Huiling Guo[1], MPH; Yee-Mun Lee[2,3], FRCSEd, MPH; Eu Chin Ho[4], MBChB, MMedEdu, FRCS(ORL); Hou Ang[5], BCh, MB, MRCS; Angela Chow[1,3], MBBS, MPH, MS, PhD

[1]Department of Clinical Epidemiology, Office of Clinical Epidemiology, Analytics, and Knowledge, Tan Tock Seng Hospital, Singapore, Singapore

[2]Department of Urology, Tan Tock Seng Hospital, Singapore, Singapore

[3]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[4]Department of Ear, Nose, and Throat, Tan Tock Seng Hospital, Singapore, Singapore

[5]Department of Emergency Medicine, Tan Tock Seng Hospital, Singapore, Singapore

Corresponding Author:
Angela Chow, MBBS, MPH, MS, PhD
Department of Clinical Epidemiology
Office of Clinical Epidemiology, Analytics, and Knowledge
Tan Tock Seng Hospital
11 Jalan Tan Tock Seng
Singapore, 308433
Singapore
Phone: 65 63577477
Email: angela_chow@ttsh.com.sg

Abstract

Background: Effective contact tracing is labor intensive and time sensitive during the COVID-19 pandemic, but also essential in the absence of effective treatment and vaccines. Singapore launched the first Bluetooth-based contact tracing app—TraceTogether—in March 2020 to augment Singapore's contact tracing capabilities.

Objective: This study aims to compare the performance of the contact tracing app—TraceTogether—with that of a wearable tag-based real-time locating system (RTLS) and to validate them against the electronic medical records at the National Centre for Infectious Diseases (NCID), the national referral center for COVID-19 screening.

Methods: All patients and physicians in the NCID screening center were issued RTLS tags (CADI Scientific) for contact tracing. In total, 18 physicians were deployed to the NCID screening center from May 10 to May 20, 2020. The physicians activated the TraceTogether app (version 1.6; GovTech) on their smartphones during shifts and urged their patients to use the app. We compared patient contacts identified by TraceTogether and those identified by RTLS tags within the NCID vicinity during physicians' 10-day posting. We also validated both digital contact tracing tools by verifying the physician-patient contacts with the electronic medical records of 156 patients who attended the NCID screening center over a 24-hour time frame within the study period.

Results: RTLS tags had a high sensitivity of 95.3% for detecting patient contacts identified either by the system or TraceTogether while TraceTogether had an overall sensitivity of 6.5% and performed significantly better on Android phones than iPhones (Android: 9.7%, iPhone: 2.7%; $P<.001$). When validated against the electronic medical records, RTLS tags had a sensitivity of 96.9% and specificity of 83.1%, while TraceTogether only detected 2 patient contacts with physicians who did not attend to them.

Conclusions: TraceTogether had a much lower sensitivity than RTLS tags for identifying patient contacts in a clinical setting. Although the tag-based RTLS performed well for contact tracing in a clinical setting, its implementation in the community would be more challenging than TraceTogether. Given the uncertainty of the adoption and capabilities of contact tracing apps, policy makers should be cautioned against overreliance on such apps for contact tracing. Nonetheless, leveraging technology to augment conventional manual contact tracing is a necessary move for returning some normalcy to life during the long haul of the COVID-19 pandemic.

(JMIR Mhealth Uhealth 2020;8(10):e23148) doi:10.2196/23148

XSL•FO
RenderX

KEYWORDS

infectious disease; real-time locating systems; electronic medical records; COVID-19; contact tracing; public health

Introduction

The fight against COVID-19 is expected to be a long haul due to its high transmissibility, despite its low virulence rates [1]. There have been predictions of potential second waves of SARS-CoV-2 infections, and several countries have experienced repeated waves of outbreaks after seemingly successful efforts of containing COVID-19 transmissions [2,3]. In the absence of effective treatment or vaccination, approaches to preventing transmissions have included social distancing, use of face masks, hygiene measures (eg, disinfection), case isolation, contact tracing, and quarantine [1]. Effective contact tracing is essential for preventing transmissions and subsequent generations of SARS-CoV-2 infections [4].

Contact tracing is the process of identifying and subsequently acquiring information from persons potentially exposed to infectious diseases [4]. Conventional contact tracing methods include a mixture of patient interviews and contact verifications to map the social and work encounters of an infected individual. These methods enable the contact tracer to obtain a comprehensive list of persons potentially exposed to infectious individuals [5,6]. Additional measures, such as patient and visitor registration system downloads, staff roster and electronic medical record (EMR) reviews, and direct observations via closed-circuit television have been employed for contact tracing in health care settings [7]. These processes are labor intensive, yet time sensitive and crucial for stemming the spread of the infectious disease at the emergence of an outbreak [4,8,9].

During the COVID-19 outbreak, contact tracing can be challenging in large communities with a limited pool of contact tracers [6,10]. Technology can potentially overcome this barrier by simplifying the process of data collection. Several health care institutions have benefitted from the use of a real-time locating system (RTLS) for contact tracing. Hellmich et al [11] were able to identify twice as many potential contacts of confirmed pertussis cases in an emergency department with RTLS compared with EMR review. Ho et al [12] also found that the use of location-based RTLSs to complement EMR review increased the sensitivity of detecting health care workers' contact with COVID-19 inpatients from 47.2% to 77.8%. However, despite the positive outcomes of using RTLSs for contact tracing [11-13], such technologies have largely been confined to health care institutions due to the high costs of infrastructure setup [14,15].

Contact tracing apps can provide a more feasible solution to the barrier of scale presented by large communities with high smartphone penetration rates. The ubiquity and relatively low developmental cost of smartphone apps allow for large-scale deployment in time-sensitive situations. Therefore, many countries have jumped on the bandwagon of contact tracing apps [16-19] to augment contact tracing capabilities in response to the COVID-19 pandemic. Likewise, Singapore launched its first Bluetooth-based contact tracing app—TraceTogether—in March 2020 [16,18]. The app is available on both the Google and Apple app stores, and approximately one-sixth of Singapore residents downloaded the app during the peak of the COVID-19 outbreak in April 2020 [20].

Despite the potential of contact tracing apps in enhancing contact tracing efforts during the COVID-19 pandemic, this potential can only be realized when 60% of the population use the app [6]. The efficacy of large-scale technology adoption depends on the law and enforcement measures of the country, public trust in personal data protection, and users' perceived utility of adopting the technology, along with other factors [21-23]. Prior to large-scale technology adoption, the validity of novel contact tracing methods should be assessed to optimize the time and resources used in tackling the COVID-19 pandemic. The effectiveness of digital contact tracing tools has yet to be established due to the paucity of published data in real-life outbreak settings [24]. Hence, we compared the performance of the TraceTogether app with that of a wearable tag-based RTLS and validated both against the EMRs at the National Centre for Infectious Diseases (NCID), the national referral center that managed the majority of Singapore's hospitalized COVID-19–positive patients in May 2020.

Methods

Setting

This study was conducted in the COVID-19 screening center of the NCID in Singapore. The NCID was the designated hospital for managing suspected and confirmed COVID-19 cases during the COVID-19 outbreak in Singapore. Physicians from the co-located Tan Tock Seng Hospital were deployed to the NCID screening center on 10-day rotating shifts to manage the center during the outbreak. Singapore entered a 2-month partial lockdown phase from April 7 through June 1, 2020 due to a surge in COVID-19 cases in the community and foreign worker dormitories.

The study was conducted over a 10-day physician-posting period from May 10 through May 20, 2020. At the time of the study, the NCID was managing more than 70% of Singapore's COVID-19–positive cases. The majority of patients who attended at the screening center during the study period were residents of foreign worker dormitories.

Study Design

We employed a cross-sectional study design to validate the TraceTogether app and the NCID's wearable RTLS wrist tag against EMRs. All patients and physicians in the NCID screening center were issued temporary RTLS tags for contact tracing (Figure 1). Wearing the RTLS tags was mandatory for entry into the NCID screening center (Figure 1). Physicians deployed to the NCID screening center from May 10 to May 19, 2020 were instructed to install the TraceTogether app (version 1.6), activate their smartphone's Bluetooth function during their shifts, and urge their patients to download and activate the TraceTogether app when they medically attended to their patients. Pictorial instructions to download and activate

XSL•FO
RenderX

the TraceTogether app were available in English, Tamil, Bengali, and Mandarin at every screening station (Figure 2).

Figure 1. Mandatory real-time locating system (RTLS) tags worn by patients and staff within the vicinity of the COVID-19 screening centre for contact tracing purposes.

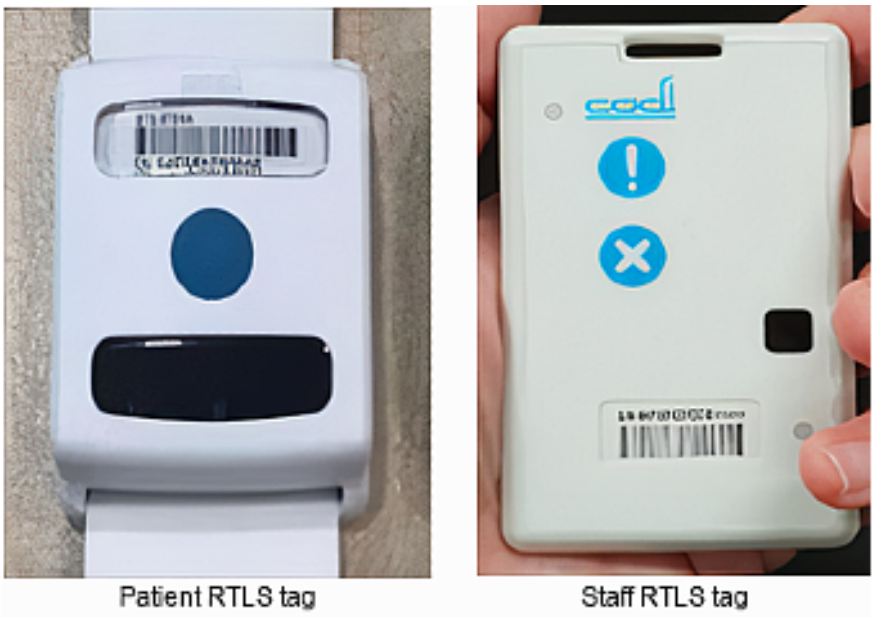


Figure 2. Pictorial instructions to download and activate version 1.6 of the TraceTogether app provided at the National Centre for Infectious Diseases COVID-19 screening centre, Singapore from May 10 through May 20, 2020.

XSL•FO
RenderX

Each patient was assigned to a designated screening station (arranged 2 meters apart from other stations), at which the patient was medically attended. Patients could download the app while awaiting attendance by their physicians using the complimentary Wi-Fi at the screening center. Attending physicians also encouraged patients to download and activate the TraceTogether app.

Tag-Based RTLS

The RTLS was fitted into the NCID building since the building's official opening in September 2019. RTLS location exciters and wireless access points were fitted throughout the NCID building to detect signals from RTLS tags (CADI Scientific). The tag receives a signal whenever it passes a location exciter and sends a radio-frequency signal to the access points to determine the exact location of the RTLS tag. Concurrently, RTLS tags are enabled with tag-tag radio-frequency identification technology to determine close contacts of <=2 meters between staff and patients. Although the location-based radio-frequency identification system in the NCID has been validated by Ho et al [12], the performance of tag-based technology has yet to be assessed.

TraceTogether App

The TraceTogether app was developed by a government-linked company—GovTech—in response to the COVID-19 outbreak. The free-to-download app worked by exchanging Bluetooth signals with other nearby app users and storing the encrypted data locally in the smartphone. Users were requested to upload the data captured on their smartphones should they be confirmed with COVID-19 infection to facilitate contact tracing. Data stored on the phone were automatically deleted after three weeks [17]. Version 1.6 of the app only stored the user's mobile phone number as identity data and collected the phone number, proximity, and duration of users' contacts. Both the mobile phone number and anonymized random user ID (generated from account creation) were stored in the secure GovTech server.

EMR System

All staff providing clinical care to patients at the NCID were issued with rights to make entries in patients' EMRs. Therefore, every patient screened at the NCID screening center had their clinical encounters recorded by the NCID EMR system. Hence, the EMR system was considered the gold standard for validating patient-physician contacts in this study.

Participants

Data validation was only possible when TraceTogether and RTLS data from physicians were available. Therefore, only physicians who uploaded their TraceTogether data and had working RTLS tags were included as participants in this study. By the end of the 10-day posting period, 18 physicians with working RTLS tags had uploaded the data captured from the TraceTogether app from their smartphones to the cloud system.

Data Analyses

We compared patient contacts identified by TraceTogether with those detected by the tag-based RTLS within the NCID screening center from 8:30 AM on May 10, 2020 through 8:30 AM on May 20, 2020 (Table 1). The mobile phone numbers of patients identified by the 18 participating physicians' RTLS tags were compared with the mobile phone numbers captured by the physicians' TraceTogether app. These mobile phone numbers were identified as those from the NCID patient registration (ie, from patients who attended the NCID during the study period).

We also reviewed the EMRs of 156 patients who attended the NCID screening center from 8:30 AM on May 12, 2020 to 8:30 AM on May 13, 2020 to determine if the eight participating physicians who were on shift during that time were among the attending physician(s) of each patient. Only records that fell within each physician's shift were considered. The mobile phone numbers of these 156 patients were compared with the mobile phone numbers of patients identified by the eight physicians' RTLS tags and TraceTogether app, respectively, to determine whether the respective technologies had detected them. We then computed the sensitivity, specificity, positive and negative predictive values, and likelihood ratios of the two digital technologies, using EMR review as the gold standard (Table 2).

XSL•FO
RenderX

Table 1. Comparison of the number of patient contacts identified by a tag-based real-time locating system within the vicinity of the NCID and the TraceTogether app on iPhones and Android phones.

Phone type and physician number	RTLS[a]	TraceTogether[b]	RTLS+TraceTogether	RTLS or TraceTogether	Sensitivity of RTLS[c] (%)	Sensitivity of TraceTogether[d] (%)
iPhones						
1	65	0	0	65	100	0
2	52	0	0	52	100	0
3	2	0	0	2	100	0
4	70	0	0	70	100	0
5	59	4	1	64	93.8	7.8
6	9	3	0	12	75	25
7	14	0	0	14	100	0
8	18	2	0	20	90	10
9	59	0	0	59	100	0
10	14	0	0	14	100	0
Total	362	9	1	372	97.6	2.7[e]
Android phones						
11	83	1	2	86	98.8	3.5
12	61	0	0	61	100	0
13	11	0	1	12	100	8.3
14	78	5	2	85	94.1	8.2
15	12	0	0	12	100	0
16	80	12	6	98	87.8	18.4
17	43	2	0	45	95.6	4.4
18	42	10	3	55	81.8	23.6
Total	410	30	14	454	93.4	9.7[e]
iPhone+Android Phone						
Total	772	39	15	826	95.3	6.5

[a]RTLS: real-time locating system.

[b]TraceTogether-only data includes contacts that were concurrently detected by the TraceTogether app and the NCID location-based RTLS to ensure that the encounters were within NCID's vicinity.

[c]RTLS sensitivity = (number of contacts detected by RTLS + (RTLS+TraceTogether))/(number of contacts detected RTLS or TraceTogether).

[d]TraceTogether sensitivity = (number of contacts detected by TraceTogether + (RTLS+TraceTogether))/(number of contacts detected by RTLS or TraceTogether).

[e]$P<.001$ when the sensitivity of TraceTogether was compared between Android phones and iPhones.

XSL•FO
RenderX

Table 2. Performance of the tag-based real-time locating system and TraceTogether app validated against electronic medical records.[a]

Digital contact tracing tool	Sensitivity[b]	Specificity[c]	Positive predictive value	Negative predictive value	LR+[d]	LR−[e]
Tag-based real-time locating system	96.9% (31/32)	83.1% (103/124)	59.6% (31/52)	99.0% (103/104)	5.73	0.04
TraceTogether app	0.0% (0/32)	98.4% (122/124)	0.0% (0/2)	79.2% (122/154)	0.00	1.02
Either digital contact tracing tool	96.9% (31/32)	81.5% (101/124)	57.4% (31/54)	99.0% (101/102)	5.24	0.04

[a]Physicians 1, 4, 5, 8, 9, 11, 12, and 18 from Table 1 were included in the Table 2 analyses.

[b]Denominator for sensitivity = number of patients attended by eight physicians who uploaded their TraceTogether data and had a work shift between 8:30 AM on May 12, 2020 and 8:30 AM on May 13, 2020.

[c]Denominator for specificity = number of patients attended by other physicians in the same work shift between 8:30 on May 12, 2020 and 8:30 AM on May 13, 2020.

[d]LR+: Positive Likelihood Ratio; LR+ = sensitivity/(1 − specificity).

[e]LR−: Negative Likelihood Ratio; LR− = (1 − sensitivity)/specificity.

Ethical Considerations

This study was approved by the National Healthcare Group Domain Specific Review Board in Singapore.

Results

Table 1 shows the comparative performance of the RTLS and TraceTogether app in identifying patient contacts of individual physicians. The RTLS had a high sensitivity of 95.3% in detecting all patient contacts identified either by the RTLS system or TraceTogether app, while TraceTogether had an overall sensitivity of 6.5%. Version 1.6 of the app performed significantly better on Android phones than iPhones (Android: 9.7%, iPhone: 2.7%, $P<.001$).

Table 2 shows the performance of the tag-based RTLS and TraceTogether app validated against EMRs. RTLS tags had high sensitivity (96.9%) and specificity (83.1%), as the tags could detect patient contacts other than those between patients and the participating physicians who medically attended to them. TraceTogether detected only 2 patient contacts with physicians who did not attend to them. Hence, the app had a sensitivity of 0% and specificity of 98.4% in a clinical setting. The sensitivity of identifying patient contacts increased to 96.9% when both digital contact tracing tools were used.

The positive predictive value and negative predictive value of the RTLS were 59.6% and 99.0%, respectively, while those of TraceTogether were 0% and 79.2%, respectively. Positive and negative predictive values are influenced by the prevalence of the disease, which in this case, was the proportion of patients in contact with physicians. A higher prevalence likely leads to higher positive predictive values. The RTLS's moderately high positive likelihood ratio of 5.73 and high negative likelihood ratio of 0.04 suggest that the RTLS is capable of ruling in and ruling out patient contacts.

Discussion

Effective and timely contact tracing is essential in slowing the spread of COVID-19 in the community. We compared the performance of a contact tracing app—TraceTogether—and NCID RTLS tags in identifying patient-physician contacts and validated both digital contact tracing tools against the EMRs at the NCID COVID-19 screening center. To our knowledge, this is the first study to assess the validity of a contact tracing app in a real-life outbreak setting (ie, the COVID-19 pandemic).

TraceTogether had a much lower sensitivity than the tag-based RTLS in identifying patient contacts in a clinical setting. High sensitivity is preferred for digital contact tracing tools to rule out the possibility of failing to detect close contacts with COVID-19 cases. The low sensitivity of TraceTogether could be attributed to the small proportion of patients who activated the app and turned on their Bluetooth, despite physicians' prompts and the reminder posters at the screening center. In contrast, all patients had RTLS tags attached to their wrists at registration. Hence, we were able to more accurately assess the validity of the RTLS tags in this study.

Although the tag-based RTLS performed well for contact tracing in a clinical setting, the high setup cost would render it less feasible for a community-wide scale-up. Distributing RTLS tags and enforcing their use in the community would be much more challenging than doing so in a clinical setting. Previous RTLS studies have been mainly confined to a defined setting [11,13,25,26]. On the other hand, Bluetooth technology is low cost, available on personal digital devices, and interoperable among Bluetooth-enabled devices [27]. This flexibility is essential for facilitating widespread adoption of contact tracing tools, as users would have the convenience of selecting their preferred form of the contact tracing tool.

A critical mass of app adoption must be achieved to increase the sensitivity and positive predictive value of TraceTogether [28]. Most people have probably not heard of the app or found it useful to download the app during the study period, as movement was restricted due to lockdown measures. The concept of contact tracing apps only actualized on a large scale after it was realized that the fight against COVID-19 was going to be a long haul. Privacy concerns regarding data storage and location tracking were likely the biggest barrier against such contact tracing apps [20,23,28]. Despite the hype surrounding contact tracing apps, health systems worldwide have not revolutionized contact tracing efforts for COVID-19. Partial regulatory enforcements, effective communication of app utility, and a good understanding of the barriers and facilitators of contact tracing app adoption are crucial for app adoption to reach a critical mass.

XSL•FO
RenderX

Given the uncertainty of the adoption and capabilities of contact tracing apps [22,29], contact tracers and policy makers should be cautioned against the overreliance on such apps for contact tracing. An infectious disease workgroup estimated that manual contact tracing would reduce COVID-19 transmissions by 61% compared with the 44% reduction provided by app-based tracing if 53% of the population uses the app [30]. Even if a contact tracing app achieved widespread adoption and high sensitivity in detecting contacts, manual verifications would still be required to ascertain the contact before actions can be undertaken to quarantine potential exposures. In reality, a mixture of contact tracing methods is required to optimize the performance of contact tracing [23].

There were limitations to this study. First, we were unable to enforce the usage of TraceTogether among all physicians at the screening center for a comprehensive review on its performance. We could only consider the patient contacts of physicians who uploaded their TraceTogether data to the cloud system. This limitation in a clinical setting reflects the challenges of wide-spread adoption of app-based contact tracing tools in the community. However, despite the difficulties in enforcing wide-spread adoption, app-based contact tracing can complement conventional contact tracing by speeding up the process. Second, the lackluster performance of TraceTogether version 1.6 on the iPhone may have decreased the performance of the app. Many patient contacts could have been missed, since half of the physicians in this study used an iPhone and the majority of the dormitory workers used an Android phone.

The performance of TraceTogether is expected to improve with app upgrades and increased use over time. We have provided feedback to the developers of TraceTogether, and the app has been upgraded to improve its performance on the iPhone. Since TraceTogether is Singapore's national contact tracing app, many firms have encouraged its use [11] among employees who had to return to work after lockdown. App use has been made mandatory among dormitory workers [31] who have the highest risk of COVID-19 transmissions, and Bluetooth tokens have been distributed to populations susceptible to SARS-CoV-2 infection, such as older adults [32]. Factors that influence people's willingness to adopt TraceTogether are being assessed to achieve higher adoption rates so that the digital contact tracing tool can be effective in the community.

In conclusion, technological and app adoption barriers must be overcome for digital contact tracing tools to be effective for contact tracing during the COVID-19 pandemic. Although the RTLS performed well for contact tracing in a clinical setting, its implementation will be confined to a defined setting. The sensitivity of contact tracing apps needs to be improved for app-based contact tracing to be viable in the community. Leveraging technology to complement conventional manual contact tracing is a necessary move for returning some normalcy to life after exiting lockdowns. The capabilities and utility of digital contact tracing tools are expected to grow over the long haul of the COVID-19 pandemic as the technology matures.

Acknowledgments

We wish to acknowledge the support from our collaborators from GovTech. We also wish to thank the participating physicians for activating their TraceTogether app and reminding their patients to use the app during their COVID-19 duties. Their participation made the study possible.

Authors' Contributions

ZH conceived the study, analyzed and interpreted the data, and drafted the manuscript with input from all authors. GH assisted with interpreting study data and provided input for the manuscript. YML and ECH provided support for study planning and data collection and provided input for the manuscript. HA provided support for study planning and provided input for the manuscript. AC conceived the study, provided overall direction and planning for the study, analyzed and interpreted the data, and critically revised the manuscript. All authors reviewed and approved the final version of the manuscript prior to submission.

Conflicts of Interest

None declared.

References

1. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491) [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
2. Xu S, Li Y. Beware of the second wave of COVID-19. The Lancet 2020 Apr;395(10233):1321-1322. [doi: 10.1016/s0140-6736(20)30845-x]
3. Wise J. Covid-19: Risk of second wave is very real, say researchers. BMJ 2020 Jun 09;369:m2294. [doi: 10.1136/bmj.m2294] [Medline: 32518177]
4. World Health Organization. Contact tracing in the context of COVID-19. COVID-19: Surveilllance, case investigation and epidemiological protocols Internet. URL: https://www.who.int/publications/i/item/contact-tracing-in-the-context-of-covid-19 [accessed 2020-10-19]

XSL•FO
RenderX

5. World Health Organization. Contact tracing. World Health Organization. 2017. URL: https://www.who.int/features/qa/contact-tracing/en/ [accessed 2020-03-17]
6. Baraniuk C. Covid-19 contact tracing: a briefing. BMJ 2020 May 13;369:m1859. [doi: 10.1136/bmj.m1859] [Medline: 32404304]
7. Chow A, Htun HL, Kyaw WM, Ang H, Tan G, Tan HN, et al. Atypical COVID-19: Preventing transmission from unexpected cases. Infect Control Hosp Epidemiol 2020 Aug 13:1-3. [doi: 10.1017/ice.2020.419] [Medline: 32787979]
8. Saurabh S, Prateek S. Role of contact tracing in containing the 2014 Ebola outbreak: a review. Afr Health Sci 2017 Mar;17(1):225-236 [FREE Full text] [doi: 10.4314/ahs.v17i1.28] [Medline: 29026397]
9. Kretzschmar M, Rozhnova G, Bootsma M, van Boven M, van de Wijgert JHHM, Bonten M. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. The Lancet Public Health 2020 Aug;5(8):e452-e459 [FREE Full text] [doi: 10.1016/s2468-2667(20)30157-2]
10. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Glob Health 2020 Apr;8(4):e488-e496. [doi: 10.1016/s2214-109x(20)30074-7]
11. Hellmich TR, Clements CM, El-Sherif N, Pasupathy KS, Nestler DM, Boggust A, et al. Contact tracing with a real-time location system: A case study of increasing relative effectiveness in an emergency department. Am J Infect Control 2017 Dec 01;45(12):1308-1311 [FREE Full text] [doi: 10.1016/j.ajic.2017.08.014] [Medline: 28967513]
12. Ho HJ, Zhang ZX, Huang Z, Aung AH, Lim W, Chow A. Use of a Real-Time Locating System for Contact Tracing of Health Care Workers During the COVID-19 Pandemic at an Infectious Disease Center in Singapore: Validation Study. J Med Internet Res 2020 May 26;22(5):e19437 [FREE Full text] [doi: 10.2196/19437] [Medline: 32412416]
13. Chang Y, Syed-Abdul S, Tsai C, Li Y. A novel method for inferring RFID tag reader recordings into clinical events. Int J Med Inform 2011 Dec;80(12):872-880. [doi: 10.1016/j.ijmedinf.2011.09.006] [Medline: 22018605]
14. Yazici HJ. An exploratory analysis of hospital perspectives on real time information requirements and perceived benefits of RFID technology for future adoption. Int J Inf Manage 2014 Oct;34(5):603-621 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2014.04.010]
15. Vilamovska AM, Hatziandreu E, Schindler HR, van Oranje-Nassau C, de Vries H, Krapels J. Study on the requirements and options for RFID application in healthcare: Identifying areas for Radio Frequency Identification deployment in health care delivery: A review of relevant literature. RAND Technical Report. 2009. URL: https://www.rand.org/pubs/technical_reports/TR608.html [accessed 2020-10-20]
16. O'Neill PH, Ryan-Mosley T, Johnson B. A flood of coronavirus apps are tracking us. Now it's time to keep track of them. MIT Technology Review. 2020 May 07. URL: https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/ [accessed 2020-10-19]
17. Brack S, Reichert L, Scheuermann B. Decentralized Contact Tracing Using a DHT and Blind Signatures. URL: https://eprint.iacr.org/2020/398.pdf [accessed 2020-05-01]
18. Dubov A, Shoptawb S. The Value and Ethics of Using Technology to Contain the COVID-19 Epidemic. Am J Bioeth 2020 Jul;20(7):W7-W11. [doi: 10.1080/15265161.2020.1764136] [Medline: 32420817]
19. Cohen IG, Gostin LO, Weitzner DJ. Digital Smartphone Tracking for COVID-19: Public Health and Civil Liberties in Tension. JAMA 2020 Jun 16;323(23):2371-2372. [doi: 10.1001/jama.2020.8570] [Medline: 32459289]
20. Mathews M, Tan A, Suhaini S. Attitudes towards the use of surveillance technologies in the fight against COVID-19. National University of Singapore, Lee Kuan Yew School of Public Policy, Institute of Policy Studies. URL: https://tinyurl.com/y3ruk5zy [accessed 2020-10-19]
21. Li J, Guo X. Global Deployment Mappings and Challenges of Contact-tracing Apps for COVID-19. SSRN. Preprint posted online on May 26, 2020. [doi: 10.2139/ssrn.3609516]
22. Osman M, Fenton NE, McLachlan S, Lucas P, Dube K, Hitman GA, et al. The thorny problems of Covid-19 Contact Tracing Apps: The need for a holistic approach. Journal of Behavioral Economics for Policy 2020;4(COVID-19 Special Issue):57-61 [FREE Full text]
23. Cohen IG, Gostin LO, Weitzner DJ. Digital Smartphone Tracking for COVID-19: Public Health and Civil Liberties in Tension. JAMA 2020 Jun 16;323(23):2371-2372. [doi: 10.1001/jama.2020.8570] [Medline: 32459289]
24. Anglemyer A, Moore THM, Parker L, Chambers T, Grady A, Chiu K, et al. Digital contact tracing technologies in epidemics: a rapid review. Cochrane Database Syst Rev 2020 Aug 18(8). [doi: 10.1002/14651858.cd013699]
25. Lowery-North DW, Hertzberg VS, Elon L, Cotsonis G, Hilton SA, Vaughns CF, et al. Measuring social contacts in the emergency department. PLoS One 2013 Aug 21;8(8):e70854 [FREE Full text] [doi: 10.1371/journal.pone.0070854] [Medline: 23990915]
26. Ebrahimzadeh F, Nabovati E, Hasibian M, Eslami S. Evaluation of the Effects of Radio-Frequency Identification Technology on Patient Tracking in Hospitals: A Systematic Review. J Patient Saf 2017 Dec 15. [doi: 10.1097/PTS.0000000000000446] [Medline: 29252967]
27. Digi-Key's North American Editors. Comparison of Proprietary RF and Bluetooth. 2018 Feb 28. URL: https://www.digikey.sg/en/articles/comparison-of-rf-and-bluetooth [accessed 2020-10-19]
28. Cho H, Ippolito D, Yu YW. Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs. arXiv. Preprint posted online on March 30, 2020 [FREE Full text]

XSL•FO
RenderX

29. Vaughan A. The problems with contact-tracing apps. New Sci 2020 Apr 25;246(3279):9 [FREE Full text] [doi: 10.1016/S0262-4079(20)30787-9] [Medline: 32372819]
30. Kucharski AJ, Klepac P, Conlan A, Kissler SM, Tang M, Fry H, CMMID COVID-19 Working Group. Effectiveness of isolation, testing, contact tracing and physical distancing on reducing transmission of SARS-CoV-2 in different settings. medRxiv. Preprint posted on April 29, 2020. [doi: 10.1101/2020.04.23.20077024]
31. FAQs on using TraceTogether for contact tracing. Ministry of Manpower. URL: https://tinyurl.com/y5t3yz88 [accessed 2020-10-19]
32. Seniors to receive their first batch of TraceTogether Tokens from 28 June 2020. gov.sg A Singapore Government Agency Website. 2020 Jul 08. URL: https://tinyurl.com/yydkufpw [accessed 2020-10-19]

Abbreviations

EMR: electronic medical record
NCID: National Centre for Infectious Diseases
RTLS: real-time locating system

Edited by G Eysenbach; submitted 03.08.20; peer-reviewed by J Li, A Johnson; comments to author 27.08.20; revised version received 09.09.20; accepted 30.09.20; published 29.10.20.

Please cite as:
Huang Z, Guo H, Lee YM, Ho EC, Ang H, Chow A
Performance of Digital Contact Tracing Tools for COVID-19 Response in Singapore: Cross-Sectional Study
JMIR Mhealth Uhealth 2020;8(10):e23148
URL: http://mhealth.jmir.org/2020/10/e23148/
doi:10.2196/23148
PMID:33006944

©Zhilian Huang, Huiling Guo, Yee-Mun Lee, Eu Chin Ho, Hou Ang, Angela Chow. Originally published in JMIR mHealth and uHealth (http://mhealth.jmir.org), 29.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX