Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 21 (2019), Issue 9 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Reviews

Corrigenda and Addenda

Original Paper

Tailoring Persuasive Electronic Health Strategies for Older Adults on the Basis of Personal Motivation: Web-Based Survey Study

Lex van Velsen[1,2], PhD; Marijke Broekhuis[1,2], MSc; Stephanie Jansen-Kosterink[1,2], PhD; Harm op den Akker[1,2], PhD

[1]eHealth Group, Roessingh Research and Development, Enschede, Netherlands

[2]Biomedical Signals and Systems Group, University of Twente, Enschede, Netherlands

Corresponding Author:
Lex van Velsen, PhD
eHealth Group
Roessingh Research and Development
PO Box 310
Enschede, 7500 AH
Netherlands
Phone: 31 0880875754
Email: l.vanvelsen@rrd.nl

Abstract

Background: Persuasive design, in which the aim is to change attitudes and behaviors by means of technology, is an important aspect of electronic health (eHealth) design. However, selecting the right persuasive feature for an individual is a delicate task and is likely to depend on individual characteristics. Personalization of the persuasive strategy in an eHealth intervention therefore seems to be a promising approach.

Objective: This study aimed to develop a method that allows us to model motivation in older adults with respect to leading a healthy life and a strategy for personalizing the persuasive strategy of an eHealth intervention, based on this user model.

Methods: We deployed a Web-based survey among older adults (aged >60 years) in the Netherlands. In the first part, we administered an adapted version of the revised Sports Motivation Scale (SMS-II) as input for the user models. Then, we provided each participant with a selection of 5 randomly chosen mock-ups (out of a total of 11), each depicting a different persuasive strategy. After showing each strategy, we asked participants how much they appreciated it. The survey was concluded by addressing demographics.

Results: A total of 212 older adults completed the Web-based survey, with a mean age of 68.35 years (SD 5.27 years). Of 212 adults, 45.3% were males (96/212) and 54.7% were female (116/212). Factor analysis did not allow us to replicate the 5-factor structure for motivation, as targeted by the SMS-II. Instead, a 3-factor structure emerged with a total explained variance of 62.79%. These 3 factors are intrinsic motivation, acting to derive satisfaction from the behavior itself (5 items; Cronbach alpha=.90); external regulation, acting because of externally controlled rewards or punishments (4 items; Cronbach alpha=.83); and a-motivation, a situation where there is a lack of intention to act (2 items; r=0.50; P<.001). Persuasive strategies were appreciated differently, depending on the type of personal motivation. In some cases, demographics played a role.

Conclusions: The personal type of motivation of older adults (intrinsic, externally regulated, and/or a-motivation), combined with their educational level or living situation, affects an individual's like or dislike for a persuasive eHealth feature. We provide a practical approach for profiling older adults as well as an overview of which persuasive features should or should not be provided to each profile. Future research should take into account the coexistence of multiple types of motivation within an individual and the presence of a-motivation.

(J Med Internet Res 2019;21(9):e11759) doi:10.2196/11759

KEYWORDS

persuasive communication; health communication; software design

Introduction

Background

In recent years, persuasive design has become an integral part of electronic health (eHealth). Persuasive technology aims to change people's attitudes and behaviors [1] and can be an extremely valuable instrument in increasing patient adherence to Web-based health interventions [2,3], improving self-management abilities [4], and more positive perceptions of usability [5]. To facilitate the design of persuasive systems, a wide range of features have been listed (most notably in a study by Oinas-Kukkonen and Harjumaa [6]). These features include rewarding target behavior, creating trustworthiness, and applying personalized content or services. However, designing a persuasive eHealth intervention is not an easy task of inserting as many persuasive elements as possible. Instead, when combining specific principles, their workings can reinforce one another (eg, combining social learning and comparison), whereas combining too many persuasive features diminishes the effect of the intervention [7]. Choosing the optimal set of persuasive features is a delicate task and might very well depend on the individual end user's characteristics. Tailoring the set of persuasive features a person has to his or her avail, therefore, seems to be a very promising approach [8].

Tailoring, as defined in a study by Hawkins et al, pertains to "any of a number of methods for creating communications individualized for their receivers, with the expectation that this individualization will lead to larger intended effects of these communications" [9]. Although the term is often interchangeably used with *personalization*, we use the term *tailoring* as an umbrella term to cover various specific concepts, such as feedback, context awareness, or user targeting, as defined in a study by op den Akker et al [10]. Research in and applications of tailoring typically focus on specific communications. The Handheld Exercise Agent from Bickmore et al is a good example of this, where tailoring is adopted by providing feedback on the user's personal physical activity and using context-aware features to provide assistive messages in certain situations [11]. Other studies have delved into using machine learning to choose opportune moments for health messages [12] or the use of different framing techniques to increase their intended effect [13]. However, irrespective of whether the technology used is basic or advanced, the principle of the process always remains the same—one must *measure* one or more characteristics of the user (user modeling), *reason* on, and adapt the specific communication to provide and *reach* the individual with the specific outcome.

Studies that aimed to identify the most effective persuasive eHealth tactics for specific user groups have mainly tried to single out a set of features that work best for a given population. For example, Karppinen et al [14] identified a subset of persuasive features that work specifically well for an eHealth intervention, aimed at persons at risk for or suffering from metabolic syndrome. Similarly, Smith et al [15] uncovered a set of features that have the best effect on melanoma checking using a digital intervention. However, they also conclude that personality traits should be taken into account when selecting the most suitable persuasion strategy for an individual. In their context, emotional stability appeared to be a decisive factor for this selection. Other research has reached similar conclusions [16,17]. A promising way to tailor a persuasive strategy, which has not been explored before, is to use an end user's motivation to work on his or her health as the main decisive factor.

Objectives

In this paper, we report on a study that aims to define a strategy for tailoring the set of persuasive features for older adults in an eHealth intervention, aimed at promoting a healthy lifestyle. The basis of this tailoring strategy is a user model in which the individual older adult's motivation to lead a healthy life is modeled. Individual motivation has been identified as a crucial factor for explaining older adults' physical activity levels [18], social participation [19], and cognitive health [20]. Therefore, in the first part of this study, we develop a method for modeling older adults' motivation to adopt a healthy lifestyle. In the second part of this study, we elicit the preferred persuasive eHealth strategies for each motivational profile that was identified.

Theoretical Background

In the psychological literature, there are ample theories and methods to classify different types of motivation. Self-determination theory (SDT), a prominent motivational theory, proposes that the level of autonomy and control [21] influences the level of motivation for a specific action or behavior. Autonomous motivation means that the individual voluntarily performs a behavior because he or she enjoys the activity or finds it interesting. Controlled motivation means that an individual performs a behavior because of external rewards or (social) pressure. SDT describes a continuum of autonomy and control. People can become more or less autonomous or controlled in their motivation. According to SDT, there are 6 types of motivation [21]: (1) *intrinsic motivation*, where one acts because one derives satisfaction from the behavior itself; (2) *extrinsic motivation—integrated regulation*, where one acts because the behavior is in line with one's life goals, objectives, and needs; (3) *extrinsic motivation—identified regulation*, where one acts because something is considered personally important and worthwhile; (4) *extrinsic motivation—introjected regulation*, where one acts to feel worthy, out of guilt, or to avoid shame; (5) *extrinsic motivation—external regulation*, where externally controlled rewards or punishments direct behavior; and (6) *a-motivation*, a situation where there is a lack of intention to act.

Integrated, identified, introjected, and external regulation are all subtypes of extrinsic motivation but differ in the level of autonomy and control. Nonetheless, their common denominator is that a person behaves in a certain way because it leads to a desired outcome, such as receiving a reward, feeling less guilty, or acting in accordance with one's personal values. People who are intrinsically motivated perform the behavior because they like the activity itself [22]. In contrast, although both intrinsic and extrinsic motivations imply the intention to act, a-motivation implies a lack of intention. More fine-grained approaches toward understanding motivation and measurement instruments for assessing this motivation have been developed for different contexts (eg, motivation to exercise [23]). At the moment of writing, however, no methods are available for classifying different kinds of motivation among older adults to lead a healthy lifestyle.

As we mentioned before, a person's type of motivation can play a role when selecting the most suitable persuasive strategy for an individual. This assertion has been studied mostly in the context of promoting exercise and physical activity although not necessarily by means of eHealth interventions. Ingledew and Markland [24] found that, in general, different types of motivation can be associated with different motives for physical exercise and that persuasive strategies should comply with these motives. Kaptein et al describe how they developed personalized persuasive systems using personal persuasion profiles and demonstrate how such systems can persuade people to eat healthier or be more physically active [25]. De Vries et al [26] used the transtheoretical model (which discerns different stages of change for an individual) to provide personalized encouragement for physical exercise through digital services. In another study, De Vries et al [27] found that the Big Five classification can predict the appreciation of different motivational messages, sometimes moderated by the demographic gender. The role of demographics in explaining older adults' health-related behaviors and their motivations for living a healthy lifestyle has been generally acknowledged [28,29].

Research Model and Hypotheses

On the basis of the theoretical background, we created a research model, as depicted in Figure 1.

The research model makes several assumptions. First, it posits that an individual older adult has a specific type of motivation to work on his or her health. This assumption has been tested numerous times for different health-related contexts, with the SDT as the basis for conceptualizing personal motivation [30].

> *Hypothesis 1: An older adult has a personal motivation for working on his or her health that can be classified as intrinsic motivation, integrated regulation, identified regulation, introjected regulation, external regulation, or a-motivation.*

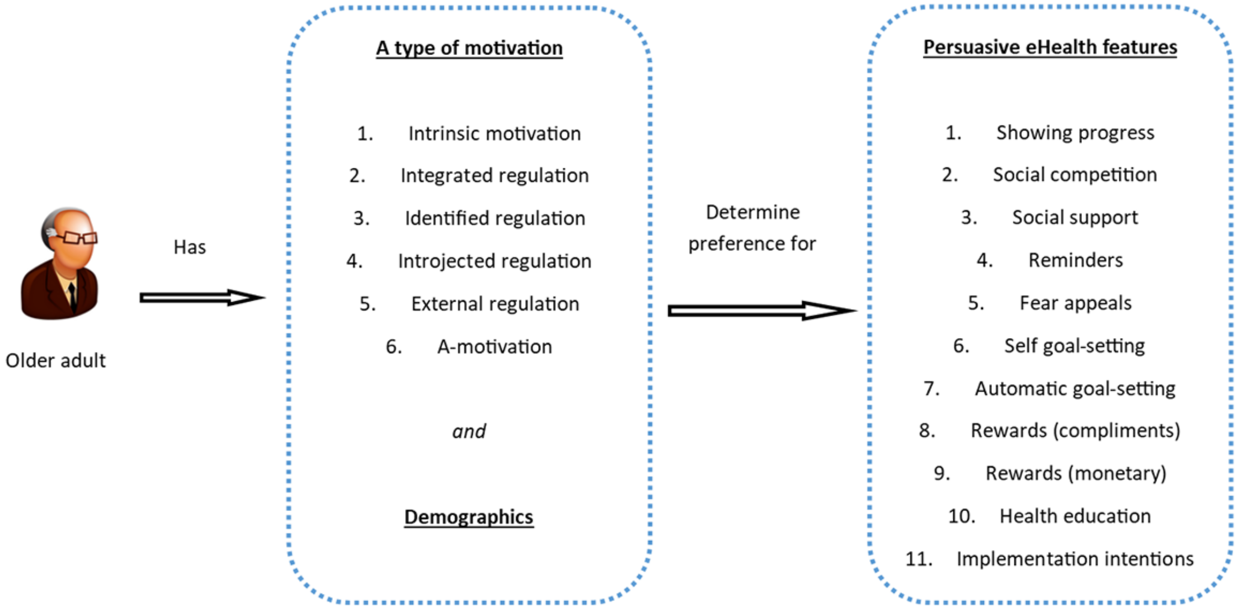
Next, we hypothesize that the persuasive features that an older adult appreciates in an eHealth service can be explained by this type of motivation. Previously, it was found, when motivating people to comply with a physical exercise regime in an offline setting, that persuasive strategies are more effective when tailored to an individual's motivation [24].

> *Hypothesis 2: Each type of motivation is linked to a unique set of highly appreciated persuasive features.*

Finally, we hypothesize that, besides an individual's motivation, demographic factors also predict appreciation of a persuasive feature. This assertion has been scarcely studied before. Age has been found to play a role in the appreciation of persuasive electronic service features in the context of energy saving [31]. Nonetheless, demographics do explain to an important extent to what degree an eHealth intervention is adopted [32,33]. Therefore, we expect them to play a similar role in the formation of appreciation of persuasive features.

> *Hypothesis 3: An older adult's demographic traits (age, gender, education level, and living situation) predict his or her appreciation of a persuasive feature within an eHealth service.*

Figure 1. Research model. eHealth: electronic health.



Methods

Overview

We conducted a Web-based survey among older adults in the Netherlands, which consisted of 3 parts. First, we collected data to test a user profiling method for classifying older adults' motivation to adopt a healthy lifestyle. Second, we provided each participant with a set of persuasive eHealth strategies and asked them to indicate their appreciation for each feature. Finally, we assessed each participant's demographics. The survey was provided through a Web-based survey environment.

User Profiling

To model personal motivation, we used the revised Sports Motivation Scale (SMS-II) [34], a validated instrument for assessing sports motivation; a domain closely related to adopting a healthy lifestyle. The survey is based on the SDT [21] and clusters people on the 6 types of motivation of the SDT (intrinsic motivation, integrated regulation, identified regulation, introjected regulation, external regulation, and a-motivation). We adapted the questionnaire to the needs of this specific context. Each type of motivation, or construct, was assessed via 3 statements, accompanied by a 7-point Likert scale (ranging from very much disagree [1] to very much agree [7]).

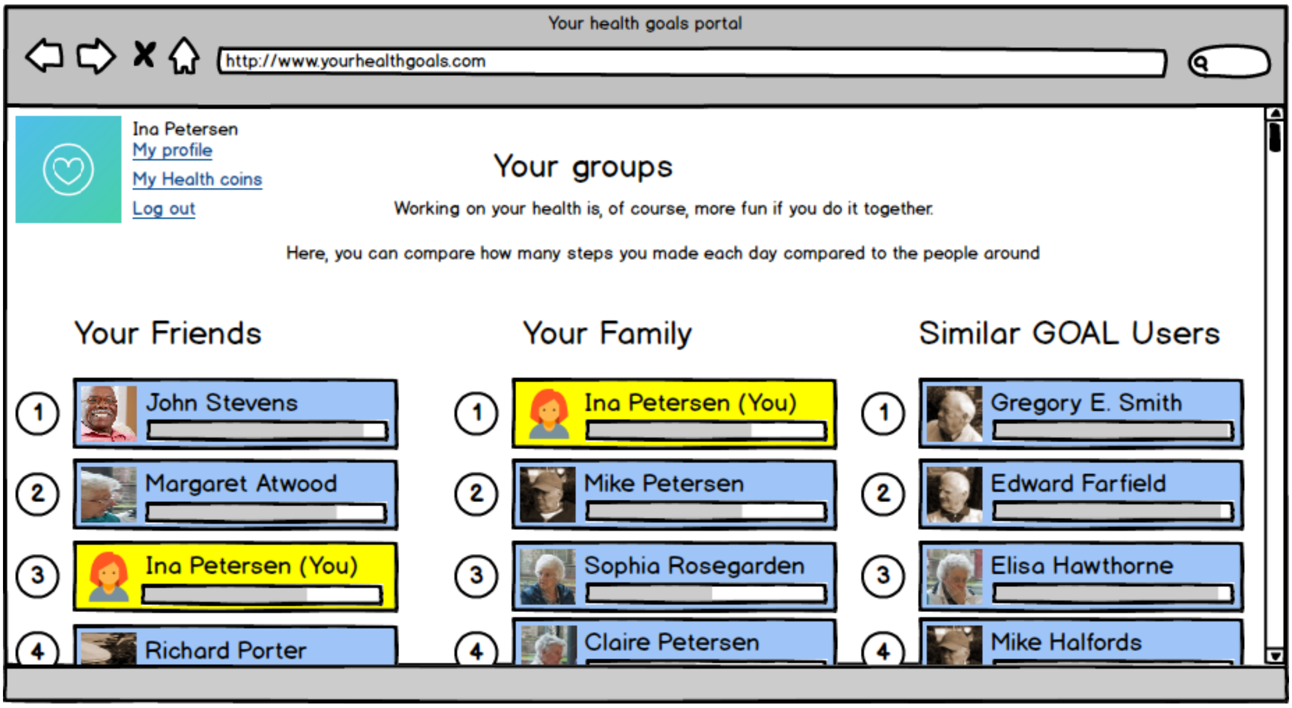
Preferred Persuasive Features

We assessed the appreciation of 11 persuasive features or strategies. The features in this list were selected based on their popularity in existing eHealth interventions.

1. Showing progress: displaying how many recent activities have helped to reach specific health goals.
2. Social competition: showing how healthy your behavior is in relation to peers, in the form of leaderboards.
3. Social support: connecting peers so that they can motivate each other to reach health goals.
4. Reminders: reminding individuals to act healthy or to do exercises by means of pop-ups on, for example, mobile devices.
5. Fear appeals: instilling fear about the current lifestyle of a person by explicating the negative consequences.
6. Self-goal setting: allowing a person to set health goals him/herself (in the mock-up visualized by setting weekly step goal).
7. Automatic goal setting: automated health goal setting for a person (in the mock-up displayed by setting a weekly step goal, based on past performance).
8. Rewards—compliments: complimenting a person on reaching health goals (in the form of badges).
9. Rewards—monetary: awarding points when reaching health goals that can be used at a Web-based shop.
10. Health education: educating persons about the benefits of healthy behavior and the way in which the body reacts to this behavior.
11. Implementation intentions: providing the possibility to plan healthy activities (eg, walking and yoga) in a Web-based, weekly planner.

Each persuasive feature was presented to the participant in the form of a simple digital mock-up, which was created in Balsamiq(Balsamiq Studios). Figures 2 and 3 show two of these mock-ups (social competition and monetary rewards) as examples. Multimedia Appendix 1 gives an overview of all mock-ups. To keep the length of the survey acceptable, we asked each participant to rate 5 of the 11 persuasive features available in total. These 5 features were randomly selected and presented. Each feature was accompanied by a statement (This [information/feature] would motivate me to work on my health) and a 7-point Likert scale ranging from very much disagree [1] to very much agree [7]).

Figure 2. Mock-up of the persuasive feature "social competition".



Your health goals portal

http://www.yourhealthgoals.com

Ina Petersen

My profile

My Health coins

Log out

Your groups

Working on your health is, of course, more fun if you do it together.

Here, you can compare how many steps you made each day compared to the people around

Your Friends	Your Family	Similar GOAL Users
1 John Stevens	1 Ina Petersen (You)	1 Gregory E. Smith
2 Margaret Atwood	2 Mike Petersen	2 Edward Farfield
3 Ina Petersen (You)	3 Sophia Rosegarden	3 Elisa Hawthorne
4 Richard Porter	4 Claire Petersen	4 Mike Halfords

Figure 3. Mock-up of the persuasive feature "monetary rewards".

Your health goals portal

http://www.yourhealthgoals.com

Ina Petersen

My Health Coins

Earn Health Coins and save up for exclusive rewards and discounts!

Current balance: 162 coins

Recent transactions: IN: 5 coins reward for walking 15,000 steps on August, 28 2017.
OUT: 25 coins for 50% discount on 1 kilo apples at Harms vegetable corner
OUT: 20 coins for unlocking a new level in the game Monument Valley

Recommendations from the marketplace

Buy a new level! 30 coins

50% off on 2 avocados! 25 coins

20% Discount on Everything! 200 coins

Demographics

For each participant, we assessed gender, age, educational level (none, elementary school, basic level high school, vocational education, high level high school, and college/university), living situation (living together/married, single, and other), and self-reported physical activity level (by asking them to choose 1 of 5 options: not sporting, not sporting but thinking about beginning, sporting <2.5 hours a week, sporting >2.5 hours a week in the last 6 months, or sporting >2.5 hours a week for more than 6 months). Then we questioned participants' self-reported health by asking participants to rate their health on each of the aspects of health, according to the Positive Health approach [35] (bodily functions, mental functions and perception, spiritual and existential health, quality of life, social and societal participation, and daily functioning), using a 10-point rating scale (ranging from 1 [very poor] to 10 [excellent]). Examples of these statements are as follows: How healthy do you rate your own body? Do you feel fit? Do you have pain somewhere? And can you sleep and eat well? (for physical health), or What do you think of your social life? Do you have enough friends? Do you have others to do fun things with? Do you get help when you need it? And do you have the feeling of belonging somewhere? (for social and societal participation). Finally, we assessed health literacy using the health literacy scale by Chew et al [36], consisting of 3 statements and accompanied by a 6-point Likert scale (ranging from 1 [negative] to 6 [positive]). These statements are as follows: How often do you have problems learning about your medical condition because of difficulty understanding written information? How confident are you filling out medical forms by yourself? How often do you have someone help you read hospital materials?

Recruitment

Participants were eligible if they were aged 60 years or older. We recruited participants through a Dutch panel of older adults that indicated they were interested in participating in research on the topic of eHealth. In addition, we used snowball sampling through social media and personal connections.

Statistical Analyses

All data were analyzed in SPSS 19.0 (SPSS Inc). Descriptive statistics were calculated for demographic variables (percentages, means, and SDs). To determine the suitability of the SMS-II for factor analysis, we assessed interitem correlations, the Kaiser-Meyer-Olkin measure for sampling adequacy, and the Bartlett test of sphericity, and we focused on the diagonals of the anti-image correlation matrix and item communalities. To uncover the factor structure underlying the SMS-II data, we conducted a factor analysis with principal axis factoring and oblimin rotation. The reliability of the resulting factors was assessed by assessing item-to-total correlations and the factor's skewness, kurtosis, and Cronbach alpha (in case a factor had more than 2 items) or by calculating Pearson correlation (in case a factor consisted of 2 items). Distribution of the scores for the different motivational factors was made insightful by means of boxplots, whereas overlap among the factors was displayed using a scatterplot and supported by Pearson correlations. To determine which type of motivation predicts the appreciation of the 11 different persuasive features, we first checked the distribution of these appreciation scores for normality (which they did). Second, we assessed mean appreciation scores and SDs and calculated correlation scores between types of motivation and the appreciation of persuasive features. Finally, we conducted linear regression analyses. At first, we created models wherein appreciation predicts one type of motivation. If significant, we ran a second analysis where we also included the demographics sex, age, living situation, and education as main effects using the Enter method. For this analysis, the variable *education* was split into low and high education, and for the variable *living situation*, the answering option *other* was classified as missing variable (which only occurred once among 212 cases) so that we only had 2 answering options. These changes enabled us to include these variables in regression analyses.

Ethics

Digital informed consent was obtained from each respondent. The nature of this general internet-based survey among healthy volunteers from the general population does not require formal medical ethical approval according to Dutch law.

Results

Demographics

Data collection resulted in 212 valid participants. Of 212 participants, 45.3% (96/212) was male, and 54.7% (116/212) was female. Their ages ranged from 60 to 93 years (mean 68.35 [SD 5.27]). Their educational background consisted of elementary school (1.4%, 3/212), lower vocational education (17.5%, 37/212), vocational education (21.7%, 46/212), high school (14.2%, 30/212), or higher vocational education or university (45.3%, 96/212). Most participants lived together with spouse (79.2%, 168/212). Others lived alone (20.3%, 43/212) or had other living arrangements (0.5%, 1/212).

The participants rated their health literacy as high with a mean of 4.91 (SD 0.63) on a 6-point scale. With regard to their current physical activity behavior, 9/212 participants stated that they did not exercise or compete in sports and were not planning to do so (4.2%); 6/212 participants were not doing this but were thinking about starting (2.8%). Most participants were already physically active. In addition, 45/212 participants (21.2%) exercised or competed in sports, but not on a regular basis (<2.5 hours a week), 44/212 participants (20.8%) exercised or competed in sports regularly between the last 6 months and now (>2.5 hours a week), and 108/212 participants (50.9%) exercised or competed in sports regularly for a period longer than 6 months (>2.5 hours a week). The participants rated their health quite high on different health dimensions:

- Bodily functions: mean 7.27 (SD 1.46)
- Mental functions and perception: mean 7.93 (SD 1.10)
- Spiritual/existential dimension: mean 8.09 (SD 1.15)
- Quality of life: mean 8.04 (SD 1.14)
- Social and societal participation: mean 7.88 (SD 1.38)
- Daily functioning: mean 8.55 (SD 1.02)

User Profiling

As a first step in developing the measurement model for motivation to live a healthy life (and hence, the user profiling method), we examined the factorability of the 18 measurement items for motivation. All items correlated with at least one other item, with a coefficient of 0.3. Next, the Kaiser-Meyer-Olkin measure for sampling adequacy was 0.89, whereas the Bartlett test of sphericity was significant (χ^2_{153}=2177.9; P<.001). The diagonals of the anti-image correlation matrix were also all over 0.6. All item communalities were larger than 0.3. These results indicate that a factor analysis can be conducted with these 18 items.

We conducted a factor analysis with principal axis factoring and oblimin rotation. Initial eigenvalues suggested 3 factors, with an explained variance of 39.67%, 15.38%, and 7.74%, respectively. Total explained variance for these 3 factors was 62.79%. Table 1 displays the item factors loadings and communalities.

The results of the factor analysis show that the 6-factor structure of the SMS-II could not be replicated for a population of older adults when adapted to the context of healthy living (and not sports, as was the original focus of the survey instrument). However, a 3-factor structure became apparent. The first factor, which we call *intrinsic motivation*, focuses on leading a healthy life for personal reasons (eg, because an older adult likes it or wants to develop himself or herself this way). This factor comprised SMS-II items that assess intrinsic motivation and 2 items that, originally, assessed identified regulation but also focus on the self (to develop personal strong suits and develop other sides of myself). The second factor consists of the items that assessed external regulation in the original SMS-II, supplemented by an item from the a-motivation scale that also focuses on external rewards (to get compliments from others). Hence, we retain the name *external regulation*. The third and final factor is *a-motivation*, which consists of the two remaining a-motivation items.

Table 1. Factor loadings and communalities (factor loadings <0.3 are suppressed).

Item	Factor			Communality
	1	2	3	
Intrinsic motivation				
Because I like to learn more about healthy living	0.79	—	—	0.60
Because I like to discover new ways to lead a healthier life	0.88	—	—	0.73
Because I think it’s very interesting to learn how to live a healthier life	0.82	—	—	0.62
Integrated regulation				
Because by living healthy I show who I am	0.40	0.50	—	0.61
Because living healthy is an important part of my life	0.52	—	−0.43	0.67
Because I think it’s very important to live a healthy life	0.43	—	−0.47	0.55
Identified regulation				
Because I chose myself to live a healthier life in order to develop myself	0.59	—	—	0.52
Because I think it is a good way to develop my strong suits	0.70	—	—	0.63
Because I think it’s one of the best ways to develop other sides of myself	0.81	—	—	0.69
Introjected regulation				
Because I would feel bad if I didn’t make time for that	0.45	0.39	—	0.53
Because I would think I am worth little if I did not lead a healthy life	—	0.43	—	0.34
Because I would feel better if I lead a healthy life	0.57	—	−0.42	0.63
External regulation				
Because the people that are important to me would be angry at me if I didn’t	—	0.71	—	0.45
Because I would then be appreciated by the people I know	—	0.65	—	0.57
Because I think other would disapprove of me if I didn’t	—	0.88	—	0.58
A-motivation				
I used to have good reasons to live a healthy life, but lately I’m doubting whether or not to continue with that	—	—	0.61	0.32
So that I get compliments from others	—	0.76	—	0.61
I don’t think that living a healthy life really is something for me	—	—	0.64	0.37

Table 2 displays the reliability scores for the emerging measurement scales for intrinsic motivation and external regulation. As the third factor comprised only 2 items, these metrics do not apply for this scale. The correlation between the two items (“I used to have good reasons to live a healthy life, but lately I’m doubting whether or not to continue with that” and “I don’t think that living a healthy life really is something for me”) was significant (r=0.50; P<.001).

Figure 4 displays 3 boxplots of the score distributions for the different types of motivation in our sample. It shows that, on average, participants scored quite high on intrinsic motivation, but there is a wide range in scores. For external regulation, we also see this wide range; but, in general, participants scored below average. With respect to a-motivation, finally, low scores were observed overall, with a small group of participants who scored high.

Furthermore, we created a scatterplot for intrinsic motivation and external regulation (Figure 5) to show overlap between the two types of motivation. It shows that there is some overlap between the two types of motivation. Participants who were intrinsically motivated sometimes also displayed a high degree of external regulation. It was also possible to be intrinsically motivated only. Calculation of the Pearson correlation between these two factors confirms this (r=0.38; P<.001).

These results partly support our first hypothesis: an older adult has a personal motivation for working on his or her health that can be classified as intrinsic motivation, integrated regulation, identified regulation, introjected regulation, external regulation, or a-motivation. We found that an older adult can be classified as being intrinsically motivated, externally regulated, and a-motivated. However, the classifications are not mutually exclusive.

Table 2. Reliability of the measurement scales for intrinsic motivation and external regulation.

Item	Item-to-total correlation	Skewness	Kurtosis	Cronbach alpha
Intrinsic motivation		−0.61	−0.22	.90
Because I like to learn more about healthy living	0.71			
Because I like to discover new ways to lead a healthier life	0.79			
Because I think it’s very interesting to learn how to live a healthier life	0.74			
Because I think it is a good way to develop my strong suits	0.70			
Because I think it’s one of the best ways to develop other sides of myself	0.79			
External regulation		0.70	0.00	.83
Because the people that are important to me would be angry at me if I didn’t	0.57			
Because I would then be appreciated by the people I know	0.59			
Because I think other would disapprove of me if I didn’t	0.72			
So that I get compliments from others	0.74			

Figure 4. Boxplots of the scores on appreciation for different types of motivation.

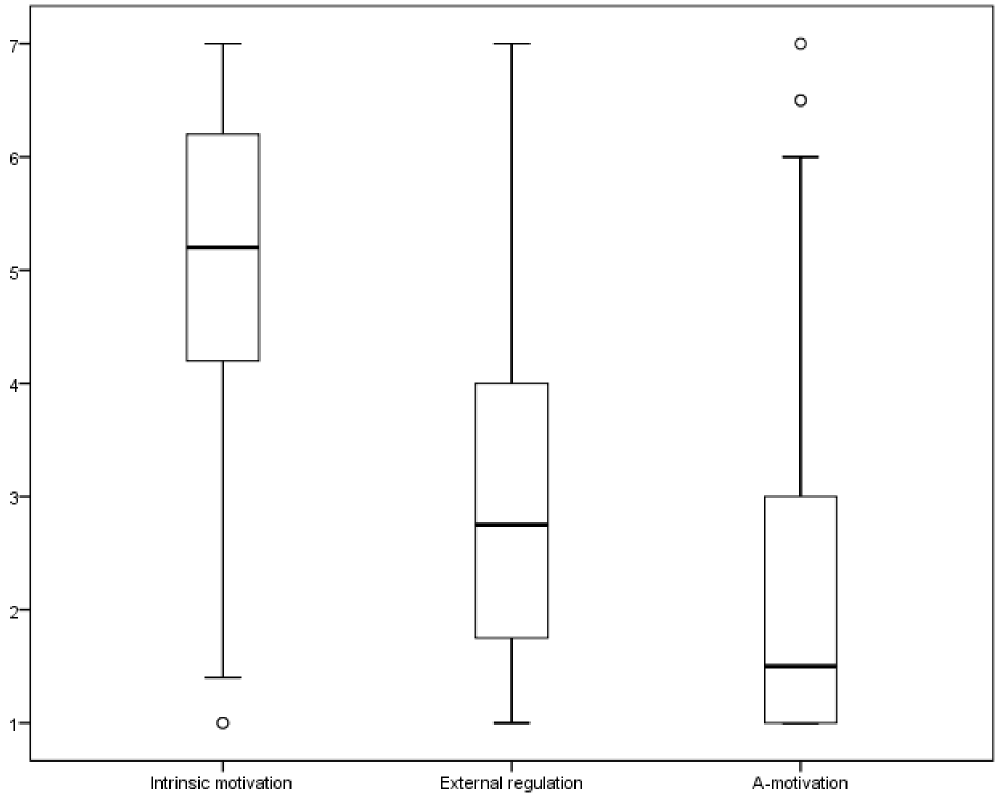
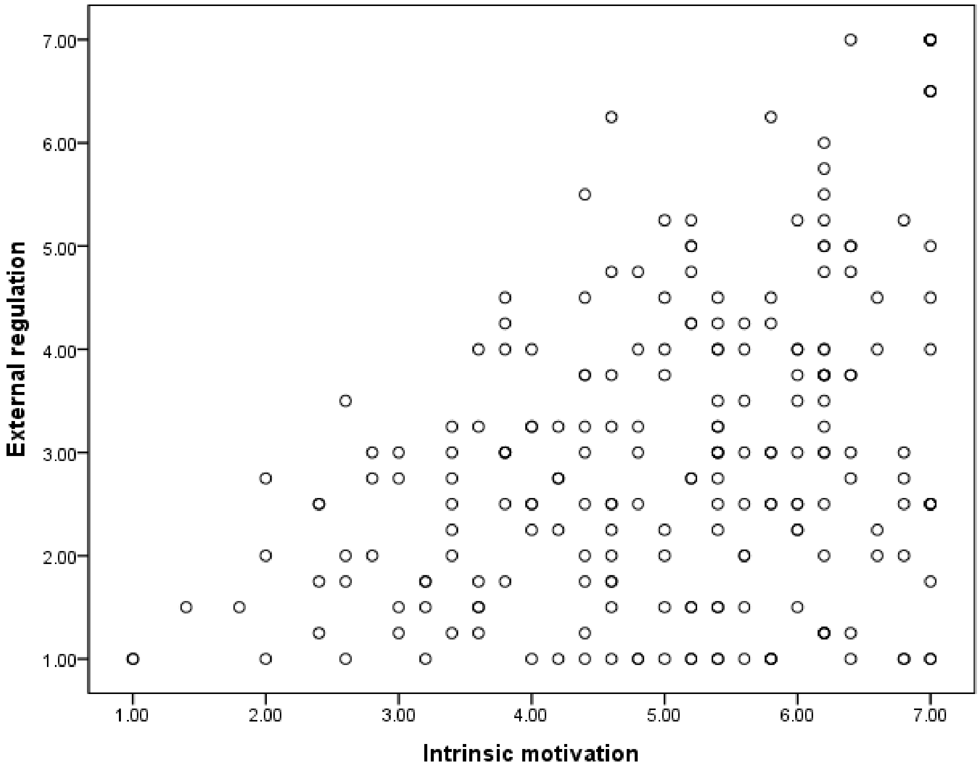


Figure 5. Scatterplot for scores on intrinsic motivation and external regulation.



Appreciation of Persuasive Features

We asked the participants to rate their appreciation for 5 of 11 persuasive features. Table 3 displays how many participants were presented each feature and how they appreciated them. It shows that the appreciations were neutral or positive, whereby fear appeals, monetary rewards, and implementation intentions were appreciated best. In all cases, SDs were high, indicating that participants' appreciations were diverse.

Furthermore, we assessed the correlations among the different types of motivation on the one hand and the appreciation of persuasive features on the other. The results are presented in Table 4. It shows that intrinsic motivation is correlated to the appreciation of all persuasive features, external regulation to some, and a-motivation to none.

Table 3. General appreciation of 11 persuasive features.

Persuasive feature	N	Appreciation, mean (SD)
1. Showing progress	105	3.34 (2.08)
2. Social competition	89	4.14 (1.79)
3. Social support	106	3.92 (1.92)
4. Reminders	106	3.61 (2.09)
5. Fear appeals	101	4.63 (1.99)
6. Self-goal setting	98	3.89 (2.07)
7. Automatic goal setting	109	4.06 (1.88)
8. Rewards—compliments	89	3.40 (1.98)
9. Rewards—monetary	100	4.75 (1.78)
10. Health education	88	3.59 (2.06)
11. Implementation intentions	94	4.40 (1.76)

Table 4. Correlations between types of motivation and appreciation of persuasive features.

Types of motivation	Persuasive feature										
	1	2	3	4	5	6	7	8	9	10	11
Intrinsic motivation	0.42[a]	0.47[a]	0.35[a]	0.25[b]	0.38[a]	0.62[a]	0.44[a]	0.42[a]	0.35[a]	0.37[a]	0.37[a]
External regulation	0.37[a]	0.26[c]	0.17	0.29[b]	0.22[c]	0.33[b]	0.26[b]	0.41[a]	0.17	0.46[a]	0.12
A-motivation	0.01	0.05	0.08	−0.01	0.13	0.06	0.02	0.07	−0.02	0.09	0.16

[a] *P*<.001.

[b] *P*<.01.

[c] *P*<.05.

To determine which type of motivation predicts the appreciation of the 11 different persuasive features, we conducted different sets of linear regression analyses: first, assessing the influence of the type of motivation on the appreciation of a feature; second, assessing the role that demographics play (for results, see Table 5). To preserve a legible overview, we only report significant results for the analyses of demographics.

The results show that being intrinsically motivated positively affected the appreciation of all persuasive features, most heavily showing progress, social competition, self-goal setting, automatic goal setting, and rewards in the form of compliments. Being externally regulated positively affected the appreciation of all persuasive features, except for monetary rewards and implementation intentions. The strongest influences were found for the features showing progress, self-goal setting, rewards in the form of compliments, and health education. A-motivation, as the third motivational style, did not affect the appreciation of any of the persuasive features. Finally, with respect to demographics, we found that health education was better appreciated by participants with a lower education and who are either intrinsically motivated or externally regulated. Intrinsically motivated participants who lived alone appreciated implementation intentions better. Gender and age did not play a significant role.

These results partly support our second hypothesis (each type of motivation is linked to a unique set of highly appreciated persuasive features). Appreciation of all the persuasive features that we studied was affected by either intrinsic motivation and/or external regulation. A-motivation did not affect the appreciation of any persuasive feature. There was a distinction between the features that were highly appreciated by people with high intrinsic motivation on the one hand and high external regulation on the other. Finally, our third hypothesis (an older adult's demographic traits [age, gender, education level, and living situation] predict his or her appreciation of a persuasive feature within an eHealth service) is partly supported by our analyses. Education affected the appreciation of health education (for both the intrinsically motivated and externally regulated persons), whereas participants who had high intrinsic motivation and lived alone appreciated implementation intentions higher.

Table 5. Results of regression analyses.

Persuasive feature	Predictor	Beta	*t* test (*df*)	*P* value	R^2
Showing progress					
1	Intrinsic motivation	.55	6.16 (90)	<.001	0.30
2	External regulation	.42	4.39 (90)	<.001	0.18
3	A-motivation	−.04	−0.40 (90)	NS[a]	0.00
Social competition					
1	Intrinsic motivation	.49	5.05 (79)	<.001	0.24
2	External regulation	.29	2.64 (79)	.010	0.08
3	A-motivation	.04	0.39 (79)	NS	0.00
Social support					
1	Intrinsic motivation	.35	3.61 (91)	.001	0.13
2	External regulation	.25	2.44 (91)	.017	0.06
3	A-motivation	.16	1.51 (91)	NS	0.02
Reminders					
1	Intrinsic motivation	.27	2.60 (84)	.011	0.07
2	External regulation	.32	3.09 (84)	.003	0.10
3	A-motivation	−.03	−0.25 (84)	NS	0.00
Fear appeals					
1	Intrinsic motivation	.36	3.46 (83)	.001	0.13
2	External regulation	.25	2.39 (83)	.019	0.06
3	A-motivation	.15	1.36 (83)	NS	0.02
Self-goal setting					
1	Intrinsic motivation	.64	7.45 (80)	<.001	0.41
2	External regulation	.40	3.86 (80)	<.001	0.16
3	A-motivation	.08	0.74 (80)	NS	0.01
Automatic goal setting					
1	Intrinsic motivation	.43	4.56 (91)	<.001	0.19
2	External regulation	.34	3.44 (91)	.001	0.12
3	A-motivation	.05	0.51 (91)	NS	0.00
Rewards — compliments					
1	Intrinsic motivation	.42	3.91 (73)	<.001	0.17
2	External regulation	.40	3.74 (73)	<.001	0.16
3	A-motivation	.06	0.52 (73)	NS	0.00
Rewards — monetary					
1	Intrinsic motivation	.36	3.60 (87)	.001	0.13
2	External regulation	.18	1.71 (87)	NS	0.03
3	A-motivation	.00	−0.02 (87)	NS	0.00
Health education					
1	Intrinsic motivation	.33	3.55 (85)	.001	0.26
	Education	−.36	−3.82 (85)	<.001	
2	External regulation	.43	4.86 (85)	<.001	0.34
	Education	−.36	−4.04 (85)	<.001	

Persuasive feature	Predictor	Beta	*t* test (*df*)	*P* value	R^2
3	A-motivation	.07	0.63 (75)	NS	0.01
Implementation intentions					
1	Intrinsic motivation	.35	3.76 (91)	<.001	0.22
	Living situation	−.29	−3.07 (91)	.003	
2	External regulation	.21	1.89 (80)	NS	0.04
3	A-motivation	.18	1.66 (80)	NS	0.03

[a] NS: nonsignificant.

Discussion

Principal Findings

The use of persuasive features in eHealth services to improve the uptake of health advice has become very popular in recent years. This study shows that different types of older adults prefer different persuasive features when improving their lifestyle by an eHealth service. To profile older adults, so that the offer of persuasive features can be tailored to the individual, we tried to classify older adults' motivation in this context. For this, we used the SMS-II. However, in this context, we were unable to replicate the classification scheme, as posited by the SDT. Rather, our analysis showed that older adults are intrinsically motivated to work on their health (they derive satisfaction from the behavior itself), externally regulated (they act because of rewards or punishments), or a-motivated (there is a general lack of intention to act). It is also possible to have multiple types of motivation. You can find the final survey for classifying older adults in Multimedia Appendix 2. Most participants could be classified as intrinsically motivated, and smaller groups were externally regulated or a-motivated. Interestingly, we found a relationship between being intrinsically motivated and externally regulated; the two seem to go hand in hand quite regularly.

The appreciation of different persuasive features that aim to motivate older adults to adopt a healthy lifestyle can, so we found, be explained by their motivation. Being intrinsically motivated turned out to be a precursor for appreciating all persuasive features that we tested. Being externally regulated explained appreciation of most persuasive features. By looking at the regression weights and explained variance for each persuasive feature, we can conclude that older adults with a high degree of intrinsic motivation are best served by means of showing progress, social competition, self-goal setting, automatic goal setting, and rewards in the form of compliments. Older adults that are externally regulated would be served best by offering them self-goal setting, rewards in the form of compliments, and health education (especially if they are lower educated).

Finally, being a-motivated turned out to have no effect on the appreciation of persuasive features. This might suggest that older adults who are a-motivated are not in a situation in which they are influenced by persuasive features. Instead, different interventions or strategies need to be installed to get them into a stage of being motivated.

Comparison With Prior Work

The growing interest in persuasive eHealth design has mainly approached the use of persuasive features as a one-size-fits-all solution, which is strange, as personalization is one of the recommendations in the seminal overview of persuasive features by Oinas-Kukkonen [6]. Only recently have some studies focused on the use of personalized approaches toward offering persuasive eHealth features [15,16]. This study, however, is the first to use the end user's motivation to act healthy as a main component in the user model and, thus, as a decisive factor for tailoring the persuasive approach.

Contrary to previous work, which has focused on modeling motivation and acting on these models, we have also taken into account a situation wherein a person is a-motivated. As it turned out, this is a specific group of people with a specific stance toward persuasive technology. In general, methods and strategies specifically designed for engaging this group of people are scarce [37]. A-motivation stems from not perceiving the benefits of the activity [38], not believing the activity will lead to certain outcomes [39], or not feeling adequately skilled for the activity [40]. Studies that analyze behavioral reasons for physical activity among different motivation clusters often do not consider a-motivation [41-43], as it entails an absence of intentionality [22]. Motivating this group of people could be done by means of motivational interviewing or changing the a-motivated person's context (eg, removing places for smoking from the workplace) [37]. Both approaches, however, seem to lie outside the realm of eHealth services, and it might well be the case that for an a-motivated person to use persuasive technology, more intensive, offline types of coaching or transformation need to be applied.

Previous research has treated the different motivation profiles as orthogonal constructs. Our analyses uncovered, however, that multiple types of motivation can exist within one individual (in our case, intrinsic motivation and external regulation mostly). Only recently have researchers recognized that different types of motivation can co-occur [44,45]. Future studies and persuasive eHealth technologies should take into account that a person can be motivated in multiple ways and that persuasive design should be tailored toward this situation. As such, the results of this study support the use of persuasion profiles, in which each type of an individual's motivation is scored on a preset scale [25].

Limitations

The operationalization of the 11 different persuasive strategies that we put to the test in this study was done by means of creating low-fidelity mock-ups. Developing high-fidelity prototypes that would allow for end user interaction would have provided stronger stimulus material. On the other hand, it would not have allowed us to explore such a broad range of persuasive strategies, as it would drastically increase the time needed to complete the survey.

Providing personalized, persuasive eHealth interventions should primarily result in health gains (eg, a healthier diet and less time spent sedentary). We used end-user appreciation as the sole indicator of the success of a persuasive strategy. This was done, as it allowed us to provide and test a set of persuasive strategies within a very short time frame. Future studies should take a longitudinal approach to assess this effect. At first, a suitable personalization strategy should be selected based on a model of the individual participant. Then, the participant should be allowed to interact with the personalized eHealth intervention over time so that the intended health effects can occur. Only at this time can the full effect of personalized persuasion be assessed.

In this study, we focused on older adults (aged >60 years) and their motivation to adopt a healthy lifestyle. Next, during recruitment, we used a panel of older adults that showed an interest in eHealth. This led to an overrepresentation of highly educated participants. This might have accounted for the high number of intrinsically motivated participants. In a sample with more lower educated and/or people with less interest in eHealth, the distributions of participants over the different motivational classifications might be somewhat different. However, we do think that for the modeling of motivation and analyses of the relationships between motivational type and the appreciation of different features, this limitation will have no or a very marginal effect. Of course, generalization of these results to other age groups should be done with caution.

Conclusions

Older adults can be classified as being intrinsically motivated, externally regulated, and/or a-motivated when it comes to working on their health. This study provides a set of survey items that can be used to model each type of motivation for an individual and shows which persuasive features can be used best to engage an older adult in working on his or her health. The fact that we found that different types of motivation can coexist within an individual is contradictory to previous research, which has treated different types of motivation as orthogonal constructs. Next, a-motivation is never considered while designing persuasive eHealth technology. This study has shown that we should take this type of motivation into account. We hope that these lessons will further mature the growing field of persuasive technology and eHealth design.

Acknowledgments

This study has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement #731656. The authors would like to thank Miriam Cabrita and Tessa Beinema for their valuable input during the design and data analysis phase of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mock-ups of persuasive features.
[PDF File (Adobe PDF File), 950 KB - jmir_v21i9e11759_app1.pdf]

Multimedia Appendix 2
Final survey for classifying older adults.
[PDF File (Adobe PDF File), 200 KB - jmir_v21i9e11759_app2.pdf]

References

1. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco: Morgan Kaufmann; 2003.
2. 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]
3. Wozney L, Huguet A, Bennett K, Radomski AD, Hartling L, Dyson M, et al. How do ehealth programs for adolescents with depression work? A realist review of persuasive system design components in internet-based psychological therapies. J Med Internet Res 2017 Aug 9;19(8):e266 [FREE Full text] [doi: 10.2196/jmir.7573] [Medline: 28793983]
4. Henkemans OA, van der Boog PJ, Lindenberg J, van der Mast CA, Neerincx MA, Zwetsloot-Schonk BJ. An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technol Health Care 2009;17(3):253-267. [doi: 10.3233/THC-2009-0545] [Medline: 19641261]

5. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in ehealth interventions combining self-tracking and persuasive ecoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017 Aug 1;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
6. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009;24:485-500. [doi: 10.17705/1CAIS.02428]
7. Wildeboer G, Kelders SM, van Gemert-Pijnen JE. The relationship between persuasive technology principles, adherence and effect of web-based interventions for mental health: a meta-analysis. Int J Med Inform 2016 Dec;96:71-85. [doi: 10.1016/j.ijmedinf.2016.04.005] [Medline: 27117057]
8. Ludden GD, van Rompay TJ, Kelders SM, van Gemert-Pijnen JE. How to increase reach and adherence of web-based interventions: a design research viewpoint. J Med Internet Res 2015 Jul 10;17(7):e172. [doi: 10.2196/jmir.4201] [Medline: 26163456]
9. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
10. op den Akker H, Jones VM, Hermens HJ. Tailoring real-time physical activity coaching systems: a literature survey and model. User Model User-Adap 2014 Jun 25;24(5):351-392. [doi: 10.1007/s11257-014-9146-y]
11. Bickmore TW, Mauer D, Brown T. Context awareness in a handheld exercise agent. Pervasive Mob Comput 2009 Jun 1;5(3):226-235 [FREE Full text] [doi: 10.1016/j.pmcj.2008.05.004] [Medline: 20161031]
12. Op den Akker H, Jones VM, Hermens HJ. Predicting Feedback Compliance in a Teletreatment Application. In: Proceedings of the 3rd International Symposium on Applied Sciences. 2010 Presented at: ICAS'10; November 7-10, 2010; Rome, Italy p. 7-10. [doi: 10.1109/ISABEL.2010.5702804]
13. Rothman AJ, Salovey P. Shaping perceptions to motivate healthy behavior: the role of message framing. Psychol Bull 1997 Jan;121(1):3-19. [doi: 10.1037/0033-2909.121.1.3] [Medline: 9000890]
14. Karppinen P, Oinas-Kukkonen H, Alahäivälä T, Jokelainen T, Keränen AM, Salonurmi T, et al. Persuasive user experiences of a health behavior change support system: a 12-month study for prevention of metabolic syndrome. Int J Med Inform 2016 Dec;96:51-61. [doi: 10.1016/j.ijmedinf.2016.02.005] [Medline: 26992482]
15. Smith KA, Dennis M, Masthoff J. Personalizing Reminders to Personality for Melanoma Self-checking. In: Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization. 2016 Presented at: UMAP'16; July 13-16, 2016; Halifax, Nova Scotia, Canada p. 85-93. [doi: 10.1145/2930238.2930254]
16. Josekutty TR, Masthoff J, Oren N. Adapting Healthy Eating Messages to Personality. In: Proceedings of the International Conference on Persuasive Technology. 2017 Presented at: PERSUASIVE'17; April 4-6, 2017; Amsterdam, The Netherlands p. 119-132. [doi: 10.1007/978-3-319-55134-0_10]
17. Orji R, Vassileva J, Mandryk RL. Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. User Model User-Adap Inter 2014 Jul 14;24(5):453-498. [doi: 10.1007/s11257-014-9149-8]
18. Dacey M, Baltzell A, Zaichkowsky L. Older adults' intrinsic and extrinsic motivation toward physical activity. Am J Health Behav 2008;32(6):570-582. [doi: 10.5555/ajhb.2008.32.6.570] [Medline: 18442337]
19. Losier GF, Bourque PE, Vallerand RJ. A motivational model of leisure participation in the elderly. J Psychol 1993 Mar;127(2):153-170. [doi: 10.1080/00223980.1993.9915551] [Medline: 8336294]
20. Hess TM. Selective engagement of cognitive resources: motivational influences on older adults' cognitive functioning. Perspect Psychol Sci 2014 Jul;9(4):388-407 [FREE Full text] [doi: 10.1177/1745691614527465] [Medline: 26173272]
21. Deci EL, Ryan RM. The Handbook of Self-Determination Research. Rochester: University Of Rochester Press; 2002.
22. Ryan RM, Deci EL. Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp Educ Psychol 2000 Jan;25(1):54-67. [doi: 10.1006/ceps.1999.1020] [Medline: 10620381]
23. Buckworth J, Lee RE, Regan G, Schneider LK, DiClemente CC. Decomposing intrinsic and extrinsic motivation for exercise: Application to stages of motivational readiness. Psychology of Sport and Exercise 2007;8(4):441-461. [doi: 10.1016/j.psychsport.2006.06.007]
24. Ingledew DK, Markland D. The role of motives in exercise participation. Psychol Health 2008;23(7):807-828. [doi: 10.1080/08870440701405704] [Medline: 25160882]
25. Kaptein M, Markopoulos P, de Ruyter B, Aarts E. Personalizing persuasive technologies: explicit and implicit personalization using persuasion profiles. Int J Hum-Comput St 2015 May;77:38-51. [doi: 10.1016/j.ijhcs.2015.01.004]
26. de Vries RA, Truong KP, Kwint S, Drossaert CH, Evers V. Crowd-Designed Motivation: Motivational Messages for Exercice Adherence Based on Behavior Change Theory. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI'16; May 7-12, 2016; San Jose, California, USA. [doi: 10.1145/2858036.2858229]
27. de Vries RA, Truong KP, Zaga C, Li J, Evers V. A word of advice: how to tailor motivational text messages based on behavior change theory to personality and gender. Pers Ubiquit Comput 2017 Jun 15;21(4):675-687. [doi: 10.1007/s00779-017-1025-1]
28. Newsom JT, Kaplan MS, Huguet N, McFarland BH. Health behaviors in a representative sample of older Canadians: prevalences, reported change, motivation to change, and perceived barriers. Gerontologist 2004 Apr;44(2):193-205. [doi: 10.1093/geront/44.2.193] [Medline: 15075416]

29. Loeb SJ, O'Neill J, Gueldner SH. Health motivation: a determinant of older adults' attendance at health promotion programs. J Community Health Nurs 2001;18(3):151-165. [doi: 10.1207/S15327655JCHN1803_02] [Medline: 11560108]
30. Ng JY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-determination theory applied to health contexts: a meta-analysis. Perspect Psychol Sci 2012 Jul;7(4):325-340. [doi: 10.1177/1745691612447309] [Medline: 26168470]
31. Shih LH, Jheng YC. Selecting persuasive strategies and game design elements for encouraging energy saving behavior. Sustainability 2017 Jul 21;9(7):1281-1288 [FREE Full text] [doi: 10.3390/su9071281]
32. Price-Haywood EG, Harden-Barrios J, Ulep R, Luo Q. eHealth literacy: patient engagement in identifying strategies to encourage use of patient portals among older adults. Popul Health Manag 2017 Dec;20(6):486-494. [doi: 10.1089/pop.2016.0164] [Medline: 28384076]
33. Reiners F, Sturm J, Bouw LJ, Wouters EJ. Sociodemographic factors influencing the use of ehealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):E645 [FREE Full text] [doi: 10.3390/ijerph16040645] [Medline: 30795623]
34. Pelletier LG, Rocchi MA, Vallerand RJ, Deci EL, Ryan RM. Validation of the revised sport motivation scale (SMS-II). Psychol Sport Exerc 2013 May;14(3):329-341 [FREE Full text] [doi: 10.1016/j.psychsport.2012.12.002]
35. Huber M, van Vliet M, Giezenberg M, Winkens B, Heerkens Y, Dagnelie PC, et al. Towards a 'patient-centred' operationalisation of the new dynamic concept of health: a mixed methods study. BMJ Open 2016 Jan 12;6(1):e010091 [FREE Full text] [doi: 10.1136/bmjopen-2015-010091] [Medline: 26758267]
36. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]
37. Hardcastle SJ, Hancox J, Hattar A, Maxwell-Smith C, Thøgersen-Ntoumani C, Hagger MS. Motivating the unmotivated: how can health behavior be changed in those unwilling to change? Front Psychol 2015;6:835 [FREE Full text] [doi: 10.3389/fpsyg.2015.00835] [Medline: 26136716]
38. Ryan RM. Psychological needs and the facilitation of integrative processes. J Pers 1995 Sep;63(3):397-427. [doi: 10.1111/j.1467-6494.1995.tb00501.x] [Medline: 7562360]
39. Seligman ME. Helplessness: On Depression, Development, and Death. San Francisco: W H Freeman; 1975.
40. Deci EL. Intrinsic Motivation: Perspectives in Social Psychology. New York: Plenum Press; 1975.
41. Ferrand C, Nasarre S, Hautier C, Bonnefoy M. Aging and well-being in French older adults regularly practicing physical activity: a self-determination perspective. J Aging Phys Act 2012 Apr;20(2):215-230. [doi: 10.1123/japa.20.2.215] [Medline: 22472581]
42. Edmunds J, Ntoumanis N, Duda JL. A test of self-determination theory in the exercise domain. J Appl Soc Psychol 2006 Sep;36(9):2240-2265. [doi: 10.1111/j.0021-9029.2006.00102.x]
43. Stephan Y, Boiché J, le Scanff C. Motivation and physical activity behaviors among older women: a self-determination perspective. Psychol Women Q 2010 Aug 2;34(3):339-348. [doi: 10.1111/j.1471-6402.2010.01579.x]
44. Gourlan M, Trouilloud D, Boiché J. Motivational profiles for physical activity practice in adults with type 2 diabetes: a self-determination theory perspective. Behav Med 2016;42(4):227-237. [doi: 10.1080/08964289.2014.1001810] [Medline: 25961448]
45. Howard J, Gagné M, Morin AJ, van den Broeck A. Motivation profiles at work: a self-determination theory approach. J Vocat Behav 2016 Aug;95-96:74-89. [doi: 10.1016/j.jvb.2016.07.004]

Abbreviations

eHealth: electronic health
SDT: self-determination theory
SMS-II: revised Sports Motivation Scale

Edited by G Eysenbach; submitted 01.08.18; peer-reviewed by R Cronkite, M Dugas; comments to author 09.10.18; revised version received 31.01.19; accepted 25.05.19; published 06.09.19.

Please cite as:
van Velsen L, Broekhuis M, Jansen-Kosterink S, op den Akker H
Tailoring Persuasive Electronic Health Strategies for Older Adults on the Basis of Personal Motivation: Web-Based Survey Study
J Med Internet Res 2019;21(9):e11759
URL: https://www.jmir.org/2019/9/e11759
doi:10.2196/11759
PMID:31493323

Original Paper

A Digital Intervention Addressing Alcohol Use Problems (the “Daybreak” Program): Quasi-Experimental Randomized Controlled Trial

Robert J Tait[1], BSc, PhD; Raquel Paz Castro[2], PhD; Jessica Jane Louise Kirkman[3], BPsych, MClinPsych, PhD; Jamie Christopher Moore[3], BPsych, Crim (Hons); Michael P Schaub[2], PhD

[1]National Drug Research Institute, Faculty of Health Sciences, Curtin University, Perth, Australia
[2]Swiss Research Institute for Public Health and Addiction, University of Zurich, Zurich, Switzerland
[3]Hello Sunday Morning, Sydney, Australia

Corresponding Author:
Robert J Tait, BSc, PhD
National Drug Research Institute
Faculty of Health Sciences
Curtin University
GPO Box U 1981
Perth, 6845
Australia
Phone: 61 892661610
Email: robert.tait@curtin.edu.au

Abstract

Background: Alcohol use is prevalent in many societies and has major adverse impacts on health, but the availability of effective interventions limits treatment options for those who want assistance in changing their patterns of alcohol use.

Objective: This study evaluated the new *Daybreak* program, which is accessible via mobile app and desktop and was developed by *Hello Sunday Morning* to support high-risk drinking individuals looking to change their relationship with alcohol. In particular, we compared the effect of adding online coaching via real-time chat messages (intervention group) to an otherwise self-guided program (control group).

Methods: We designed the intervention as a randomized control trial, but as some people (n=48; 11.9%) in the control group were able to use the online coaching, the main analysis comprised all participants. We collected online surveys at one-month and three-months follow-up. The primary outcome was change in alcohol risk (measured with the alcohol use disorders identification test–consumption [AUDIT–C] score), but other outcomes included the number of standard drinks per week, alcohol-related days out of role, psychological distress (Kessler-10), and quality of life (EUROHIS-QOL). Markers of engagement with the program included posts to the site and comments on the posts of others. The primary analysis used Weighted Generalized Estimating Equations.

Results: We recruited 398 people to the intervention group (50.2%) and 395 people to the control group (49.8%). Most were female (71%) and the mean age was 40.1 years. Most participants were classified as probably dependent (550, 69%) on the AUDIT–10, with 243 (31%) classified with hazardous or harmful consumption. We followed up with 334 (42.1%) participants at one month and 293 (36.9%) at three months. By three months there were significant improvements in AUDIT–C scores (down from mean 9.1 [SD 1.9] to 5.8 [SD 3.1]), alcohol consumed per week (down from mean 37.1 [SD 28.3] to mean 17.5 [SD 18.9]), days out of role (down from mean 1.6 [SD 3.6] to 0.5 [SD 1.6]), quality of life (up from 3.2 [SD 0.7] to 3.6 [SD 0.7]) and reduced distress (down from 24.8 [SD 7.0] to 19.0 [SD 6.6]). Accessing online coaching was not associated with improved outcomes, but engagement with the program (eg, posts and comments on the posts of others) were significantly associated with improvements (eg, in AUDIT–C, alcohol use and EUROHIS-QOL). Reduced alcohol use was found for both probably dependent (estimated marginal mean of 40.8 to 20.1 drinks) and hazardous or harmful alcohol users (estimated marginal mean of 22.9 to 11.9 drinks).

Conclusions: Clinically significant reductions in alcohol use were found, as well as reduced alcohol risk (AUDIT–C) and days out of role. Importantly, improved alcohol-related outcomes were found for both hazardous or harmful and probably dependent drinkers. Since October 2016, *Daybreak* has reached more than 50,000 participants. Therefore, there is the potential for the program to have an impact on alcohol-related problems at a population health level, importantly including an effect on probably dependent drinkers.

Trial Registration: Australian New Zealand Clinical Trials Registry: ACTRN12618000010291; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=373110

International Registered Report Identifier (IRRID): RR2-10.2196/9982

(J Med Internet Res 2019;21(9):e14967) doi:10.2196/14967

KEYWORDS

alcohol consumption; internet; digital health; intervention study; social marketing; health promotion

Introduction

The use of alcohol is one of the leading causes of disease burden and deaths globally, and among those aged 15-49 years it is the leading risk factor for deaths for both males and females [1]. While there are increasing doubts that alcohol use confers any health benefits [2], national guidelines tend to be framed in terms of reducing the risks to health from alcohol use [3,4]. Despite these guidelines, many people still consume alcohol in a manner that increases their risk of adverse consequences: in Australia, about 38% of those aged 12 years or older exceeded either the single occasion consumption guideline or the regular use guideline in 2016 [5]. Further, the current availability of treatment resources is estimated to only fulfil 26%-48% of the needs for treatment required [6]. In the United States, about 24.5% of those aged 12 years or older are classified as having binged in the last 12 months, about 6.1% have engaged in heavy alcohol use and about 5.1% have an alcohol use disorder [7]. Yet, even among those with an alcohol use disorder, less than 10% are likely to have received treatment in the last year [8].

In order to provide interventions, particularly for those with at-risk patterns of alcohol use (as opposed to those with alcohol use disorders, such as alcohol dependence), opportunistic screening and brief interventions have been developed [9,10]. Subsequently, these face to face interventions have been adapted for Internet delivery, with the intensity varying from personalized feedback to multistage interventions [11,12]. An early meta-analysis found that compared with single session interventions (Hedges' *g*=.27), extended self-help interventions (Hedges' *g*=.61) offered significantly greater reductions in alcohol use [11] but that the addition of guidance did not significantly improve outcomes compared with unguided interventions [11]. In contrast, a recent individual patient meta-analysis based on 14,198 participants found that guided interventions resulted in greater reductions in alcohol use than fully automated interventions [13]. Further, no difference in outcomes was found between those with different drinking profiles (described as regular at-risk drinking, heavy drinking or binge drinking.)

Even where interventions have been shown to be efficacious, there are few publicly available websites, with a recent review identifying 72 trials of alcohol interventions but only eight (11%) with functioning websites [14]. Further, the importance of interventions reaching the whole of the target population has been emphasized if public health impacts are to be realized [15]. To this end, *Hello Sunday Morning* took a different approach to conventional web-based interventions and provided both a social media–based, health promotion movement combined with an online environment that incorporated blogging, social networking, peer support and interventions (eg, behavioral experiments) [16].

Hello Sunday Morning, established in 2010, asked participants to set a public alcohol consumption goal (eg, abstinence or reduction) for a set period of time (eg, three, six or twelve months [16,17]). Previous reports have provided descriptive analysis of participants [16], and an observational study reported a significant improvement in scores on the alcohol use disorders identification test (AUDIT) [18], down from a mean of 20.3 (SD 6.7) at baseline to a mean of 8.9 (SD 8.9) at seven months [19]. In August 2018, the original version of *Hello Sunday Morning* closed, having registered more than 100,000 people, and was superseded by the *Daybreak* program that was accessible via mobile app and desktop.

As with the legacy program, the *Daybreak* program aimed to support individuals in changing their relationship with alcohol via access to peer support, self-guided experiments, and individualized health coaching. *Daybreak* was developed in response to the growing number of probably dependent drinkers engaging with *Hello Sunday Morning*, to ensure that these members had the level of clinical support they required (eg, access to health coaches, psychologists and appropriate levels of peer moderation) that the previous version of the program could not provide.

The objectives of the current study were: (1) to assess the effectiveness of the new *Daybreak* program; and (2) to evaluate the effect of adding a professional, clinical component (termed online coaching) to the program, whereby members had access to counsellors or psychologists via smartphone chat. We hypothesized that those receiving coaching would show greater improvements in alcohol-related measures than those who did not receive coaching. As additional objectives, we evaluated whether engagement with the program improved outcome measures. Further, given the need to provide resources for both those with at-risk alcohol use and those with more severe problems (probable dependence), outcomes were compared for these two groups. *Daybreak* is designed as a standalone intervention.

Methods

Design

The study was planned as a randomized control trial to compare the *Daybreak* program plus online coaching via real-time chat messages (Intervention group) and the *Daybreak* program without online coaching (Control group) at one, three and six months. However, due to a programming error, some (n=48; 11.9%) people in the Control group were able to access online coaching and we could not tell if other members of the Control group found they could access the online coaching but did not opt to use it. Therefore, with the approval of the human research ethics committee, the trial was ended after the last person to consent reached their three-month follow-up rather than the planned six-month period.

Participants and Randomization

We invited new registrants to *Daybreak* to join the study (see Multimedia Appendix 1). To be eligible, participants had to be 18 years or older, a resident in Australia, had to provide a valid email address and had to have access to the Internet. However, we excluded those who reported a history of treatment for cardiovascular disease, as this was likely to confer additional health risks during alcohol withdrawal [20]. Participants had to exceed the at-risk threshold (>7) on the 10-item AUDIT [18]. Scores on the AUDIT between 8-19 indicate a pattern of drinking that is likely to fulfil the International Classification of Diseases criteria for hazardous or harmful drinking, while higher scores equate to alcohol dependence [21], but the AUDIT is a screening tool and not a diagnostic indicator. Finally, as part of our duty of care, we assessed suicide risk with the P4 suicidality screening survey [22]. Those participants classified as above minimal risk were still eligible for the study, but they were also provided with the contact details for Lifeline (a 24-hour support service). Those people classified in the highest AUDIT group (≥20: probable dependence) had it recommended to them to speak to their general practitioner or other health professional prior to reducing their use of alcohol in order to minimize potentially serious complications during alcohol withdrawal. The study planned to use a simple, fully randomized allocation (eg, no blocking). Participants were blind to their condition and follow-up data were collected online by an automated survey and hence blind to condition. Those not meeting the eligibility criteria could still access *Daybreak* but were not part of the study cohort. The processes of screening, enrolment and randomization were fully automated.

Procedure

The *Daybreak* program is available from the Google and Apple App Stores or https://www.daybreakprogram.org/, but it can also be accessed from sites such as healthdirect.gov or other health directories. Alternatively, visitors to the *Hello Sunday Morning* website (www.hellosundaymorning.org) are directed to the program. *Daybreak* is a self-guided program that can be accessed as frequently as required by the participant, and access is controlled via username (email address) and password. Those in the Intervention group also had access to an online health coach between 7:00 and 19:00 on weekdays, but as noted above, some of the Control group also accessed coaching. We commenced recruitment in February 2018 and closed it in November 2018. The *Daybreak* version used throughout the study was version 1.5.8, build 90. Since late 2016, *Daybreak* has reached more than 50,000 participants.

We emailed and sent a text to participants with a link to the relevant follow-up survey after one month and after three months. If required, we sent a reminder text message one week later. If there was no response, a research assistant who was blind to study group allocation telephoned the participant to ask them to complete the follow-up, with a maximum of three telephone reminders calls allowed to be made. At each follow-up, participants were eligible to enter a draw to win an Ipad2. The study received the required institutional ethics approval (Curtin University 2017-0855), and the trial procedure was registered (ACTRN12618000010291). In compliance with Australian ethical guidelines, as all participant were screened with at risk alcohol use, we used an active rather than placebo control group [23].

Sample Size

The sample size estimation was calculated for the planned outcomes at six months. For online alcohol interventions, typical effect size values are in the range of Cohen d=0.3-0.4 at 6 months [11,12,24]. We were not aware of previous investigations involving social networks as a means of reducing alcohol use; however, investigations with other behaviors (eg, diet, physical activity) typically report small but not significant changes [25]. Therefore, we based our sample calculation on a small effect (Cohen f=.10; equivalent to Cohen d=0.2) and assumed that the repeated measures would be correlated at r=0.5. To achieve a power of 0.80 with an alpha P<.05 would require 60 people per group, however, given the clinical interest in the effectiveness of brief alcohol interventions for those with more entrenched problems (probable dependence), and potential gender differences, we aimed to recruit 60 people to the smallest cell (sex by AUDIT risk level by study group). To achieve this, we projected recruitment of 300 participants in each group (600 in total), and assuming that 35% would be lost to follow-up, we targeted 467 per group (N=934).

Outcome Measures

Participants were screened with the 10-item AUDIT, which has been validated in Australia [18] and elsewhere. Scores range from 0-40 (0-7=low risk, 8-19=hazardous or harmful alcohol use, 20-40=probable dependence), and outcomes were assessed as change in the scores for the first three items (termed the AUDIT–C). Prior research has shown that the AUDIT–C can predict clinical outcomes at 12 months [26]. We also assessed secondary outcomes and other measures, which are listed in Textbox 1.

Textbox 1. Secondary outcomes and other measures assessed in the study.

Secondary Outcomes:

- Self-reported alcohol consumption in standard drinks (10 g alcohol) collected via a 7-day drinking diary [27,28]. Australian guidelines recommend no more than two standard drinks per day in the general adult population [3].
- Mental distress was assessed with the Kessler’s K-10 [29]. The K-10 scores have a range of 10-50. We interpreted values of 20-24 as showing mild distress, 25-29 as moderate distress and ≥30 as severe mental health distress [30].
- We used Kessler’s Days out of role [31] to determine the number of days either wholly or partially out of role due to alcohol consumption during the last 30 days. Research shows that people with an alcohol disorder have significantly more days either wholly or partially out of role than those without a disorder [31]. From national Australian data, for those with alcohol dependence, the mean number of days out of role is 3.8 [32].
- Quality of life was assessed with the eight item EUROHIS-QOL (also known as the World Health Organization QOL-8). This has been recommended for use in alcohol and other drug treatment services [33] and has been validated in Australia [34], with a depressed sample having a mean score of 2.71 (SD 0.69) versus a non-depressed group with a mean of 3.30 (SD 0.64).

Other Measures:

- The use of health services was quantified with a checklist of health professionals seen in the last eight weeks. This was adapted from a preexisting checklist [35] by the addition of alcohol or other drug treatment services and alcohol pharmacotherapy.
- A four-point rating item (very bad to very good) from the Pittsburgh Sleep Quality Index was used assess to sleep quality in the last four weeks [36].
- We used the Godin Leisure-Time Exercise survey to estimate total exercise in the last seven days [37]. This allows the metabolic equivalents (METs) from different types of exercise to be combined, with one MET defined as the energy used while sitting at rest [38].
- We quantified adverse events arising from alcohol use, only at baseline, using the CORE survey [39]. As this was originally developed for college populations, two items were modified (ie, “missed a class” changed to “missed a class or work” and “been in trouble with police, resident hall, or other college authorities” changed to “been in trouble with police or other authorities”). Further, as the current study focused on alcohol, the reference to drug use (lead-in statement and from one item “Thought I might have a drinking or other drug problem”) was deleted. For United States college students, more than 30% reported driving under the influence, and between 1/5 and 1/3 report being in an argument or fight[39].

Engagement with the *Daybreak* program was recorded as: (1) engagement with coaching (defined as at least one message sent by the participant to the coach); (2) The number of experiments completed; (3) number of blog shares (posts on their own blog); and (4) number of blog comments (posts on another person’s blog).

Content of the Interventions

Overview

The rationale behind the *Daybreak* program is to help people change their relationship with alcohol. This is facilitated by encouraging participants to establish a goal (eg, abstinence, reduced use), to reflect on their mood, and also to give and receive peer support. Four mechanisms were used to help achieve behavior change, which included weekly check-ins, peer support, behavioral experiments and health coaching.

Weekly Check-Ins

The *Daybreak* program includes self-reported questionnaires to encourage participants to undertake self-reflection to explore their intrinsic motivators for change.

Peer Support

The *Daybreak* program enables participants to connect with other users of the program through a blog function. Internal data showed that 45% of shares on the blog received five or more comments in less than 60 minutes. In addition to supporting others going through the process of becoming a nonconsumer of alcohol, the narrative process enables individuals to construct new self-identities, with transitions in the narratives often noted [40].

Behavioral Experiments

Two components of the *Daybreak* program are self-guided experiments and associated learnings. The experiments draw on a range of theoretical perspectives (eg, cognitive behavioral therapy, acceptance and commitment therapy) and cover five areas (mindfulness, connectedness, resilience, situational strategies and health). For example, a mindfulness experiment might guide participants to be in the moment during a period of craving. There are also some experiments that take a broader perspective to help participants with their general health (eg, fitness routines, healthy eating, sleep hygiene).

Health Coaching

The online health coaching was the critical difference between the Intervention Group and Control group. As previously described [41], all the health coaches fulfilled the relevant guidelines for low intensity mental health providers [42], and some coaches were registered general and clinical psychologists. In-house training was provided to develop skills and knowledge on delivering online services (eg, ethical considerations, forming connections online, coaching procedures, risk management, and platform specific training). Novice coaches received supervision, with further periodic feedback from a senior health coach.

The role of the online coaches was to partner with participants in order to assist them with goal setting and assist them in reaching their goals. All coaching interactions occurred through real-time chat messages on a secure platform. The coaches tailored support to the individual’s requirements and drew on a range of techniques (eg, cognitive behavioral therapy, motivational interviewing and acceptance and commitment therapy techniques) as appropriate. *Daybreak* has a written Risk

Management Protocol covering all members, not just the research cohort, to automatically detect trigger words and alert the clinical team [41]. In addition, forum posts were monitored and the clinical team responded to alerts from other participants.

Analysis

Descriptive data and analyses (eg, *t* tests, Mann-Whitney *U* tests, χ^2 tests) were provided for each group. However, due to concerns over the randomization process, the main repeated measures analysis was for all participants rather than by randomized groups. In addition to the overall results, Multimedia Appendix 2 provides online descriptive data at 1-month and 3-months follow-up by randomized group for each outcome measure.

We used Weighted Generalized Estimating Equation (WGEE) analyses to investigate longitudinal changes in primary and secondary outcomes over the study period. WGEE is a repeated measures regression model that takes into account the correlation of repeated measures within each subject [43]. In contrast to repeated measures analysis of variance (ANOVA), WGEE has minimal assumptions about time dependence and uses all available longitudinal data, irrespective of single missing values at follow-up. WGEE is also more robust than unweighted generalized estimating equation, when the assumption of missing completely at random (MCAR) is violated. To control for attrition at 1-month and 3-months follow-up, we estimated weights as suggested by Salazar et al (2016). Attrition analysis revealed that 1-month assessments were more likely to be completed by participants who were older ($t_{721.92}$= 3.67; $P<.001$), not single (χ^2_3=11.467; P=.009), who used less alcohol in the 7 days prior to baseline ($t_{786.919}$=–2.676; P=.008), who had higher sleep quality values ($t_{703.512}$=–1.953; P=.05), who had more blog shares ($t_{482.08}$=3.744; $P<.001$), blog comments ($t_{550.511}$=4.058; $P<.001$), experiments taken (χ^2_1=15.65; $P<.001$), and finally, who stayed longer in the program ($t_{692.773}$=5.252; $P<.001$). In addition, the 3-month assessments were also more likely completed by participants who were female (χ^2_2=20.366; $P<.001$), who had minimal initial suicide risk (χ^2_2=8.706; P=.01), and who had lower psychological distress ($t_{657.61}$=–2.328; P=.02).

In the first round of analysis, WGEE models included only the time variable to examine significant changes in outcomes over the study course. In the second round, WGEE models included the following predictors: (1) time (baseline versus follow-up assessment); (2) study group; and (3) the interaction term for time by group. In the last round of analyses, WGEE models included: (1) time; (2) baseline variables; (3) program use variables; and (4) interaction terms for time by program use variables. We then used a hierarchical, backwards procedure where we removed predictors with the highest *P*-value one at a time until only significant predictors were retained within the model. In the WGEE parameters, the intercept shows the mean change in the outcome for all participants (and on a population level). Betas are added or subtracted to the intercept, which reveals the change in the outcome for the particular group.

Outcomes by AUDIT risk group (probably dependent versus hazardous or harmful) were assessed with a standard, repeated measures ANOVA, as we were interested in the subject level outcomes rather than the population interpretation from the WGEE. In addition, an alpha level of 0.05 (two tailed) was chosen for all statistical tests conducted in the study. All analyses were performed using the statistical tools SPSS version 22 and R version 3.3 via geepack [44].

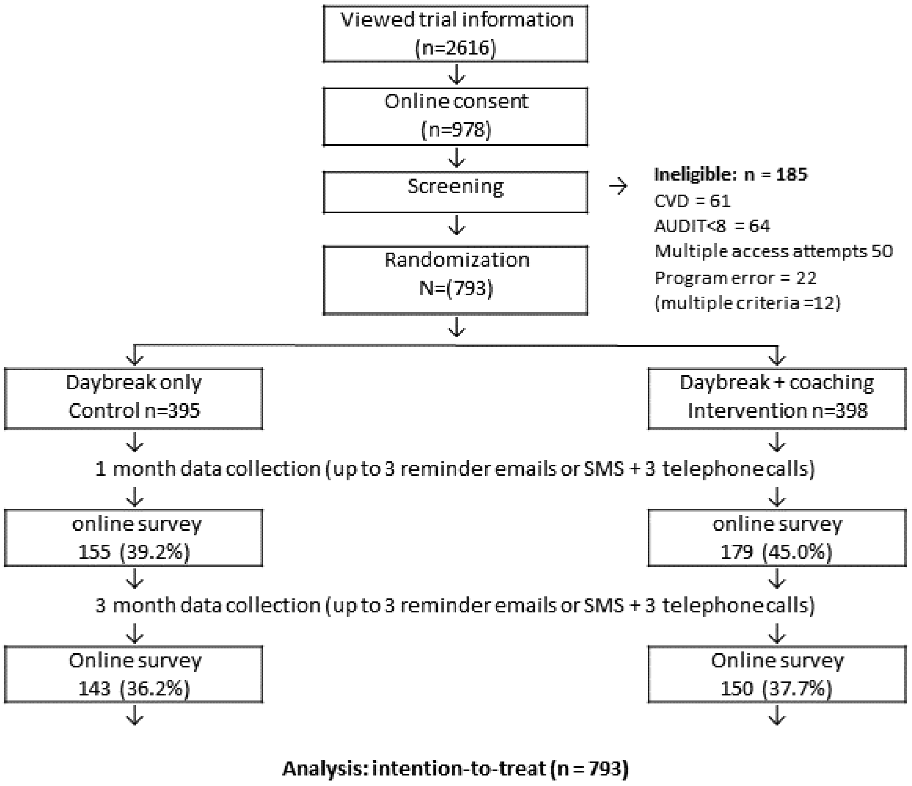
Results

Study Participation, Sample Characteristics, and Attrition

There were 2616 people who viewed the invitation page. Of these, 1278 (49%) accessed additional information and 978 (77%) provided consent. However, not all were eligible for the study. Of everyone who provided consent, 61 indicated a previous cardiovascular disease diagnosis, 28 scored below the AUDIT cut-off score, and 36 did not complete the AUDIT (12 cases failed more than 1 eligibility criteria). Further, there were 50 people who attempted to enroll multiple times after failing an eligibility criterion. We excluded these people, plus an additional 22 people who were not randomized because of a programming error. Thus, the study cohort consisted of 793 people (81.1% of those who consented; see Figure 1).

There were 398 people randomized to the Intervention group (50.2%) and 395 to the Control group (49.8%). Of all the participants, 71% were female (one person selected other) with a mean age of 40.1 years (SD 10.0), and most participants (77%) had at least commenced tertiary education. The mean Kessler's K-10 score was 24.8 (SD 7.0), with 19 participants (2.4%) reporting a serious suicide attempt, and nearly 30% of participants were in the highest risk category on the P4 suicide screen. A total of 69% (n=550) of participants were classified in the AUDIT highest risk category (probably dependent), with their mean alcohol consumption being 37.1 (SD 28.3) standard drinks per week. At baseline, the only significant difference between the groups was in terms of days completely out of role due to alcohol use (Table 1). Follow-up data were subsequently collected from 334 (42.1%) participants at 1 month, and from 293 (36.9%) at 3 months.

Figure 1. Consolidated standards of reporting trials diagram. CVD: cardiovascular disease. AUDIT: alcohol use disorders identification test.



Program Use

Overall, 106 participants engaged with coaching (eg, at least one message sent by the participant to the coach), including 68 from the Intervention and 38 from the Control groups (Table 2). A greater proportion of those in the Intervention group engaged with an online coach (17.1%; χ^2_1=9.5; *P*=.002) than the Control group (9.6%). Blog posts were made by 526 people, with 258 in the Intervention group (65%) and 268 in the Control group (68%). In terms of shares, 230 people (58%) in the Intervention and 227 people (58%) in the Control group commented on the posts of others. There were no significant differences in the median number of experiments taken, shares, comments or days in the study.

Program Effectiveness

Pre-post comparisons of primary and secondary outcomes are displayed in Table 3. The WGEE models revealed significant changes over the study period for all outcomes at *P*<.001, except for amount of exercise at baseline to 1 month (*P*=.15) and at baseline to 3 months (*P*=.65). Participants significantly reduced their alcohol use and their psychological distress, and also gained in sleep quality and quality of life. Further, WGEE revealed that neither study group, nor the interaction term between study group and time, were associated with changes in primary or secondary outcomes. WGEE accounted for missing patterns at T1 and T2, with weights for T1 accounting for age, marital status, alcohol use days (last 7 days), sleep quality, retention, blog shares, blog comments, and experiments seen or done, and weights for T2 accounting for age, gender, marital status, suicide risk, psychological distress (K-10), retention, blog shares, blog comments, and experiments seen or done. Effect size was calculated via Psychometrica [45].

Table 1. Baseline characteristics (N=793).

Variable	Intervention (n=398)	Control (n=395)	*P* value
Sex (female)[a], n (%)	285 (72)	276 (70)	.51
Age (years), mean (SD)	40.9 (10.0)	41.0 (10.1)	.89
Marital status[a], n (%)			**.74**
Single or divorced	139 (35)	129 (33)	
Married	162 (41)	174 (44)	
De facto	87 (22)	80 (20)	
Remarried	10 (3)	12 (3)	
Highest education level[a], n (%)			**.71**
Primary or high school or trade	88 (22)	97 (25)	
Some or complete university	191 (48)	185 (47)	
Some or complete higher degree	119 (30)	113 (29)	
Suicide risk (highest category)[a], n (%)	119 (30)	109 (28)	.30
Kessler-10, mean (SD)	24.5 (7.0)	25.0 (7.1)	.32
Kessler-10 category[a], n (%)			**.62**
Low	35 (9)	36 (9)	
Mild	106 (27)	91 (23)	
Moderate	154 (39)	168 (43)	
Severe	103 (26)	100 (25)	
EUROHIS-QOL, mean (SD)	3.2 (0.7)	3.1 (0.7)	.70
Sleep Quality, mean (SD)	1.7 (0.7)	1.7 (0.8)	.70
Exercise (METS[b])[c], median (IQR)	36 (17-56)	35 (16-56)	.76
Health service use[c], median (IQR)	2 (0-6)	2 (0-5)	.60
Core adverse events, mean (SD)	27.5 (11.3)	26.5 (12.1)	.24
Days out of role[c]			**.02**
Median (IQR)	1 (0-2)	0 (0-1)	
Mean (SD)	1.8 (4.1)	1.4 (2.8)	
Part days out of role[c]			**.47**
Median (IQR)	1 (0-5)	1 (0-4)	
Mean (SD)	3.6 (5.5)	3.5 (5.6)	
AUDIT[d] (10 items; initial screen), mean (SD)	23.0 (6.0)	23.0 (6.5)	.61
AUDIT 10 category[a]			**.45**
8-19 hazardous or harmful	117 (29)	126 (32)	
20-40 dependent	281 (71)	269 (69)	
AUDIT–C (3 items), mean (SD)	9.0 (1.9)	9.2 (1.9)	.29
7-day standard drinks, mean (SD)	37.5 (31.1)	36.8 (25.3)	.74

[a] Assessed with chi-square test.

[b] METS: metabolic equivalents.

[c] Assessed with non-parametric Mann-Whitney *U*.

[d] AUDIT: alcohol use disorders identification test.

Table 2. Program use by study group (N=793).

Variable	Intervention (n=398)	Control (n=395)	*P* value
Engaged with coaching (yes), n (%)	68 (17)	38 (10)	.002
Experiments taken (any)			
n (%)	147 (37)	134 (34)	.40
Median (IQR[a])	0 (0-1)	0 (0-1)	.50
Blog shares, median (IQR)	1 (0-4)	1 (0-10)	.13
Blog comment, median (IQR)	1 (0-10)	2 (0-10)	.74
Time in study (days), median (IQR)	32 (6-87)	30 (4-90)	.51

[a] IQR: interquartile range.

Table 3. Means and standard deviations for outcomes, with significance tests of changes from T0 to T1 and T1 to T2, plus effect size (T0-T2). Please note: not all participants responded to all items in the survey, resulting in missing values.

Variable	Baseline (T0) (n=792)	1 month (T1) (n=333)	3 months (T2) (n=293)	*P* value (T0-T1)	*P* value (T0-T2)	Pre-post effect size (T0-T2)
AUDIT–C[a], mean (SD)	9.11 (1.92)	6.03 (3.02)	5.78 (3.02)	<.001	<.001	–1.53
Alcohol days out of role, mean (SD)	1.59 (3.55)[b]	0.60 (2.46)[b]	0.48 (1.63)	<.001	<.001	–0.262
Alcohol use (last 7 days), mean (SD)	37.10 (28.34)[c]	17.06 (21.57)	17.49 (18.89)	<.001	<.001	–0.689
Kessler K-10, mean (SD)	24.80 (7.03)	19.84 (6.75)	18.97 (6.60)	<.001	<.001	–0.937
EUROHIS, mean (SD)	3.15 (0.70)	3.51 (0.73)	3.57 (0.70)	<.001	<.001	0.745
Sleep Quality, mean (SD)	1.70 (0.74)[d]	1.28 (0.74)	1.27 (0.75)	<.001	<.001	–0.54
Exercise (METS[e]), mean (SD)	55.41 (153.19)	76.20 (246.16)	52.05 (97.88)	.15	.65	–0.015

[a] AUDIT–C: alcohol use disorders identification test-communication.

[b] 11 missing values.

[c] 1 missing value.

[d] 4 missing values.

[e] METS: metabolic equivalents.

Change in alcohol consumption was compared for those who were as classified on the AUDIT–10 at baseline as hazardous or harmful or as probably dependent. Retention was similar for the two groups at three months, with hazardous or harmful at 38.3% (χ^2_1=0.3; *P*=.61) versus probably dependent at 36.4 % (χ^2_1=0.3; *P*=.61). There were significant reductions in alcohol use for both groups (F_{1288}= 24.7; *P*<.001), with estimated marginal means of 40.8 down to 20.1 drinks for probably dependent and 22.9 down to 11.9 drinks for hazardous or harmful, and there was also a significant time by group interaction (F_{1288}= 8.3; *P*=.004) reduction for the probable dependence group. The time by sex and the time by group by sex interactions were not significant.

Factors that predicted significant changes in alcohol-related variables are displayed in Table 4. Online coaching was not significant in any of the models and was therefore not retained. Significant predictors for AUDIT–C were time (reduced at each follow-up), gender (males report higher scores than females), age (older participants report higher scores than younger participants), and the interaction between blog comments and time (more comments were associated with lower AUDIT–C scores at 1-month follow-up). Significant predictors for alcohol days out of role were time (reduced at each follow-up), gender (males report more days out of role than females), marital status (single participants report more days out of role than partnered participants), education (higher education was associated with lower scores), and the interaction between taking part in experiments and time (taking part in experiments was associated with fewer days out of role at 1-month and 3-months follow-up). Significant predictors for alcohol use in standard drinks were time (lower at each follow-up), age (increased with age), marital status (lower for partnered people), and the interaction term between blog shares and time (more shares were associated with lower alcohol use at 1-month and 3-months follow-up).

Table 4. Factors relating to change in alcohol use disorders identification test (AUDIT–C), alcohol days out of role and alcohol use (last 7 days) for all participants (Weighted Generalized Estimating Equation models).

Model, predictors, categories	Beta	SE	*P* value
AUDIT–C			
Model intercept	8.03	0.4	<.001
Time			
Baseline (ref[a])	—[b]	—	—
1 month	–2.88	0.17	<.001
3 months	–3.03	0.21	<.001
Gender			
Female (ref)	—	—	—
Male	0.76	0.23	.001
Age	0.02	0.009	.04
Blog comments	0.003	0.0009	<.001
Blog comments × time			
Baseline (ref)	—	—	—
1 month	–0.008	0.003	.003
3 months	–0.004	0.003	.13
Alcohol days out of role			
Model intercept	2.03	0.21	<.001
Time			
Baseline (ref)	—	—	—
1 month	–0.58	0.27	.03
3 months	–0.94	0.17	<.001
Gender			
Female (ref)	—	—	—
Male	0.45	0.22	.04
Marital status			
Single or divorced (ref)	—	—	—
Married or remarried or de facto	–0.83	0.19	<.001
Education level			
Lower (ref)	—	—	—
Higher	–0.46	0.16	.004
Experiments			
No (ref)	—	—	—
Yes	0.32	0.27	.23
Experiments × time			
Experiments (no) × baseline (ref)	—	—	—
Experiments (yes) × 1 month	–0.86	0.36	.01
Experiments (yes) × 3 months	–0.25	0.3	.40
Alcohol use last 7 days			
Model intercept	26.15	3.19	<.001
Time			
Baseline (ref)	—	—	—

Model, predictors, categories	Beta	SE	*P* value
1 month	–19.17	1.46	<.001
3 months	–18.06	1.42	<.001
Age	0.26	0.08	.002
Marital status			
Single or divorced (ref)	—	—	—
Married or remarried or de facto	–3.93	1.97	.04
Blog shares	0.05	0.05	.26
Blog shares × time			
Baseline (ref)	—	—	—
1 month	–0.13	0.06	.04
3 months	–0.14	0.05	.01

[a] Reference category.

[b] Not applicable.

Factors that significantly predicted changes in psychological distress and quality of life are displayed in Table 5. Significant predictors for the K-10 were time (lower scores at both follow-ups), age (lower with increasing age), marital status (lower for partnered participants), and education level (lower for those with more education). None of the interaction terms for program use by time were retained in the model. Significant predictors for EUROHIS-QOL were time (higher scores at each follow-up), marital status (higher for partnered participants), education (higher for those with more education), and the interaction between blog comments and time (more comments were associated with higher quality of life scores at 1-month follow-up).

Table 5. Factors relating to changes in K-10 and EUROHIS-QOL scores for all participants (WGEE models).

Model, predictor, and categories	Beta	SE	*P* value
Kessler K-10			
Model intercept	28.76	1.14	<.001
Time			
Baseline (ref[a])	—[b]	—	—
1 month	–4.77	0.36	<.001
3 months	–5.33	0.42	<.001
Age	–0.06	0.03	.04
Marital status			
Single/divorced (ref)	—	—	—
Married/remarried/de-facto	–1.62	0.57	.005
Education level			
Lower (ref)	—	—	—
Higher	–1.86	0.56	<.001
EUROHIS-QOL			
Model intercept	2.9	0.04	<.001
Time			
Baseline (ref)	—	—	—
1 month	0.32	0.03	<.001
3 months	0.39	0.04	<.001
Marital status			
Single/divorced (ref)	—	—	—
Married/remarried/de-facto	0.3	0.05	<.001
Education level			
Lower (ref)	—	—	—
Higher	0.17	0.05	.001
Blog comments	–0.0007	0.003	.03
Interaction blog comments × time			
Baseline (ref)	—	—	—
1 month	0.001	0.0005	.03
3 months	0.0007	0.0005	.17

[a] Reference category.

[b] Not applicable.

Discussion

Principal Findings

Use of the *Daybreak* program resulted in significantly improved outcomes in terms of alcohol measures, mental health and quality of life. An important finding was that the program was effective in those who were classified as probably dependent at baseline. Face to face brief interventions are recommended for those with at-risk alcohol use, but they are regarded as ineffective for those with alcohol use disorders [46,47]. It is therefore important to determine the range of alcohol use for which electronic health (eHealth) interventions are effective. Overall, these findings need to be tempered with the fact that the trial was not implemented as designed, and the longest period of follow-up was only three months.

With its multiple components, the *Daybreak* program does not fit within the typical definitions of a brief intervention, certainly not as described for face to face brief interventions [46,48]. In the context of eHealth interventions, the concept of a number of sessions of treatment or time to deliver treatment is ambiguous; however, Riper et al note a distinction is often drawn between "single session e-personalized normative feedback" and extended interventions that draw on a range of therapeutic techniques [11]. The *Daybreak* program lies within the more extensive group. Previously, the extent of engagement with the *Hello Sunday Morning* program (eg, in terms of blog posts, blog shares, following other participants) has been shown to correlate with improved outcomes (ie, reduced AUDIT scores) [19]. Despite the low overall level of engagement (see Table 2), these measures still predicted a range of improved outcomes (eg, in AUDIT–C score, in EUROHIS-QOL score with blog comments about alcohol use, and in blog shares at one month). These findings suggest that the community aspect of *Daybreak* has an important therapeutic role in supporting behavior change.

The outcomes for the study also need to be considered in light of the cohort who participated. Those who access eHealth interventions often have less extensive alcohol use, with a range of 9.1-43.6 standard drinks (mean 18.3) [12], than those in face to face interventions (mean 24.4 standard (10g) drinks) [49]. However, in the recent analysis by Riper et al, the mean consumption was 38.1 standard drinks per week [13]. Most of our participants were classified as probably dependent on alcohol, with mean consumption being 37.1 standard drinks, but the prevalence of severe alcohol problems was higher than typically found in eHealth trials. Thus, many of our participants (69%) were classified as probably dependent on the AUDIT whereas the respective figure reported by Riper et al was 22%, but they also reported that 34% were heavy drinkers (defined as >35/50 standard drinks for females or males) [13]. Also notable was the high prevalence of women in our trial, representing 71% of participants compared with typical values of about 50% [12,13] and even compared to the previous assessment of *Hello Sunday Morning* clients (64%) [19]. There are concerns that eHealth alcohol interventions may not be as effective with women as men [13]. Although men had higher AUDIT–C scores and more days out of role than women, gender was not significant as an interaction term in the current study.

Also notable was the high level of distress as measured with the K-10 and the suicide screening scores. Given the evident distress, it is important that eHealth interventions in this area provide appropriate support and referral pathways. In contrast, the days out of role were not as extensive as those reported for people with alcohol dependence [32], and the mean quality of life scores, while lower at baseline than the general population, were above those for a cohort with depression [34]. Further, by three months the mean score was above that for a general population sample [32].

One of the main aims of this trial was to evaluate if outcomes could be enhanced by giving participants access to online coaching. Few participants opted to use this facility, with 17.1% of the Intervention group accessing the service. Overall, online coaching was not associated with any significantly improved outcomes. The lack of an effect for online coaching is similar to previous studies where guidance has not improved alcohol-related outcomes [24], but remains in contrast to both eHealth interventions for other mental health problems where guidance appears to be beneficial (eg, standardized mean difference compared to unguided=–0.27) [50] and the most recent analysis of alcohol interventions [13]. The reason for the potential difference between those with alcohol use problems and other types of mental health issues is unclear but given that improved outcomes were obtained without the provision of this additional resource-intensive component, this means that *Daybreak* can be provided more widely.

Limitations

Clearly the most significant limitation was that the trial was not implemented as designed, with some members of the Control group accessing the online coaching, so this necessitated a change in the planned analysis. In addition, we also stopped the trial at the completion of the three-month follow-up rather than at six months due to this error. Compared with other studies in the field, the level of attrition for this study was high (63%) compared to other alcohol studies of at least three months duration (range=8%-45%; mean 25.2% [12]) [24], thus limiting the generalizations that can be made from these data.

Conclusions

The current Australian guidelines do not provide a formal recommendation for the number of drinks per week that constitute low risk drinking, but they do recommend that average daily consumption should not exceed two drinks per day [3], equating to 14 per week. On this basis, the average reported here (17.5 drinks) still exceeds that figure, albeit from a high starting point. We do note that those with hazardous or harmful alcohol use did reduce their consumption to below the guideline (mean 11.9 at three months). The *Daybreak* program, due to its extensive reach, member engagement and clinical safety, has the potential to realize a population level impact for both at risk drinkers and for probably dependent drinkers [15].

Conflicts of Interest

The trial was funded by *Hello Sunday Morning* under support from the nib foundation and Ian Potter Foundation. RJT is supported by funding from the Australian Government under the Substance Misuse Prevention and Service Improvement Grants Fund through employment at The National Drug Research Institute at Curtin University. JJLK and JCM are employees of *Hello Sunday Morning*.

Multimedia Appendix 1
Participant information.
[PDF File (Adobe PDF File), 325KB - jmir_v21i9e14967_app1.pdf]

Multimedia Appendix 2
Outcomes by randomized groups.

[PDF File (Adobe PDF File), 74KB - jmir_v21i9e14967_app2.pdf]

Multimedia Appendix 3
CONSORT-eHealth checklist (V1.6).
[PDF File (Adobe PDF File), 2MB - jmir_v21i9e14967_app3.pdf]

References

1. GBD 2016 Alcohol Collaborators. Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2018 Dec 22;392(10152):1015-1035. [doi: 10.1016/S0140-6736(18)31310-2] [Medline: 30146330]
2. Chikritzhs T, Stockwell T, Naimi T, Andreasson S, Dangardt F, Liang W. Has the leaning tower of presumed health benefits from 'moderate' alcohol use finally collapsed? Addiction 2015 May;110(5):726-727. [doi: 10.1111/add.12828] [Medline: 25613200]
3. National Health and Medical Research Council. Australian Guidelines to Reduce Health Risks from Drinking Alcohol. Canberra: NHMRC; 2009:ISBN: 1864963743.
4. UK Chief Medical Officers. UK Department of Health. 2015. Alcohol Guidelines Review: Summary of the proposed new guidelines URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/489795/summary.pdf [accessed 2019-06-06] [WebCite Cache ID 78wNxa4Iw]
5. Australian Institute of Health and Welfare. National Drug Strategy Household Survey. Canberra: Government of Australia; 2017:ISBN: 2205-507X.
6. Ritter A, Berends L, Chalmers J, Hull P, Lancaster K, Gomez M. Drug Policy Modelling Program, National Drug and Alcohol Research Centre. 2014. New Horizons: The review of alcohol and other drug treatment services in Australia URL: https://ndarc.med.unsw.edu.au/sites/default/files/ndarc/resources/New%20Horizons%20Final%20Report%20July%202014.pdf [accessed 2019-06-06] [WebCite Cache ID 78wda7yDv]
7. Substance Abuse and Mental Health Services Administration. Results from the 2017 National Survey on Drug Use and Health: Detailed Tables. Rockville: SAMHSA, Center for Behavioral Health Statistics and Quality; 2018.
8. Substance Abuse and Mental Health Services Administration. National Survey on Drug Use and Health (NSDUH) 2017 Data Archive. Rockville: SAMHSA; 2019.
9. Landy MSH, Davey CJ, Quintero D, Pecora A, McShane KE. A Systematic Review on the Effectiveness of Brief Interventions for Alcohol Misuse among Adults in Emergency Departments. J Subst Abuse Treat 2016 Feb;61:1-12. [doi: 10.1016/j.jsat.2015.08.004] [Medline: 26482134]
10. O'Donnell A, Anderson P, Newbury-Birch D, Schulte B, Schmidt C, Reimer J, et al. The impact of brief alcohol interventions in primary healthcare: a systematic review of reviews. Alcohol Alcohol 2014;49(1):66-78 [FREE Full text] [doi: 10.1093/alcalc/agt170] [Medline: 24232177]
11. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 2011;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]
12. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. J Med Internet Res 2010;12(5):e62 [FREE Full text] [doi: 10.2196/jmir.1479] [Medline: 21169175]
13. Riper H, Hoogendoorn A, Cuijpers P, Karyotaki E, Boumparis N, Mira A, et al. Effectiveness and treatment moderators of internet interventions for adult problem drinking: An individual patient data meta-analysis of 19 randomised controlled trials. PLoS Med 2018 Dec;15(12):e1002714 [FREE Full text] [doi: 10.1371/journal.pmed.1002714] [Medline: 30562347]
14. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-Delivered Health Interventions That Work: Systematic Review of Meta-Analyses and Evaluation of Website Availability. J Med Internet Res 2017 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
15. Sundström C, Blankers M, Khadjesari Z. Computer-Based Interventions for Problematic Alcohol Use: a Review of Systematic Reviews. Int J Behav Med 2017 Dec;24(5):646-658 [FREE Full text] [doi: 10.1007/s12529-016-9601-8] [Medline: 27757844]
16. Carah N, Meurk C, Hall W. Profiling Hello Sunday Morning: who are the participants? Int J Drug Policy 2015 Feb;26(2):214-216. [doi: 10.1016/j.drugpo.2014.07.019] [Medline: 25175599]
17. Hamley B, Carah N. Fare. 2012. One Sunday at a time: Evaluating Hello Sunday Morning URL: http://fare.org.au/one-sunday-at-a-time-evaluating-hello-sunday-morning/ [accessed 2017-12-01] [WebCite Cache ID 6vNaxKfyw]
18. Babor T, Higgins-Biddle J, Saunders J, Monteriro M. AUDIT The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care, 2nd Ed. Geneva: World Health Organization; 2001.
19. Kirkman JJL, Leo B, Moore JC. Alcohol Consumption Reduction Among a Web-Based Supportive Community Using the Hello Sunday Morning Blog Platform: Observational Study. J Med Internet Res 2018 Dec 17;20(5):e196 [FREE Full text] [doi: 10.2196/jmir.9605] [Medline: 29773530]

20. Kähkönen S, Zvartau E, Lipsanen J, Bondarenko B. Effects of alcohol withdrawal on cardiovascular system. Prog Neuropsychopharmacol Biol Psychiatry 2011 Mar 30;35(2):550-553. [doi: 10.1016/j.pnpbp.2010.12.015] [Medline: 21185902]
21. The ICD-10 Classification Of Mental And Behavioural Disorders: Diagnostic Criteria For Research. Geneva: World Health Organization; 2019:ISBN: 1-280-11558-0.
22. Cole S, Raju M, Barrett J, Gerrity M, Dietrich A. The MacArthur Foundation Depression Education Program for Primary Care Physicians: background and rationale. Gen Hosp Psychiatry 2000;22(5):299-358. [Medline: 11203035]
23. National Health and Medical Research Council. National Statement on Ethical Conduct in Human Research (updated May 2015). Canberra: NHMRC; 2007:ISBN: 1864962690.
24. Riper H, Blankers M, Hadiwijaya H, Cunningham J, Clarke S, Wiers R, et al. Effectiveness of guided and unguided low-intensity internet interventions for adult alcohol misuse: a meta-analysis. PLoS One 2014;9(6):e99912 [FREE Full text] [doi: 10.1371/journal.pone.0099912] [Medline: 24937483]
25. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
26. Bradley KA, Rubinsky AD, Lapham GT, Berger D, Bryson C, Achtmeyer C, et al. Predictive validity of clinical AUDIT-C alcohol screening scores and changes in scores for three objective alcohol-related outcomes in a Veterans Affairs population. Addiction 2016 Nov;111(11):1975-1984. [doi: 10.1111/add.13505] [Medline: 27349855]
27. Redman S, Sanson-Fisher RW, Wilkinson C, Fahey PP, Gibberd RW. Agreement between two measures of alcohol consumption. J Stud Alcohol 1987 Mar;48(2):104-108. [Medline: 3560945]
28. Sobell LC, Sobell MB, Leo GI, Cancilla A. Reliability of a timeline method: assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. Br J Addict 1988 Apr;83(4):393-402. [Medline: 3395719]
29. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [Medline: 12214795]
30. Australian Mental Health Outcomes and Classification Network. 2005. Kessler - 10 Training Manual URL: https://www.corc.uk.net/media/1275/kessler10_manual.pdf [accessed 2019-06-07] [WebCite Cache ID 78wROUb6E]
31. Kessler RC, Frank RG. The impact of psychiatric disorders on work loss days. Psychol Med 1997 Jul;27(4):861-873. [Medline: 9234464]
32. Slade T, Johnston A, Teesson M, Whiteford H, Burgess P, Pirkis J. The Mental Health of Australians 2. Report on the 2007 National Survey of Mental Health and Wellbeing. Canberra: Department of Health and Ageing; 2009:ISBN: 1-74186-903-X.
33. Network of Alcohol and Other Drugs Agencies. Using the Client Outcomes Management System (COMS). Strawberry Hills: NSW Health; 2012. URL: https://www.nada.org.au/wp-content/uploads/2018/06/nada_coms_instruction_manual_int_final_web.pdf [accessed 2018-01-30] [WebCite Cache ID 6wqWXsA5X]
34. da Rocha NS, Power MJ, Bushnell DM, Fleck MP. The EUROHIS-QOL 8-item index: comparative psychometric properties to its parent WHOQOL-BREF. Value Health 2012 May;15(3):449-457 [FREE Full text] [doi: 10.1016/j.jval.2011.11.035] [Medline: 22583455]
35. Wade AG, Fernández J, François C, Hansen K, Danchenko N, Despiegel N. Escitalopram and duloxetine in major depressive disorder: a pharmacoeconomic comparison using UK cost data. Pharmacoeconomics 2008;26(11):969-981. [Medline: 18850765]
36. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [Medline: 2748771]
37. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985 Sep;10(3):141-146. [Medline: 4053261]
38. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. Clin Cardiol 1990 Aug;13(8):555-565 [FREE Full text] [doi: 10.1002/clc.4960130809] [Medline: 2204507]
39. Presley CA, Meilman PW, Lyerla R. Development of the Core Alcohol and Drug Survey: initial findings and future directions. J Am Coll Health 1994 May;42(6):248-255. [doi: 10.1080/07448481.1994.9936356] [Medline: 8046164]
40. Pennay A, MacLean S, Rankin G. Hello Sunday Morning: Alcohol, (non)consumption and selfhood. Int J Drug Policy 2016 Feb;28:67-75. [doi: 10.1016/j.drugpo.2015.10.008] [Medline: 26619752]
41. Tait RJ, Kirkman JJL, Schaub MP. A Participatory Health Promotion Mobile App Addressing Alcohol Use Problems (The Daybreak Program): Protocol for a Randomized Controlled Trial. JMIR Res Protoc 2018 May 31;7(5):e148 [FREE Full text] [doi: 10.2196/resprot.9982] [Medline: 29853435]
42. Primary Health Network Brisbane North. Low intensity mental health services. Brisbane: PHN Brisbane North; 2015.
43. Salazar A, Ojeda B, Dueñas M, Fernández F, Failde I. Simple generalized estimating equations (GEEs) and weighted generalized estimating equations (WGEEs) in longitudinal studies with dropouts: guidelines and implementation in R. Stat Med 2016 Dec 30;35(19):3424-3448. [doi: 10.1002/sim.6947] [Medline: 27059703]
44. Halekoh U, Højsgaard S, Yan J. The Package for Generalized Estimating Equations. J. Stat. Soft 2006 Dec 22;5(2):1-11. [doi: 10.18637/jss.v015.i02]

45. Psychometrica. 2019. Computation of Effect Sizes URL: https://www.psychometrica.de/effect_size.html [accessed 2019-06-04] [WebCite Cache ID 78s7xDgQl]
46. Moyer A, Finney JW, Swearingen CE, Vergun P. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 2002 Mar;97(3):279-292. [Medline: 11964101]
47. Saitz R. Alcohol screening and brief intervention in primary care: Absence of evidence for efficacy in people with dependence or very heavy drinking. Drug Alcohol Rev 2010 Nov;29(6):631-640 [FREE Full text] [doi: 10.1111/j.1465-3362.2010.00217.x] [Medline: 20973848]
48. Riper H, Tait RJ. Using the Internet for alcohol and drug prevention. In: Miller P, editor. Interventions for Addictions. Amsterdam: Elsevier; 2012:855-863.
49. Darke S. Self-report among injecting drug users: a review. Drug Alcohol Depend 1998 Aug 1;51(3):253-63; discussion 267. [Medline: 9787998]
50. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — A systematic review. Internet Interventions 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]

Abbreviations

ANOVA: analysis of variance
AUDIT: alcohol use disorders identification test
eHealth: electronic health
MCAR: missing completely at random
WGEE: Weighted Generalized Estimating Equation

Edited by G Eysenbach; submitted 11.06.19; peer-reviewed by M Duplaga; comments to author 09.07.19; revised version received 10.07.19; accepted 12.07.19; published 04.09.19.

Please cite as:
Tait RJ, Paz Castro R, Kirkman JJL, Moore JC, Schaub MP
A Digital Intervention Addressing Alcohol Use Problems (the "Daybreak" Program): Quasi-Experimental Randomized Controlled Trial
J Med Internet Res 2019;21(9):e14967
URL: https://www.jmir.org/2019/9/e14967/
doi:10.2196/14967
PMID:31486406

Original Paper

Using Relational Agents to Promote Family Communication Around Type 1 Diabetes Self-Management in the Diabetes Family Teamwork Online Intervention: Longitudinal Pilot Study

Debbe Thompson[1], PhD; Chishinga Callender[1], MSc; Caroline Gonynor[2], MPH; Karen W Cullen[1], DrPH; Maria J Redondo[3], MD, PhD; Ashley Butler[2], PhD; Barbara J Anderson[2], PhD

[1]United States Department of Agriculture/Agricultural Research Service Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX, United States

[2]Department of Pediatrics, Baylor College of Medicine, Houston, TX, United States

[3]Diabetes and Endocrinology Section, Department of Pediatrics, Baylor College of Medicine, Houston, TX, United States

Corresponding Author:
Debbe Thompson, PhD
United States Department of Agriculture/Agricultural Research Service Children's Nutrition Research Center
Baylor College of Medicine
1100 Bates St
Houston, TX, 77030
United States
Phone: 1 713 798 7076
Fax: 1 713 798 7098
Email: dit@bcm.edu

Abstract

Background: Family conflict can reduce adolescent adherence to type 1 diabetes management tasks. The Family Teamwork in-person intervention was shown to be efficacious in reducing conflict and low adherence to diabetes-related tasks. Its reach and potential impact, however, were limited by the need to deliver the intervention sessions in person. Relational agents (ie, computerized versions of humans) have been shown to appeal to diverse audiences and may be an acceptable replacement for a human in technology-based behavior change interventions.

Objective: The purpose of this paper is to present the results of a pilot study assessing feasibility and acceptability of Diabetes Family Teamwork Online, an adapted version of the Family Teamwork intervention, delivered over the internet and guided by a relational agent.

Methods: Parent-adolescent dyads were recruited through a diabetes care clinic at a large tertiary care hospital in the southwestern United States. A one-group design, with assessments at baseline, immediate postintervention, and 3 months later, was used to assess feasibility. A priori feasibility criteria included an assessment of recruitment, completion, attrition, program satisfaction, therapeutic alliance, attitudes toward the relational agent, and data collection. The institutional review board at Baylor College of Medicine approved the protocol (H-37245).

Results: Twenty-seven adolescents aged 10 to 15 years with type 1 diabetes and their parents were enrolled. Criteria used to assess feasibility were (1) recruitment goals were met (n=20), (2) families completed ≥75% of the modules, (3) attrition rate was ≤10%, (4) program satisfaction was high (≥80% of families), (5) therapeutic alliance was high (average score of ≥60/84), (6) families expressed positive attitudes toward the relational agent (average item score of ≥5 on ≥4 items), (7) ≥80% of data were collected at post 1 and post 2, and (8) few technical issues (≤10%) occurred during intervention delivery. All feasibility criteria were met. Qualitative data confirmed that adolescents and parents had positive reactions to both the content and approach.

Conclusions: The Diabetes Family Teamwork Online intervention proved to be a feasible and acceptable method for enhancing communication around diabetes management tasks in families with an adolescent who has type 1 diabetes.

International Registered Report Identifier (IRRID): RR2-10.2196/resprot.5817

(J Med Internet Res 2019;21(9):e15318) doi:10.2196/15318

KEYWORDS

adolescents; family communication; preadolescents; relational agent; type 1 diabetes

Introduction

According to the most up-to-date statistics from the SEARCH for Diabetes in Youth Study Group [1,2], about 154,000 youth are living with type 1 diabetes (T1D), and each year approximately 15,000 youth aged 20 years and younger are diagnosed with T1D, which is second only to asthma as the most common chronic disease of childhood in the United States [3]. Among youth, from 2001 to 2015, 27,000 new cases were diagnosed annually, with children aged 10 to 14 years having the greatest incidence of T1D [4].

Results from the Diabetes Control and Complications Trial (DCCT) increased awareness of how important maintaining near-normal blood glucose levels is to the management of T1D, particularly for delaying and/or preventing T1D complications [5]. However, translation of the rigorous DCCT regimen has been limited by low patient adherence to treatment plans [6]. Adolescents are particularly at risk for low adherence to the constant demands associated with T1D disease management. An efficacy trial of continuous glucose monitoring systems (CGMSs) reported that the adult cohort, but not the adolescent cohort, improved blood glucose control using CGMSs. This may be partially explained by the significantly lower CGMS adherence in the adolescent cohort [7], emphasizing that technology alone, without impacting the environments in which adolescents live and manage T1D, may not produce improved blood glucose control.

It is well documented that the family is important in adolescent adherence [8]. Family conflict and negative communication around diabetes management, especially blood glucose monitoring, are common barriers to adolescent adherence to the diabetes regimen [9]. A meta-analysis of T1D interventions promoting adherence in youth concluded that behavioral interventions without a focus on the personal and interpersonal aspects of the disease were less likely to affect blood glucose control [10]. Therefore, interventions focusing on adolescent adherence and disease management should not ignore the importance of the family, particularly family communication around T1D management tasks, in this process [8].

The Family Teamwork (FT) intervention is an 8-session, in-person intervention delivered by a trained research assistant during routine clinic visits to youth aged 10 to 14 years with T1D and their parents. It was designed to improve hemoglobin A_{1c} (HbA_{1c}) by promoting adherence and blood glucose control in adolescents with T1D [11]. Sessions focused on helping parents and adolescents work together as a team to increase positive parent involvement and reduce family conflict around T1D management. Its efficacy has been demonstrated in two randomized controlled trials, with significant improvements observed in blood glucose monitoring adherence, HbA_{1c} [12,13], and self-reported quality of life [9] in adolescents who participated in FT as compared with those who received standard care only. There was no increase in diabetes-related family conflict, and parents maintained or increased involvement in diabetes management tasks [9]. Reach, however, was limited by the costs associated with having a trained research assistant available to deliver the intervention, the number of families the intervention could be delivered to, and the need for families to travel to the clinic to participate in the intervention.

Communication technologies offer an opportunity to strengthen the reach and effectiveness [14] of interventions while reducing delivery cost. A review found effective eHealth technologies extended the reach of adult diabetes management interventions [15]. However, no studies of online self-management interventions for youth with T1D that met methodological criteria were identified [16], and most studies found lacked details on intervention components and how interventions were tailored for individual patients. An online program to meet the social and informational needs of older adolescents and young adults with T1D was found to be feasible, but only with multiple reminders from the clinic team [17]. Several small online interventions with youth with T1D improved self-management behavior [18,19]. One reported that translating an in-person coping skills intervention for adolescents with T1D that improved glycemic control to the internet was feasible [20].

Self-management is essential to successful treatment and control of chronic diseases like T1D [5,21]. A meta-analysis of pediatric eHealth interventions suggested that interventions with behavioral methods like self-monitoring, goal-setting, and problem-solving were much more effective than those that were solely educational [22], complementing the findings of the meta-analysis cited previously [10]. Both support that adherence and blood glucose control would be improved by intervening with the parent/adolescent dyad to address family barriers around adolescent adherence, particularly when supported by behavioral self-management (goal-setting, problem-solving, self-monitoring). The gold standard for health behavior change is face-to-face interaction with a health care provider [23,24]; however, limitations such as reach [24], time [23], cost, and fidelity can reduce effectiveness [23].

Relational agents (Ragts) may address these limitations. Ragts are computer representations of health care providers that simulate face-to-face interaction with an individual in the real world [23]. They mimic characteristics of face-to-face interactions, including verbal and nonverbal behaviors associated with trust-, rapport-, and relationship-building. Programs delivered by Ragts are convenient, readily accessible, and relatively low in cost, particularly when delivered online [23]. Ragts have been found to be feasible and acceptable with a variety of age groups and health behaviors [23,25-30]. The purpose of this paper is to report the feasibility of an intervention, Diabetes Family Teamwork Online (FTO), an updated version of FT adapted for online delivery by a Ragt.

Methods

Adaptation of Family Teamwork Into Diabetes Family Teamwork Online

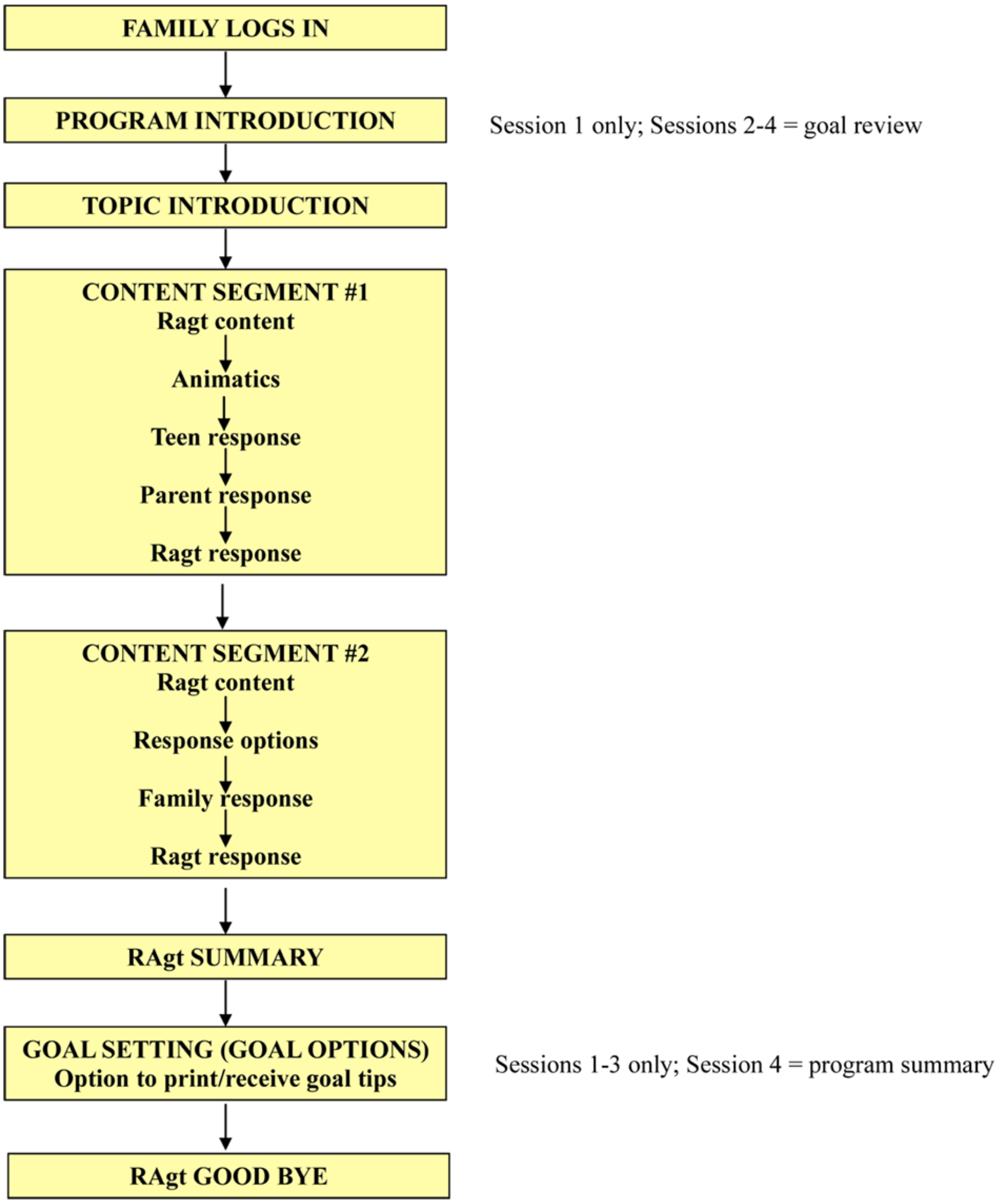
Using an approach informed by the authors’ previous work in the development of technology-based interventions, computers as persuasive technologies [31], and self determination theory (SDT) [32], FT was systematically adapted for online delivery. Informed by social cognitive theory [33], FT contained 8 informational sessions and 1 review session focused on healthy family communication around key T1D management tasks and common T1D-related issues encountered by families, such as avoiding diabetes burnout (Table 1). A trained research assistant delivered the sessions to parent/adolescent dyads during routine visits to the adolescent’s diabetes care provider, meaning that an extended period of time likely elapsed between clinic visits (eg, 3 to 4 months). Therefore, by necessity, there was a fair amount of repetition included in the sessions. After reviewing FT session content, the research team determined that the sessions could be collapsed into 4 online modules delivered to families every 2 weeks without loss of fidelity to the original program.

To facilitate translation and ensure a balance between didactic components and interactivity, a flow diagram was developed to guide the online program (Figure 1). Using feedback obtained from interviews with families during development [34], a team consisting of a key member of the FT development team (BJA), a pediatric endocrinologist (MJR), and a child psychologist (AB) collapsed the 8 sessions into 4 online modules (Figure 2). Using the flow diagram as a guide, scripts were written for each module. Consistent with SDT, scripts emphasized satisfaction of the basic psychological needs of autonomy (choice and control), competence (knowledge, skill), and relatedness (connection to important others). This was done by providing families with choices as they navigated the modules, offering guidance regarding responses to common situations families encounter regarding T1D management, and emphasizing how to work together to manage blood glucose.

Table 1. Family Teamwork session topics.

Session	Topic
1	• Diabetes and the family • Challenges of diabetes
2	• Tools • Blood sugar monitoring • Hemoglobin A_{1c}
3	• Checking blood sugars • Talking about blood sugars • Avoiding blame
4	• Sharing the burden • Identifying blood sugar patterns • Individualizing care
5	• Flexibility in meal planning • Carbohydrate counting • Exercise
6	• Reassessment of the burden • Preventing burnout • Achieving flexibility
7	• Miscarried helping • Interdependence versus independence • Reducing conflict
8	• Review
9	• Research and technology update • Advances in monitoring devices

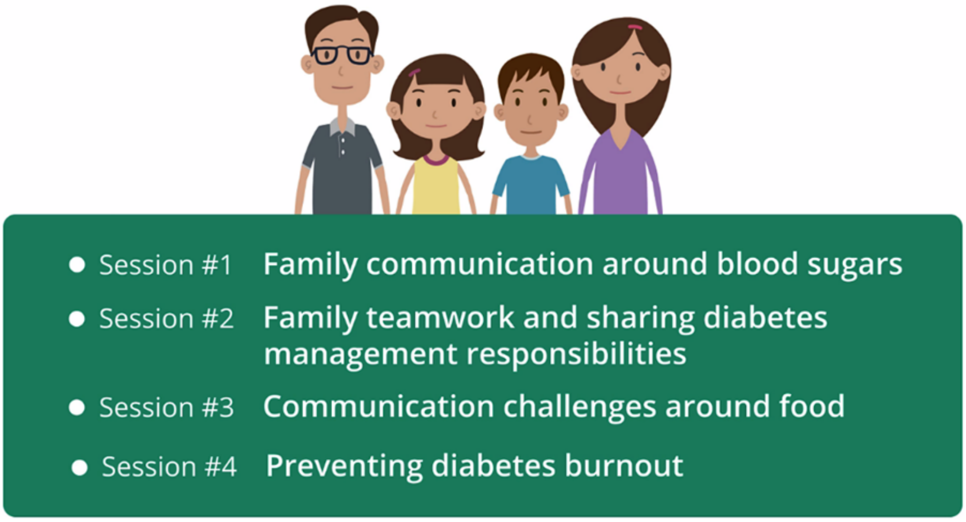
Figure 1. Family Teamwork Online module flow.



Session 1 only; Sessions 2-4 = goal review

Sessions 1-3 only; Session 4 = program summary

Figure 2. Family Teamwork Online module topics.



As part of the development work [34], parents and adolescents were interviewed separately to elicit their opinions about what the Ragt should look like, including appearance, sex, age, and clothing. They were also asked for suggestions regarding names for the Ragt. Using this information as a guide, a female Ragt named Ashley, with a professional, yet approachable, appearance was created (Figure 3). She was animated and designed to mimic in-person delivery by a human diabetes care provider. Prerecorded vocal segments by a professional voice actor were synced to Ashley's character, enabling her to voice the scripted segments and convey emotion, concern, and respect for the families through vocalizations, facial expressions, and body movements. Consistent with past feedback, Ashley was designed to be "exotic" (ie, she was designed so that she appeared to be of no particular race or ethnicity), thus enabling families to assign their preferred race/ethnicity to her. In addition to Ashley, there were eight other characters—a family, consisting of a father, mother, son, and daughter, and four adult characters, consisting of a male and female physician, a teacher, and a coach (Figure 4).

Each module lasted approximately 20 minutes and focused on a different topic relevant to family communication around diabetes management (Figure 2). Guided by the flow diagram (Figure 1), modules were delivered in a set format that included topic introduction by Ashley; typical family scenarios common to a particular problem or issue relevant to the topic of a particular module; animatics, where typical family reactions (ranging from ideal to less than ideal) to common parent/adolescent situations arising around T1D were portrayed; and interactive components, where Ashley posed questions and the parent/adolescent dyad selected responses from a preselected list of responses, followed by a summary delivered by Ashley regarding the question or scenario. Each module ended with a joint parent/adolescent goal-setting task related to the module topic and family communication around diabetes management. Families were provided with several goals from which they could select. They could then print their goal and a tip sheet. At the beginning of the next module, the family reported goal attainment and received feedback from Ashley. Following completion of the module, they also received an email with the goal they selected and the tip sheet to minimize potential problems associated with not having ready access to a printer.

The program was designed to be viewed over a high-speed internet connection on a variety of devices (eg, desktop computer, laptop, mobile device). It was hosted over a secure, password-protected website. Parents and adolescents received separate passwords with which to log on and view the module. For the initial viewing of module, parents and adolescents were required to log on together. After completing the initial viewing, parents and adolescents could log on separately to view previously completed modules unlimited times. Families were eligible to view a new module and set a new goal 2 weeks after completing the previous module. If a family logged on before they were eligible to view the next module, they had access to previously completed modules only. Families were given approximately 3 months to complete the program.

A secure, password-protected database associated with FTO recorded each family's information as they navigated the program, including log-ons, responses, and module completion. The researcher accessed the database through a secure dashboard that enabled them to view each family's progress through the program including log-ons, responses, and goals. After completion of baseline data collection, the intervention coordinator entered the family into the database and assigned passwords. To protect confidentiality, each family received a unique study identifier, and this identifier, rather than their actual names, was entered into the database. A master list of actual names, identifiers, and passwords was stored on a secure, password-protected server available only to the study team.

Figure 3. Image of Ashley, the relational agent in Diabetes Family Teamwork Online.

Figure 4. Cast of characters.

Recruitment

Families were recruited from a diabetes clinic at a large, tertiary care hospital for children. The study coordinator identified eligible families using the upcoming clinic schedule and the electronic medical record. Inclusionary criteria for the adolescents included being a current patient, aged 10 to 15 years old, diagnosed with T1D for at least 1 year, with English fluency, having high speed internet access, staying in the area for the duration of the study, and having a parent or legal guardian willing to participate in the study. Exclusionary criteria for the adolescents included having an average HbA_{1c} over the past year greater than 12% or less than 7%, inability to attend regular clinic visits, unwillingness to have interviews audio recorded, or having a physical or mental disease or condition that would conflict with study activities.

Inclusionary criteria for the parents included being the primary diabetes caregiver of the adolescent participating in the study, willingness to participate in study activities, with English fluency, having access to high speed internet, and staying in the area for the duration of the study. Exclusionary criteria included inability to attend regular clinic visits, unwillingness to have interviews audio recorded, or having a physical or mental disease or condition that would conflict with study activities.

Eligible families were sent a letter and flyer in the mail notifying parents that a study coordinator would meet the family at their upcoming clinic visit to speak about the study. Prior to the clinic visit, a study coordinator contacted the parent, informed them about the study, and if the parent was interested, screened for eligibility. If the parent was not available by phone, a study coordinator met the family at their clinic visit, described the study, and if the family was interested, screened for eligibility. Parents provided written informed parental consent and each child assented prior to study participation. The institutional review board at Baylor College of Medicine approved the protocol (H-37245).

Intervention Procedures

After completing baseline data collection, the intervention coordinator sent the parent and adolescent separate emails with the website link, username, and private password. Families were provided with a program guide with instructions on how to use the website, including how to log on, how to view the modules, how to click the Back and Next buttons, and how to print the goal and tip sheets.

After the family completed a module, the intervention coordinator sent a thank you email to the parent’s email address with the available date for the next module and a reminder that the family could log on between modules and view the previously completed modules. The goal (set by the family) and tip sheets for the module were also attached to the thank you email. The intervention coordinator kept track of the families logging on to the modules. If families were not logging on regularly to watch the next eligible module, the coordinator sent reminder emails to both the parent and adolescent and followed up with a phone call after the initial reminder email.

Data Collection Procedures

Several types of data were collected: self-report data at baseline, post 1 (immediately after completing the intervention), and post 2 (3 months after post 1); interviews (post 1, post 2); intervention (ie, log-ons, in-module responses); and staff logs (Table 2). Parents and adolescents each received monetary incentives for completing surveys and/or interviews at each data collection time point ($50 baseline, $70 post 1, $80 post 2). No incentives were provided for completing intervention modules.

Table 2. Data sources.

Type	Who	Method	Recruitment	Baseline	Intervention	Post 1	Post 2
Self-report	Parent; adolescent	Online		✓		✓	✓
Interviews	Parent; adolescent	Telephone				✓	✓
Module	Family	Backend database			✓		
Logs	Staff	Tracking system	✓	✓	✓	✓	✓

Feasibility Criteria

A priori feasibility criteria were established to guide the determination of whether FTO was feasible and acceptable [34]. Minor adjustments were made to the criteria to account for differences in measures or scoring approaches. Final criteria used to assess feasibility were (1) recruitment goals were met (n=20), (2) families completed ≥75% of the modules, (3) attrition rate was ≤10%, (4) program satisfaction was high (≥80% of families), (5) therapeutic alliance with the Ragt was high (average score of ≥60/84), (6) families expressed positive attitudes toward the Ragt (average item score of ≥5 on ≥4 items), (7) ≥80% of data were collected at post 1 and post 2, and (8) few technical issues (<10%) occurred during intervention delivery.

Measures

Recruitment was assessed with logs maintained by the research staff (eg, number of eligible families with clinic visits in the recruitment period, number of eligible families who enrolled and reasons, number of eligible families who did not enroll and reasons). The recruitment goal was 20 families. This criterion was assessed by comparing the number of eligible families who enrolled to the criterion.

Module completion was defined as parent and adolescent logging on to the program website at the same time to complete the module activities. This information was automatically collected when families logged on and interacted with the program. Module completion was calculated by dividing the total number of modules completed by the family by the total number of available modules.

Attrition rate was defined as the number of families who did not complete the intervention (ie, all 4 modules) and all data collection activities (ie, baseline, post 1, post 2). It was calculated by dividing the total number of families who did not complete all these activities by the total number of families who enrolled in the study.

Program satisfaction was assessed among parents and adolescents at post 1 with a 10-item measure of satisfaction used in previous studies [30,35,36]. Items were rated using a 3-point Likert scale (yes, not sure, no). Responses were summed to determine overall program satisfaction. Scores could range from 10-30, with lower scores representing higher satisfaction. Separate scores were calculated for parents and adolescents.

Therapeutic alliance was assessed among parents and adolescents at post 1 with the Bond subscale of the Working Alliance Inventory, a 12-item measure assessing trust and belief the participant can work with their provider (ie, Ragt) to achieve the desired outcomes [37]. Items were rated on a 7-point Likert scale, ranging from always hard to always easy. For 3 items, the desired response was never instead of always; therefore, these items were reverse coded, with never being the desired response. Responses were summed to create a total scale score. The total score could range from 12 to 84, with higher scores representing greater therapeutic alliance. Separate scores were computed for parents and adolescents.

Attitudes toward the Ragt were assessed among parents and adolescents at post 1 with a 6-item survey adapted from Bickmore et al [27]. Attitudes were assessed with items addressing different dimensions of the Ragt, including usability, continuance, relationship, preference, adherence, and satisfaction [27]. Items were rated on a 7-point scale, ranging from less desirable to more desirable attitudes (Table 3). Response scales differed for items; therefore, responses were scored and reported individually. Higher scores represented more positive attitudes.

Data completeness was defined as the percentage of families with complete self-report data at all assessment points (baseline, post 1, post 2).

Technical issues were defined as the number of incidents reported by families that limited access to the program and were determined from a review of staff logs.

Table 3. Attitude scale (adapted from Bickmore et al [27]).

Attitude	Questions	Response anchors
Usability	How easy was it to “talk” with Ashley?	1 (always hard) to 7 (always easy)
Continuance	How much would you like to continue working with Ashley?	1 (not at all) to 7 (very much would like)
Relationship	How would you describe your relationship with Ashley?	1 (complete stranger) to 7 (always a friend)
Preference	Would you have rather talked with your doctor or nurse rather than Ashley?	1 (definitely prefer doctor or nurse) to 7 (definitely prefer Ashley)
Adherence	How likely is it that you will follow Ashley’s advice?	1 (very unlikely) to 7 (very likely)
Satisfaction	How satisfied were you with Ashley?	1 (very dissatisfied) to 7 (very satisfied)

To obtain an understanding of parent and adolescent perceptions toward the intervention, telephone interviews were conducted following completion of post 1 and post 2 data collection surveys. The purpose of the post 1 interview was to identify perceptions related to the intervention, while the post 2 interview explored perceptions related to maintenance. Only post 1 interviews will be reported here to provide a deeper understanding of program feasibility and acceptability from the perspective of the adolescents and parents who participated in the study. Interviews were scripted and conducted by trained interviewers. Probes and prompts were used as needed to expand, clarify, and understand responses.

Statistical Analyses

Descriptive statistics were calculated on survey data. Separate analyses were calculated for parents and adolescents. Results were compared to the criteria to make a determination of feasibility. Audio recordings of interviews were reviewed and key point summaries created to capture important thoughts that emerged from adolescent and parent interviews to enhance understanding of results [38]. Verbatim quotes were used to provide voice to adolescent and parent perceptions and experiences.

Results

Participants

Parents were mostly female (25/26, 96%), white (21/26, 81%), non-Hispanic (22/26, 85%), and married or living with significant other (22/26, 85%). Families had college-level education or higher (20/26, 77%) and reported household incomes of greater than $61,000 (20/26, 80%; Table 4).Adolescents were mostly female (19/26, 73%), white (22/26, 85%), and non-Hispanic (21/26, 81%). There were slightly more adolescents aged 13 to 15 years in the study (14/26, 54%). Disease duration ranged from 2 (4/26, 15%) to 13 years (1/26, 4%; Table 5).

Table 4. Descriptive statistics of parents who completed the Diabetes Family Teamwork Online study (n=26).

Characteristic	Value, n (%)
Gender	
Male	1 (4)
Female	25 (96)
Race/ethnicity	
American Indian/Alaskan Native	1 (4)
Black	3 (12)
White	21 (81)
Other	1 (4)
Hispanic	
Hispanic	4 (15)
Non-Hispanic	22 (85)
Marital status	
Married/living with significant other	22 (85)
Single, never married	2 (3)
Divorced, separated, or widowed	1 (4)
Other	1 (4)
Highest household education	
High school graduate or equivalent	3 (12)
Technical school	1 (4)
Some college	2 (8)
College graduate	11 (42)
Postgraduate study	9 (35)
Household income[a]	
$21,000-$41,000	3 (12)
$42,000-$61,000	2 (8)
Greater than $61,000	20 (80)

[a] One parent did not report household income.

Table 5. Descriptive statistics of adolescents who completed the Diabetes Family Teamwork Online study (n=26).

Characteristic	Value, n (%)
Gender	
Male	7 (27)
Female	19 (73)
Race/ethnicity	
American Indian/Alaskan Native	1 (4)[a]
Black	3 (12)
White	22 (85)
Hispanic	
Hispanic	5 (19)
Non-Hispanic	21 (81)
Age (years)	
10-12	12 (46)
13-15	14 (54)
Disease duration (years)	
2	4 (15)
3	8 (31)
4	2 (8)
5	3 (12)
6	3 (12)
7	1 (4)
8	2 (8)
10	1 (4)
12	1 (4)
13	1 (4)

[a] Participant also identifies as white.

Feasibility Outcomes

Feasibility criteria and status are summarized in Table 6. Feasibility of each criterion is briefly described below.

Recruitment began in October 2016 and ended in December 2016. A review of clinic records indicated that 48 families met general eligibility criteria and were notified of the study and their eligibility to participate. Of these, 5 declined to participate, 9 could not be contacted, 3 were enrolled in another study that made them ineligible for our study, and 4 were not included for reasons ranging from clinic visit scheduled after recruitment closed (2), not meeting inclusionary criteria (1), and not returning recruitment packet (1). Of the eligible families, 27 enrolled, completed baseline data collection, and were allocated to the intervention (Figure 5). Our recruitment goal was 20 families; therefore, this criterion was exceeded.

FTO consisted of 4 modules. All 4 modules were completed by 26 out of 27 families enrolled in this study. The feasibility criterion was >75%; therefore, this criterion was met.

Of the 27 families who enrolled in the study, 26 completed the intervention and all data collection (baseline, post 1, post 2) and intervention activities. The feasibility criterion was ≤10%; therefore, this criterion was met.

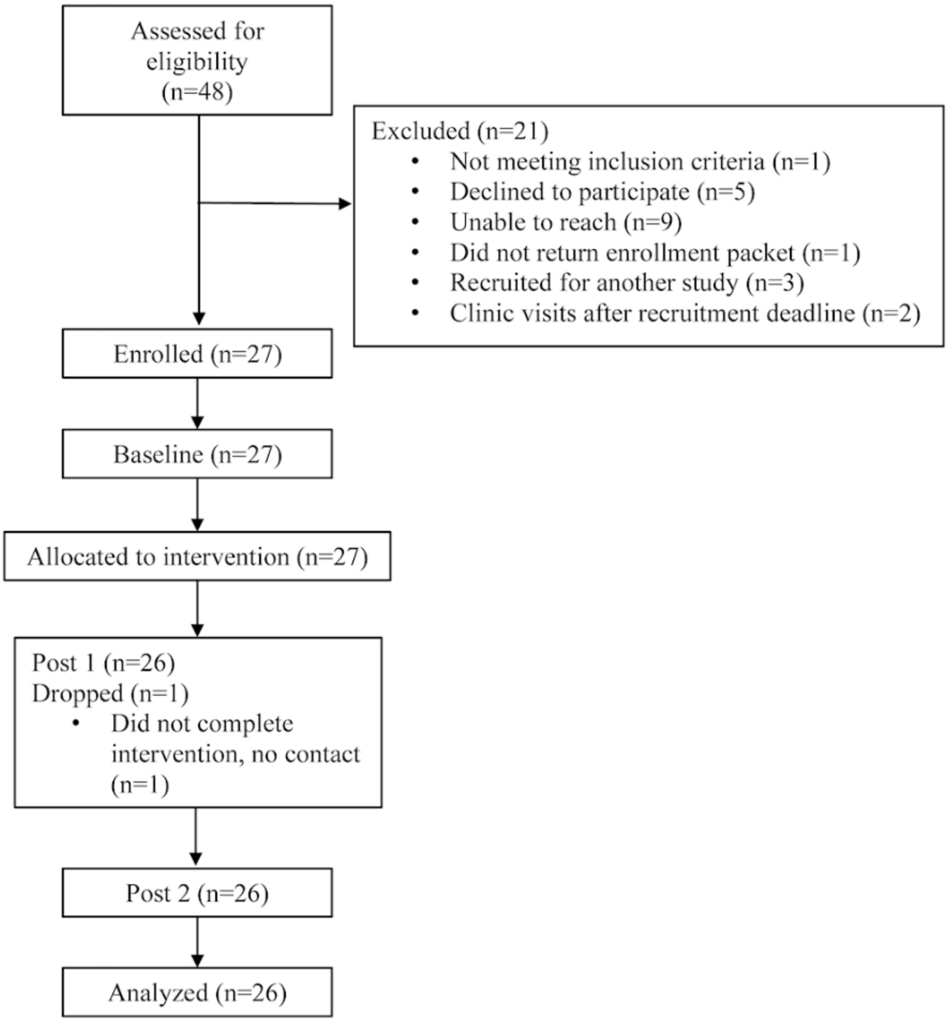
The mean program satisfaction scores were 11.08 (SD 1.32) for adolescents and 10.46 (SD 0.65) for parents, indicating that over 90% of adolescents and parents reported high program satisfaction. The feasibility criterion was ≥80% of adolescents and parents reported high program satisfaction. Therefore, this criterion was met, with both groups reporting high program satisfaction.

The mean therapeutic alliance score for adolescents was 64.58 (SD 14.43) and 64.65 (SD 14.33) for parents. The feasibility criterion was an average total scale score of ≥60/84. Therefore, this criterion was met for both adolescents and parents.

Table 6. Feasibility criteria.

Criterion	Standard	Status
Recruitment	20 families	Exceeded
Module completion	≥75% modules	Met
Attrition rate	≤10%	Met
Program satisfaction	≥80%	Met
Therapeutic alliance	Average score of ≥60/84	Met
Attitudes toward Ragt	Score of ≥5 on ≥4 items	Met
Data completeness	≥80%	Met
Technical issues	≤10%	Met

Figure 5. Family Teamwork Online CONSORT (Consolidated Standards of Reporting Trials) diagram.



Both adolescents and parents reported attitude scores toward Ashley of at least 5 on 5 of the 6 attitude items, respectively: usability (5.85 [SD 1.43], 5.50 [SD 1.39]), continuance (5.42 [SD 1.17], 5.23 [SD 1.03]), relationship (5.42 [SD 1.55], 5.08 [SD 1.52]), preference (2.73 [SD 1.59], 3.54 [SD 1.49]), adherence (5.73 [SD 1.12], 6.00 [SD 0.69]), and satisfaction (6.0 [SD 1.02], 5.96 [SD 0.72]). Therefore, this criterion was met.

Twenty-six of the 27 families had complete self-report data at all assessment points. The feasibility criterion was ≥80%; therefore, this criterion was met.

No technical issues were reported by the families. The feasibility criterion was “few technical issues”, defined as <10%. Therefore, this criterion was met.

The qualitative findings generally supported the survey results. Most adolescents and parents liked the program. They generally thought the content was appropriate, easy to understand, and realistic.

> *Well, I thought it was useful in what we deal with every day—high blood sugars or the low blood sugars or what we need to be doing better when it comes to managing [daughter's name] diabetes. The examples were pretty common to what we deal with, so it helped to see it being played out for us and how families responded and if our response was appropriate for that situation. I thought it was helpful in helping us understand how to react with the high blood sugar or low blood sugar or just certain situations we deal with every day.* [Parent]

They thought that the module number and length were appropriate.

> *I think they were just about right. I mean, I liked them as they were. I wouldn't have liked them to be longer or shorter.* [Adolescent]

This sentiment seemed to be echoed by parents.

> *I liked the time. I didn't think they were too long or too short 'cause it was so detailed with every session, but it didn't take a lot of my time or my child's.* [Parent]

A few families wanted longer or more modules on certain topics. For example, a parent suggested adding information on tricky topics.

> *I think four is okay. I think for some of the more tricky topics, maybe adding a follow-up for that that goes into more detail might be helpful.* [Parent]

A adolescent expressed also wanting more information.

> *I think it was too little...6 or 7...cause I wanted to learn more about what they were talking about.* [Adolescent]

Parents and adolescents also seemed to like the family focus taken by the program.

> *I liked it. It was pretty fun experiencing it with my mom and how we got to like work out my diabetes better. It was pretty fun.* [Adolescent]

Overall, parents shared positive feedback on the Ragt, including her appearance, responses, and voice.

> *I thought she was very helpful. She was very helpful in answering things, and like for something quick, like a quick question or whatever, I mean her response to things was better than having a doctor right there in your face. She was professional. She was child friendly and adult friendly.* [Parent]

Adolescents had similar reactions.

> *I think she was helpful, and I think that—I don't know if it's just me—since she's a girl, it's easier to listen to someone the same as me.* [Adolescent]

A few adolescents expressed that although they liked Ashley, she had limitations.

> *I think she was fine, but I would rather have an actual person...because Ashley didn't really give that much detail.* [Adolescent]

Families were also asked what information in the program was most helpful or useful. Parents shared a variety of information that was most helpful for them, including the information presented on parent reactions to high and low blood sugars, not referring to blood sugars as good and bad, communication between the parent and adolescent, setting goals, and diabetes burnout. Two parent quotes eloquently summarize parent thoughts regarding what they perceived as most helpful:

> *Well for me it would be the reminder to let the teen take charge and let the parent sort of step back and say what would you do differently instead of managing it myself.* [Parent]
>
> *That [diabetes burnout] was probably the best one for us because we really didn't know a lot about that.* [Parent]

Adolescents had similar reactions and identified a variety of topics that were helpful. They seemed to particularly focus on issues related to family communication around diabetes management.

> *How to control like no blaming and everyone's part of the team.* [Adolescent]

The focus on using different language was also viewed as important.

> *The most helpful activity was the one where you had to use different words instead of it was like a bad or good blood sugar, it was a high or a low.* [Adolescent]

When asked what was least helpful in the program, few parents or adolescents had many suggestions to offer. Most thought the program was helpful.

> *I don't really think there was just anything so unhelpful that I would leave it out.* [Parent]

This sentiment was echoed by adolescents.

> *I don't know. I don't think anything was bad about the program. I liked it myself, so I don't think anything was least helpful. Everything was pretty helpful to me.* [Adolescent]

A few suggestions were offered, however.

> *For me, it was all a good reminder, but the probably least helpful was looking at the numbers as good or bad. I always try to look at them as information...that was probably the least helpful because I kind of already look at it that way.* [Parent]
>
> *Taking the snack and checking your blood sugars and stuff like that when you're exercising.* [Adolescent]

Discussion

Principal Findings

This paper reports on the feasibility of delivering a program to enhance communication around blood glucose management tasks in families with an adolescent who has T1D. Feasibility criteria were met or exceeded. Therefore, we concluded this program was a feasible method for promoting enhanced communication around diabetes management tasks.

Recruitment goals for this study were exceeded in a relatively short period of time. Conflict around diabetes management tasks in families with adolescents who have T1D and its association with poorer diabetes outcomes have been well documented in the literature [9,39,40]. There is also clear evidence that both parents and adolescents perceive there is conflict around diabetes management [41,42]. The short time period within which recruitment goals were exceeded suggests that parents and adolescents are interested in learning to manage or reduce this conflict.

Participation by families was high, as indicated by the low attrition rate and high program completion rates. This suggests that families found the program helpful and a beneficial use of their time. This is not surprising, however, given the success of FT, the program on which FTO was based [9,12,13]. This finding also adds to the body of literature supporting communication technologies as a potentially effective way to extend the reach of diabetes management programs for adolescents [16]. Future research is needed to assess the efficacy of this approach.

Therapeutic alliance is the emotional dimension associated with working with a provider and their perceived ability to help the patient achieve a desired outcome [37]. Although parents and adolescents were aware that Ashley was a digital image and not an actual person, therapeutic alliance was established. This was similar to results reported by Bickmore et al [43] when a Ragt was used to encourage exercise in a healthy population and deliver hospital discharge instructions to patients with depressive symptoms [27]. Given the differences between initiating a healthy behavior such as exercise and receiving instructions for posthospital care versus potential consequences of poor glycemic control [44], this suggests that Ragts may be an acceptable way to extend care for a variety of diverse behaviors outside of traditional settings. In the management of a chronic condition, such as T1D, patient-provider communication is an essential component of care [45]. Therefore, our goal was not to supplant this relationship but to extend care outside the clinical setting. Although this study suggests Ragts are a feasible and acceptable method of achieving this goal, future research is needed to examine ways in which Ragts may further extend the patient-provider relationship outside the clinical setting. One way to achieve this may be to design Ragts that respond emphatically in real time, based on the participant’s expressed emotions [46]. Further, as suggested by the qualitative results, it may be that therapeutic alliance takes on a different form with a Ragt. Future research is needed to more fully investigate this possibility.

Both parents and adolescents reported positive attitudes regarding ease of use, desire to continue working with Ashley, their relationship with Ashley, the likelihood of adhering to her advice, and satisfaction with her. The exception was their stronger preference for talking with their doctor or nurse, rather than Ashley, about T1D. As discussed above, given that the patient-provider relationship is essential to the care of T1D, particularly in adolescents [45], this finding is not surprising. The goal when working with chronic disease management should not be to supplant or interfere with the patient-provider relationship but to augment and support it and extend care beyond the clinical setting. Given that this is the first study to use Ragts to enhance communication around blood glucose management tasks in families with an adolescent who has T1D, it adds important information to the literature regarding expectations around chronic disease management and ways in which to augment and extend care between visits with the health care provider. Future research with providers may provide insight into ways Ragts could be used to further extend care into the home environment.

Although missing data are a common problem in research that can present challenges in the analytic stage [47], we had complete data on all but one family. Consistent with guidelines developed to minimize missing data [48], data collection was online, making participation convenient, and the families received notifications when they were eligible to participate in data collection activities. We also ensured contact information for each family was current, study coordinators were experienced and well trained in study procedures, participation in data collection activities was monitored and tracked, and incentives were provided to the families for completing data collection at each assessment point. All of these activities likely contributed to the high data collection rates.

Finally, families reported no technical issues accessing or using the program to the intervention team. This is likely due to working with a highly skilled local Web design company with whom we have a history of successful collaboration. Further, experience has taught us the importance of knowing in advance how the individual pieces of the intervention fit together, as well as the overall program flow. A template was created for the modules, and this template guided each module. This consistency may have contributed to fewer technical glitches. A program guide was also created and shared with families prior to their viewing of the first module, which may have further contributed to fewer technical issues. Because technical issues can be annoying and may disrupt program participation, particularly when they occur multiple times or at critical points in program delivery, the lack of technical issues may have also contributed to our high program participation and satisfaction ratings.

Limitations

Although this study has strengths, including being an adaptation of an intervention previously found to be efficacious and its theoretical grounding, it is not without limitations. The small sample size and one-group design limits its ability to find statistical significance in key outcome variables, such as family conflict and communication around blood glucose. However, this was designed to be a feasibility study, with the goals of assessing the viability of this approach, needed changes, and whether it should move forward for an assessment of efficacy with a larger, more fully powered randomized controlled trial [49]. Parents and adolescents in the study were predominantly female and the Ragt was female, which may have influenced the appeal of this approach to families who participated in the study. However, the appearance, sex, and age of the Ragt designed for this study was informed by formative research with parents and adolescents other than those who participated in the feasibility study; this suggests the Ragt may have broad appeal regardless of sex. This study was also conducted in a limited geographic region with a sample of families who were predominately white, of higher socioeconomic status, and well educated, limiting the ability to draw conclusions about the generalizability of this approach with other populations. Finally, because this study uses technology, it may not be appropriate for use in locations that may have limitations regarding internet access or speed, such as some rural areas, and it may have broader appeal among families with adolescents who use technology to facilitate blood glucose management (eg, continuous glucose monitoring, insulin pumps).

Conclusion

An online intervention to enhance communication around blood glucose management task in families with an adolescent who has diabetes was found to be feasible and acceptable. The contribution of this research to the literature is that it demonstrates that Ragts are a feasible adjunct to traditional care and may be a way to extend care outside the clinical setting. Future research is needed to identify the effects on family communication, quality of life, blood glucose, and diabetes management in a larger, more fully powered sample.

Acknowledgments

This work was supported by the US Department of Agriculture/Agricultural Research Service (USDA/ARS) and the Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine (Houston, Texas), funded in part with federal funds from the USDA/ARS under Cooperative Agreement No 58-3092-5-001. This project was also supported by National Institute of Diabetes and Digestive and Kidney Diseases, grant #DK106505 (to BJA and DT). The contents of this paper do not necessarily reflect the views or policies of the USDA, and the mention of trade names, commercial products, or organizations does not imply endorsement from the US Government. We would like to thank Jerald Reichstein, Jackson Lau, and the team at Bouncing Pixel for their assistance in developing the program, as well as the parents and adolescents who participated in this study.

Conflicts of Interest

None declared.

References

1. Maahs DM, West NA, Lawrence JM, Mayer-Davis EJ. Epidemiology of type 1 diabetes. Endocrinol Metab Clin North Am 2010 Sep;39(3):481-497 [FREE Full text] [doi: 10.1016/j.ecl.2010.05.011] [Medline: 20723815]
2. Halfon N, Newacheck PW. Evolving notions of childhood chronic illness. JAMA 2010 Feb 17;303(7):665-666. [doi: 10.1001/jama.2010.130] [Medline: 20159877]
3. Miller GF, Coffield E, Leroy Z, Wallin R. Prevalence and costs of five chronic conditions in children. J Sch Nurs 2016 Oct;32(5):357-364 [FREE Full text] [doi: 10.1177/1059840516641190] [Medline: 27044668]
4. Rogers MAM, Kim C, Banerjee T, Lee JM. Fluctuations in the incidence of type 1 diabetes in the United States from 2001 to 2015: a longitudinal study. BMC Med 2017 Dec 08;15(1):199 [FREE Full text] [doi: 10.1186/s12916-017-0958-6] [Medline: 29115947]
5. Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 1993 Sep 30;329(14):977-986. [doi: 10.1056/NEJM199309303291401] [Medline: 8366922]
6. Narayan KM, Gregg EW, Engelgau MM, Moore B, Thompson TJ, Williamson DF, et al. Translation research for chronic disease: the case of diabetes. Diabetes Care 2000 Dec;23(12):1794-1748 [FREE Full text] [doi: 10.2337/diacare.23.12.1794] [Medline: 11128355]
7. Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group, Beck RW, Buckingham B, Miller K, Wolpert H, Xing D, et al. Factors predictive of use and of benefit from continuous glucose monitoring in type 1 diabetes. Diabetes Care 2009 Nov;32(11):1947-1953 [FREE Full text] [doi: 10.2337/dc09-0889] [Medline: 19675206]
8. Lansing AH, Crochiere R, Cueto C, Wiebe DJ, Berg CA. Mother, father, and adolescent self-control and adherence in adolescents with Type 1 diabetes. J Fam Psychol 2017 Jun;31(4):495-503 [FREE Full text] [doi: 10.1037/fam0000292] [Medline: 28080079]
9. Anderson BJ, Vangsness L, Connell A, Butler D, Goebel-Fabbri A, Laffel LMB. Family conflict, adherence, and glycaemic control in youth with short duration Type 1 diabetes. Diabet Med 2002 Aug;19(8):635-642. [Medline: 12147143]
10. Hood KK, Rohan JM, Peterson CM, Drotar D. Interventions with adherence-promoting components in pediatric type 1 diabetes: meta-analysis of their impact on glycemic control. Diabetes Care 2010 Jul;33(7):1658-1664 [FREE Full text] [doi: 10.2337/dc09-2268] [Medline: 20587726]
11. Anderson BJ, Brackett J, Ho J, Laffel LM. An office-based intervention to maintain parent-adolescent teamwork in diabetes management. Impact on parent involvement, family conflict, and subsequent glycemic control. Diabetes Care 1999 May;22(5):713-721 [FREE Full text] [Medline: 10332671]

12. Laffel LMB, Connell A, Vangsness L, Goebel-Fabbri A, Mansfield A, Anderson BJ. General quality of life in youth with type 1 diabetes: relationship to patient management and diabetes-specific family conflict. Diabetes Care 2003 Nov;26(11):3067-3073. [Medline: 14578241]
13. Laffel LMB, Vangsness L, Connell A, Goebel-Fabbri A, Butler D, Anderson BJ. Impact of ambulatory, family-focused teamwork intervention on glycemic control in youth with type 1 diabetes. J Pediatr 2003 Apr;142(4):409-416. [doi: 10.1067/mpd.2003.138] [Medline: 12712059]
14. King DK, Glasgow RE, Leeman-Castillo B. Reaiming RE-AIM: using the model to plan, implement, and evaluate the effects of environmental change approaches to enhancing population health. Am J Public Health 2010 Nov;100(11):2076-2084. [doi: 10.2105/AJPH.2009.190959] [Medline: 20864705]
15. Lorig K, Ritter PL, Laurent DD, Plant K, Green M, Jernigan VBB, et al. Online diabetes self-management program: a randomized study. Diabetes Care 2010 Jun;33(6):1275-1281 [FREE Full text] [doi: 10.2337/dc09-2153] [Medline: 20299481]
16. Stinson J, Wilson R, Gill N, Yamada J, Holt J. A systematic review of internet-based self-management interventions for youth with health conditions. J Pediatr Psychol 2009 Jun;34(5):495-510 [FREE Full text] [doi: 10.1093/jpepsy/jsn115] [Medline: 19029142]
17. Gerber BS, Solomon MC, Shaffer TL, Quinn MT, Lipton RB. Evaluation of an internet diabetes self-management training program for adolescents and young adults. Diabetes Technol Ther 2007 Feb;9(1):60-67. [doi: 10.1089/dia.2006.0058] [Medline: 17316099]
18. Mulvaney SA, Rothman RL, Osborn CY, Lybarger C, Dietrich MS, Wallston KA. Self-management problem solving for adolescents with type 1 diabetes: intervention processes associated with an Internet program. Patient Educ Couns 2011 Nov;85(2):140-142 [FREE Full text] [doi: 10.1016/j.pec.2010.09.018] [Medline: 21030194]
19. Mulvaney SA, Rothman RL, Wallston KA, Lybarger C, Dietrich MS. An internet-based program to improve self-management in adolescents with type 1 diabetes. Diabetes Care 2010 Mar;33(3):602-604 [FREE Full text] [doi: 10.2337/dc09-1881] [Medline: 20032275]
20. Whittemore R, Grey M, Lindemann E, Ambrosino J, Jaser S. Development of an Internet coping skills training program for teenagers with type 1 diabetes. Comput Inform Nurs 2010;28(2):103-111 [FREE Full text] [doi: 10.1097/NCN.0b013e3181cd8199] [Medline: 20182161]
21. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet 1998 Sep 12;352(9131):837-853. [Medline: 9742976]
22. Cushing CC, Steele RG. A meta-analytic review of eHealth interventions for pediatric health promoting and maintaining behaviors. J Pediatr Psychol 2010 Oct;35(9):937-949 [FREE Full text] [doi: 10.1093/jpepsy/jsq023] [Medline: 20392790]
23. Bickmore T. Relational agents for chronic disease self management. In: Hayes BM, Aspray W, editors. Health Informatics: A Patient-Centered Approach to Diabetes. Cambridge: MIT Press; 2010:181-204.
24. Bickmore T, Giorgino T. Some novel aspects of health communication from a dialogue systems perspective. 2004 Presented at: Proceedings of the AAAI Fall Symposium on Dialogue Systems for Health Communication; 2004; Arlington.
25. Bickmore TW, Schulman D, Sidner CL. A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. J Biomed Inform 2011 Apr;44(2):183-197 [FREE Full text] [doi: 10.1016/j.jbi.2010.12.006] [Medline: 21220044]
26. Bickmore T, Gruber A. Relational agents in clinical psychiatry. Harv Rev Psychiatry 2010;18(2):119-130. [doi: 10.3109/10673221003707538] [Medline: 20235777]
27. Bickmore TW, Mitchell SE, Jack BW, Paasche-Orlow MK, Pfeifer LM, Odonnell J. Response to a relational agent by hospital patients with depressive symptoms. Interact Comput 2010 Jul 1;22(4):289-298 [FREE Full text] [doi: 10.1016/j.intcom.2009.12.001] [Medline: 20628581]
28. Puskar K, Schlenk EA, Callan J, Bickmore T, Sereika S. Relational agents as an adjunct in schizophrenia treatment. J Psychosoc Nurs Ment Health Serv 2011 Aug;49(8):22-29. [doi: 10.3928/02793695-20110705-01] [Medline: 21766736]
29. Watson A, Bickmore T, Cange A, Kulshreshtha A, Kvedar J. An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial. J Med Internet Res 2012;14(1):e1 [FREE Full text] [doi: 10.2196/jmir.1629] [Medline: 22281837]
30. Thompson D, Bhatt R, Lazarus M, Cullen K, Baranowski J, Baranowski T. A serious video game to increase fruit and vegetable consumption among elementary aged youth (Squire's Quest! II): rationale, design, and methods. JMIR Res Protoc 2012;1(2):e19 [FREE Full text] [doi: 10.2196/resprot.2348] [Medline: 23612366]
31. Fogg B. Persuasive Technology: Using Computers to Change What We Think and Do. 1st Edition. San Francisco: Morgan Kaufmann Publishers; 2002.
32. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
33. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs: Prentice Hall; 1986.
34. Thompson D, Cullen KW, Redondo MJ, Anderson B. Use of relational agents to improve family communication in type 1 diabetes: methods. JMIR Res Protoc 2016;5(3):e151 [FREE Full text] [doi: 10.2196/resprot.5817] [Medline: 27468762]

35. Thompson D, Mahabir R, Bhatt R, Boutte C, Cantu D, Vazquez I, et al. Butterfly Girls; promoting healthy diet and physical activity to young African American girls online: rationale and design. BMC Public Health 2013;13:709 [FREE Full text] [doi: 10.1186/1471-2458-13-709] [Medline: 23915235]
36. Cullen KW, Thompson D, Boushey C, Konzelmann K, Chen T. Evaluation of a web-based program promoting healthy eating and physical activity for adolescents: teen choice: food and fitness. Health Educ Res 2013 Aug;28(4):704-714 [FREE Full text] [doi: 10.1093/her/cyt059] [Medline: 23748162]
37. Horvath AO, Greenberg LS. Development and validation of the working alliance inventory. J Couns Psychol 1989;36(2):223-233. [doi: 10.1037/0022-0167.36.2.223]
38. Schneider M, Hall WJ, Hernandez AE, Hindes K, Montez G, Pham T, HEALTHY Study Group. Rationale, design and methods for process evaluation in the HEALTHY study. Int J Obes (Lond) 2009 Aug;33 Suppl 4:S60-S67 [FREE Full text] [doi: 10.1038/ijo.2009.118] [Medline: 19623191]
39. Hilliard ME, Holmes CS, Chen R, Maher K, Robinson E, Streisand R. Disentangling the roles of parental monitoring and family conflict in adolescents' management of type 1 diabetes. Health Psychol 2013 Apr;32(4):388-396 [FREE Full text] [doi: 10.1037/a0027811] [Medline: 22545980]
40. Williams LB, Laffel LMB, Hood KK. Diabetes-specific family conflict and psychological distress in paediatric Type 1 diabetes. Diabet Med 2009 Sep;26(9):908-914 [FREE Full text] [doi: 10.1111/j.1464-5491.2009.02794.x] [Medline: 19719712]
41. Ye CY, Jeppson TC, Kleinmaus EM, Kliems HM, Schopp JM, Cox ED. Outcomes that matter to teens with type 1 diabetes. Diabetes Educ 2017 Dec;43(3):251-259 [FREE Full text] [doi: 10.1177/0145721717699891] [Medline: 28520550]
42. Weissberg-Benchell J, Nansel T, Holmbeck G, Chen R, Anderson B, Wysocki T, Steering Committee of the Family Management of Diabetes Study. Generic and diabetes-specific parent-child behaviors and quality of life among youth with type 1 diabetes. J Pediatr Psychol 2009 Oct;34(9):977-988 [FREE Full text] [doi: 10.1093/jpepsy/jsp003] [Medline: 19270028]
43. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
44. Rausch JR, Hood KK, Delamater A, Shroff PJ, Rohan JM, Reeves G, et al. Changes in treatment adherence and glycemic control during the transition to adolescence in type 1 diabetes. Diabetes Care 2012 Jun;35(6):1219-1224 [FREE Full text] [doi: 10.2337/dc11-2163] [Medline: 22474040]
45. Patel NJ, Datye KA, Jaser SS. Importance of patient-provider communication to adherence in adolescents with type 1 diabetes. Healthcare (Basel) 2018 Mar 28;6(2):30 [FREE Full text] [doi: 10.3390/healthcare6020030] [Medline: 29597246]
46. Scholten MR, Kelders SM, Van Gemert-Pijnen JE. Self-guided web-based interventions: scoping review on user needs and the potential of embodied conversational agents to address them. J Med Internet Res 2017 Nov 16;19(11):e383 [FREE Full text] [doi: 10.2196/jmir.7351] [Medline: 29146567]
47. Kang H. The prevention and handling of the missing data. Korean J Anesthesiol 2013 May;64(5):402-406 [FREE Full text] [doi: 10.4097/kjae.2013.64.5.402] [Medline: 23741561]
48. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med 2012 Oct 4;367(14):1355-1360 [FREE Full text] [doi: 10.1056/NEJMsr1203730] [Medline: 23034025]
49. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]

Abbreviations

CGMS: continuous glucose monitoring system
DCCT: Diabetes Control and Complications Trial
FT: Family Teamwork
FTO: Diabetes Family Teamwork Online
HbA$_{1c}$: hemoglobin A$_{1c}$
Ragt: relational agent
SDT: self-determination theory
T1D: type 1 diabetes
USDA/ARS: US Department of Agriculture/Agricultural Research Service

Edited by G Eysenbach; submitted 01.07.19; peer-reviewed by S Brigitte, T Oser; comments to author 24.07.19; revised version received 30.07.19; accepted 13.08.19; published 13.09.19.

Please cite as:
Thompson D, Callender C, Gonynor C, Cullen KW, Redondo MJ, Butler A, Anderson BJ
Using Relational Agents to Promote Family Communication Around Type 1 Diabetes Self-Management in the Diabetes Family Teamwork Online Intervention: Longitudinal Pilot Study
J Med Internet Res 2019;21(9):e15318
URL: https://www.jmir.org/2019/9/e15318/
doi:10.2196/15318
PMID:31538940

Original Paper

An Internet-Based Self-Help Intervention for Skin Picking (SaveMySkin): Pilot Randomized Controlled Trial

Christina Gallinat[1], MSc; Markus Moessner[1], PhD; Holger A Haenssle[2], MD; Julia K Winkler[2], MD; Matthias Backenstrass[3,4], PhD; Stephanie Bauer[1], PhD

[1]Center for Psychotherapy Research, University Hospital Heidelberg, Heidelberg, Germany

[2]Department of Dermatology, University Hospital Heidelberg, Heidelberg, Germany

[3]Institute of Clinical Psychology, Hospital Stuttgart, Stuttgart, Germany

[4]Institute of Psychology, Department of Clinical Psychology and Psychotherapy, Heidelberg University, Heidelberg, Germany

Corresponding Author:
Christina Gallinat, MSc
Center for Psychotherapy Research
University Hospital Heidelberg
Bergheimer Str 54
Heidelberg
Germany
Phone: 49 6221567381
Email: christina.gallinat@med.uni-heidelberg.de

Abstract

Background: In spite of the psychosocial burden and medical risks associated with skin picking disorder, the health care system does not provide sufficient treatment for affected individuals to date. Therefore, an internet-based self-help program for skin picking was developed to offer easily accessible support for this population.

Objective: This pilot study evaluated the internet-based self-help program SaveMySkin. The 12-week program is based on cognitive-behavioral therapy and contains comprehensive information and exercises, a daily supportive monitoring system, and dermatological and psychological counseling via internet chat. Primary objectives were the investigation of attitudes and expectations toward the program, intervention effects on skin picking severity, user satisfaction, adherence, and willingness to participate. Secondary outcomes included the feasibility of study procedures, adequacy of assessment instruments, effects on skin picking–related impairment, dimensions of skin picking, and general psychological impairment.

Methods: A two-arm randomized controlled trial was conducted in a sample of 133 participants (female: 124/133, 93.2%; mean age 26.67 [SD 6.42]) recruited via the internet. Inclusion required a minimum age of 17 years and at least mild skin picking severity. Participants were randomly allocated to the intervention (64/133, 48.1%) or waitlist control group (69/133, 51.9%). All assessments were conducted online and based on self-report.

Results: The willingness to participate was very high in the study, so the initially planned sample size of 100 was exceeded after only 18 days. Participant expectations indicate that they believed the program to be beneficial for them (131/133, 98.5%) and provide a feeling of support (119/133, 89.5%). Reasons for study participation were insufficient outpatient health care (83/133, 62.4%) and flexibility regarding time (106/133, 79.7%) and location (109/133, 82.0%). The post-assessment was completed by 65.4% (87/133) of the sample. The majority of the intervention group who completed the entire post-assessment were satisfied with SaveMySkin (28/38, 74%) and agreed that the program is an appropriate support service (35/38, 92%). On average, participants viewed 29.31 (SD 42.02) pages in the program, and 47% (30/64) of the intervention group used the monitoring at least once a week. In comparison with the control group, the intervention group displayed substantial improvements in the skin picking severity total score (Cohen d=0.67) and especially on the subscale Symptom Severity (Cohen d=0.79). No effects on secondary outcomes were found.

Conclusions: This study confirms the need for easily accessible interventions for skin picking disorder and the high interest in internet-based self-help within the target population. It provides important insights into the attitudes toward online support and actual user experiences. Participant feedback will be used to further enhance the intervention. Our results point to the preliminary efficacy of SaveMySkin and may lay the foundation for future research into the efficacy and cost-effectiveness of the program in a multicenter clinical trial.

Trial Registration: German Clinical Trial Register DRKS00015236; https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00015236

International Registered Report Identifier (IRRID): RR2-10.1016/j.conctc.2018.100315

(J Med Internet Res 2019;21(9):e15011) doi:10.2196/15011

KEYWORDS

skin picking; excoriation disorder; dermatillomania; internet-based; self-help; cognitive-behavioral therapy

Introduction

Background

Skin picking is a body-focused repetitive behavior receiving increased attention since “excoriation (skin picking) disorder” was recognized as a distinct category within the Obsessive-Compulsive and Related Disorders of the *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition* (DSM-5). The core symptom of the disorder is a recurrent behavioral pattern of manipulating one's own skin (eg, scratching, squeezing, excoriating, picking), which causes skin damage including wounds, erosive skin lesions, and scars in the long term. Individuals with skin picking are not able to resist the urge or stop the behavior [1]. Skin picking disorder predominantly affects females [2-4], and its lifetime prevalence is estimated at 1.2% to 1.4% [5,6].

Skin picking symptomatology is associated with impairment on psychological, social, and physical levels: affected individuals suffer not only from wounds, infections, and scarring [7,8] but also from embarrassment, guilt, and symptoms of depression and anxiety [2,3,9,10]. Occupational interferences, academic impairment, and a financial burden due to skin picking–related costs (eg, concealing cosmetics and clothing, fees for treatment) have also been reported [9,10]. In light of the substantial psychosocial impairment and risk for chronicity [2,11], the need for affected individuals to receive timely professional treatment is evident.

However, research on interventions for skin picking disorder has been very scarce until now. A limited number of studies have investigated pharmacological and behavioral interventions [12], including habit reversal training [13,14], acceptance and commitment therapy [15], cognitive-behavior therapy (CBT) [16,17], and combined approaches (eg, acceptance-enhanced CBT) [18,19]. Noteworthy, most of these previous studies showed severe methodological shortcomings (eg, small sample sizes, lack of control conditions), and most were conducted before the official DSM-5 criteria for skin picking disorder became available. Thus, the current evidence base for treatment of skin picking disorder is rather weak. Only one study investigated an internet- and CBT-based self-help intervention for skin picking; the study reported substantial improvements in symptom severity for 63% of the sample [17]. However, the study was uncontrolled, and only 4% of the initial sample (15/372) completed the entire intervention, so the results should be interpreted with caution. So far, two meta-analyses suggest an overall beneficial effect of behavioral treatments on skin picking severity, but these studies must also be seen as preliminary due to the small number of included original studies and their limited validity [20,21]. Overall, CBT seems to be the most promising approach for the treatment of skin picking disorder so far. This is also plausible in light of the fact that behavioral interventions have demonstrated efficacy and are currently considered as the method of choice in the treatment of trichotillomania [22-24], which shows substantial overlap in clinical characteristics and a high co-occurrence with skin picking disorder [25].

The scarcity of research on skin picking disorder reflects an overall lack of awareness and knowledge of this disorder. Affected individuals barely receive adequate treatment and face manifold difficulties in finding appropriate help [9]. Internet-based interventions have the potential to improve the health care situation for skin picking disorder due to their reach, accessibility, and availability. The efficacy and cost effectiveness of internet-based treatment approaches have already been proven for other psychiatric conditions including depression, anxiety, and eating disorders [26-29]. Therefore, we considered an internet-based self-help program a promising opportunity to provide support to individuals affected by skin picking. The program is conceptualized as a stand-alone intervention in order to complement conventional health care for skin picking disorder.

Objectives

We developed an internet-based self-help program for skin picking, and conducted a pilot randomized controlled trial (RCT) to investigate the adequacy of the intervention SaveMySkin and the feasibility of the study procedures. Primary objectives of our study were the investigation of attitudes and expectations toward SaveMySkin before randomization, intervention effects on skin picking severity, and user satisfaction. Further outcomes were program adherence (intervention use) and willingness to participate.

The feasibility of study procedures (eg, recruitment, randomization), appropriateness of applied questionnaires, effects on skin picking–related impairment, dimensions of skin picking (focused vs automatic skin picking), and general psychological impairment were investigated as secondary outcomes.

Methods

Study Design

This pilot study followed a two-arm randomized controlled design with a 1:1 allocation to either intervention or waitlist control group. The design is fully described in the study protocol [30]. No essential changes have been made to the study protocol after study commencement. The actual sample size exceeded the initial aim due to a very fast recruitment via internet, so the original plan to expand recruitment to dermatological clinics was not pursued.

Participant Selection

Inclusion required at least mild self-reported skin picking severity (Skin Picking Scale–Revised [SPS-R] score ≥7 [31,32]) and a minimum age of 17 years. Sufficient German language skills, home access to the internet, a smartphone, and literacy on internet and computer use were applied as implicit eligibility criteria. Potential participants were recruited via online advertisement (eg, specific forums, support groups, university mailing lists) and at a conference for skin picking and trichotillomania. In case of interest, individuals could directly access an openly available online screening questionnaire checking for eligibility. Eligible individuals were invited to register for the study and give the required informed consent for participation. Study participation did not include any restrictions concerning additional treatment use. The use of conventional treatment was assessed as part of the final questionnaire after 12 weeks.

Study Arms

Intervention

Participants randomized to the intervention group received immediate access to the internet-based intervention, SaveMySkin, for 12 weeks. The program is based on CBT techniques and consists of several modules:

- Psychoeducation: information about skin picking disorder, treatment options, and dermatological topics
- Self-management: a module with three submodules (Skills: information materials and online exercises; Tools: downloadable offline trainings; and Emotions: online exercises on emotion regulation) aiming at the reduction of skin picking behavior and the enhancement of self-management skills based on classic CBT methods like self-observation, cognitive restructuring, and behavioral strategies
- Supportive monitoring: daily support via email including a motivational message in the morning and a short monitoring questionnaire in the evening, combined with an automatically generated, tailored feedback message
- Counseling via internet chat: optional personal support in individual chat sessions with psychologists or psychological and dermatological group chats

Overall, the intervention follows a flexible and demand-oriented design. Participants were therefore expected to use the program depending on their individual needs. Recommendations on the use of certain program modules or exercises were given within chat sessions or in the monitoring feedback. Additional information on all modules of SaveMySkin is provided in Gallinat et al [30].

Control Condition

Participants in the control group did not have access to the intervention until the final assessment after 12 weeks. In the final questionnaire, participants in the control group were asked if they would still like to use the intervention. If this was the case, intervention access was activated, and participants could use the program for 12 weeks.

Primary Outcomes

All assessments—scheduled at t0 (screening), t1 (baseline), t2 (after 6 weeks), and t3 (after 12 weeks)—were performed as self-report online questionnaires. Screening data (t0) and data derived from the t1 assessment (right after screening and registration) are both referred to as baseline data. A detailed plan listing all assessments and instruments can be found in Gallinat et al [30].

Attitudes and Expectations

Attitudes and expectations toward SaveMySkin were investigated with 10 statements rated on a 4-point Likert scale from “does not apply” to “totally applies.” In addition, participants could report further reasons for participation.

Skin Picking Severity

Skin picking severity was assessed with the German version of the SPS-R [31,32]. The first 4 items of the scale refer to the subscale Symptom Severity and assess the frequency and intensity of the urge to pick the skin, time spent on skin picking, and control over skin picking behavior. The other 4 items form the subscale Impairment and assess impairing consequences caused by skin picking (avoidance, interference in social and occupational life, emotional distress, skin damage). All items are rated on a 5-point Likert scale from 0 to 4 in reference to the last 7 days. In our study, a very good internal consistency with a coefficient of α=.81 was observed for the total scale (Symptom Severity: α=.72; Impairment: α=.83).

User Satisfaction

User satisfaction was measured with self-designed items assessing overall satisfaction criteria (eg, recommendation to others, fulfillment of expectations). Satisfaction with single modules was assessed with statements rated on a 4-point Likert scale from “does not apply” to “totally applies” (eg, “I like the idea that individual chat sessions with psychologists are offered”). Participants could also indicate “not able to evaluate.”

Adherence and Use

Adherence and program use were automatically documented within the program. Monitoring compliance was assessed by the number of completed monitoring questionnaires. Chat use was evaluated by the number of booked individual chat appointments and log-ins into group chats. The use of other modules and of the overall program was investigated by the number of page views per module and user as well as log-ins per user.

Secondary Outcomes

Skin Picking–Related Impairment

Skin picking–related impairment was assessed with a German translation of the Skin Picking Impact Scale (SPIS) [33,34]. The 10 items are rated on a 5-point Likert scale from “not at all” (0) to “severe” (4) and refer to the past 7 days. The SPIS demonstrated an excellent internal consistency in our study (α=.94).

Focused Versus Automatic Skin Picking

Modes of skin picking relating to the awareness of performing the behavior were assessed with a German translation of the Milwaukee Inventory for the Dimensions of Adult Skin Picking (MIDAS) [35]. The 12 items are rated on a 5-point Likert scale from “not true for any of my skin picking” (1) to “true for all of my skin picking” (5) and form the two subscales Focused Skin Picking and Automatic Skin Picking with 6 items each. Our study revealed acceptable internal consistencies for both subscales (Focused: α=.73; Automatic: α=.69).

General Psychological Impairment

General psychological impairment was assessed with the Clinical Psychological Diagnosis System 38 (KPD-38) [36,37]. The scale consists of 38 items assessing psychological impairment, social problems, general physical condition, general life satisfaction, competence skills, and social support. The items are rated on a 4-point Likert scale from “does not apply” (1) to “totally applies” (4). In our sample, internal consistency of the KPD-38 total score was excellent, with a Cronbach alpha coefficient of α=.94.

Sample Size

The sample size in this pilot trial was determined on the basis of practical considerations and questions of feasibility, as Leon et al [38] recommend for pilot studies. These considerations led to a targeted sample size of 100 participants, which was assumed to be sufficient for the investigation of the study objectives.

Randomization

Participants were equally (1:1) randomized to one of the study groups by the software based on an a priori defined list (intervention vs waitlist control) after they completed the registration and baseline questionnaire. Randomization was stratified by gender and followed a permuted block design. The allocation sequence was produced with a computerized random number generator.

Statistical Analysis

Descriptive statistics were used to analyze data on attitudes, expectations, user satisfaction, and program use. Continuous variables were dichotomized by splitting the Likert scale (eg, “agree” contains “applies mostly” and “totally applies”; “helpful” contains “a little helpful” and “very helpful”) to analyze data with frequencies. Efficacy was tested by mixed models. Intervention effects were analyzed as cross-level interactions (group × time). The control group was coded 0, and the intervention group was coded 1. Assessment points were coded as follows: baseline (t0, t1)=0; t2=1; t3=2. In accordance with the recommendations of Lorah [39], another run of mixed-models analyses was conducted with standardized outcome variables to calculate the Cohen *d* based on the estimated coefficient per time span (Cohen *d* = standardized coefficients of the time × group interaction * max[time]). It should be noted that one participant in each study group did not complete the entire post-assessment. The analyses regarding user satisfaction and help-seeking at t3 therefore refer to n=38 (intervention group) and n=47 (control group). Statistical analyzes were conducted with SPSS Statistics version 25.0 (IBM Corp).

Ethical Considerations

This trial was approved by the ethics committee of the Medical Faculty of Heidelberg University and registered at the German Clinical Trials Register [DRKS00015236]. The study protocol was published before recruitment was completed [30].

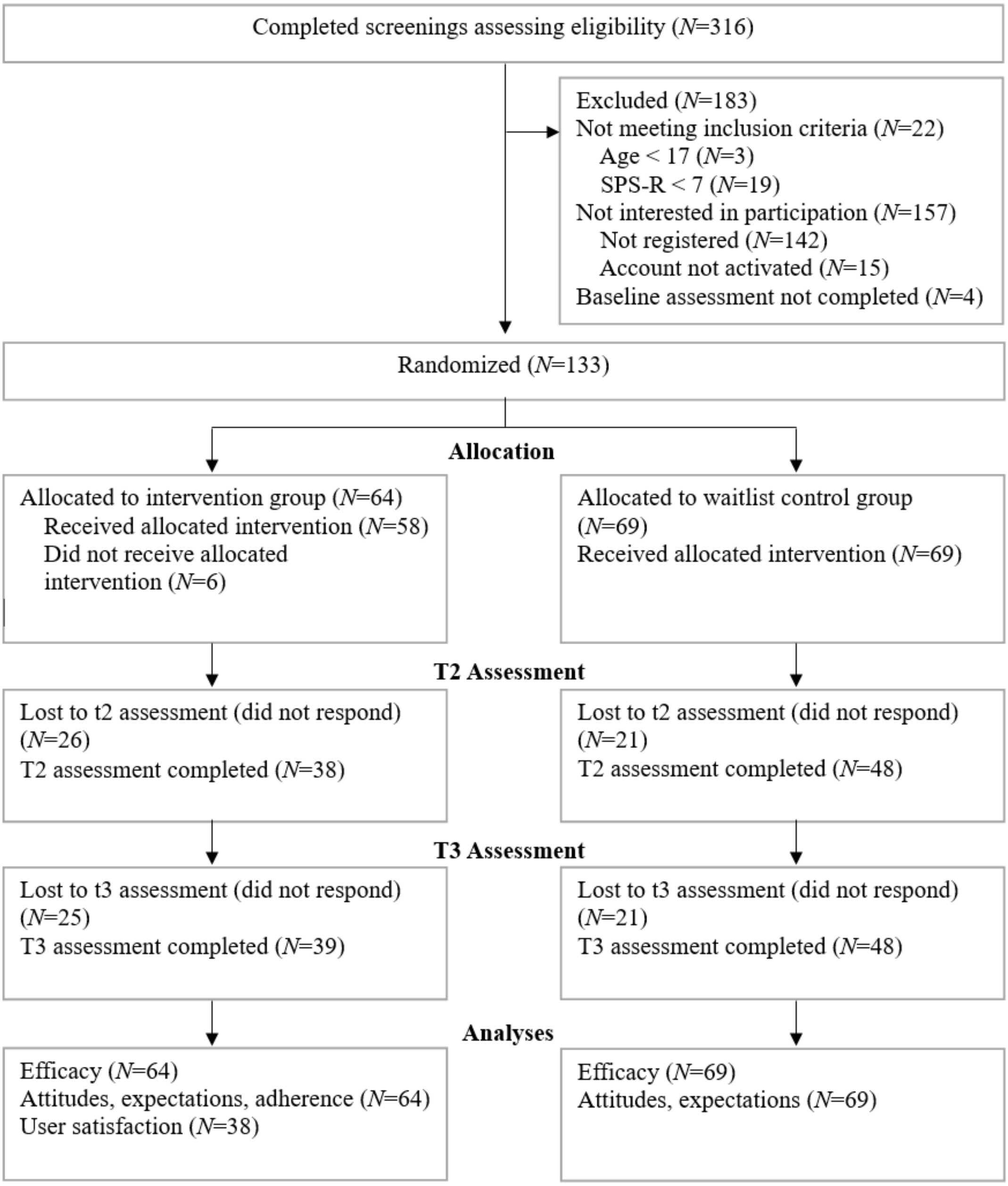
Results

Recruitment and Participant Flow

Participant enrollment started in October 2018. The planned sample size of 100 participants was achieved after only 18 days of recruitment. Advertisement was stopped then, but due to ethical considerations, screening and registration were not closed until December 2018.

Out of 316 individuals who completed the screening questionnaire, 294 were eligible for study participation. More than half of this subsample registered for study participation (152/294, 51.7%), but 15 individuals did not activate their account and 4 did not complete the baseline assessment. In the end, 42.1% (133/316) of the individuals screened for eligibility completed the entire inclusion process and were therefore randomized to one of the study groups; 64 participants were allocated to the intervention group and 69 participants to the waitlist control group. Six participants in the intervention group neither logged into the program after the initial registration process nor completed any of the daily monitoring assessments, so they did not receive the allocated intervention. The response rate for the assessment at t2 (6 weeks after randomization) was 59% in the intervention group (38/64) and 70% in the control group (48/69). The final assessment (t3: 12 weeks after randomization) was completed by 61% of the intervention group (39/64) and 70% of the control group (48/69). A detailed overview of the participant flow including the number of analyzed cases for each objective is provided in Figure 1.

Figure 1. Participant flow diagram. SPS-R: Skin Picking Scale–Revised.



Baseline Characteristics

Tables 1 and 2 show the demographic and clinical characteristics for each study group. The majority of the sample was female (124/133, 93.2%) with a mean age of 26.67 (SD 6.42) years (range 17-56).

The total sample displays high levels of skin picking severity (SPS-R scores) and skin picking–related impairment (SPIS scores). Furthermore, the participants show significant psychological impairment with regard to the KPD-38 total score (mean 2.44 [SD 0.55]), which corresponds to the 87th percentile of the norm data for women in the general population (age category: 14-34 years) [36].

Nearly one-quarter of the sample (32/133, 24.1%) was currently in psychotherapeutic or psychiatric treatment at the time of baseline assessment. Except for the SPS-R subscale Symptom Severity, the study groups did not differ in the assessed baseline characteristics.

Table 1. Demographic characteristics (N=133).

Characteristic	Intervention group (n=64)	Control group (n=69)	Test statistic	*P* value
Female, n (%)	61 (95)	63 (91)	0.85[a]	.36
Age (years), mean (SD)	26.19 (6.52)	27.12 (6.34)	0.83[b]	.41
Education, n (%)			3.07[c]	.55
Still in school	1 (2)	1 (1)		
Middle secondary	7 (11)	7 (10)		
Highest secondary	29 (45)	23 (33)		
University	27 (42)	37 (54)		
Other	—[d]	1 (1)		
Occupational status, n (%)			4.54[e]	.60
Employed	15 (23)	27 (39)		
Trainee	5 (8)	3 (4)		
School student	1 (2)	1 (1)		
University student	38 (59)	35 (51)		
Housewife/househusband	2 (3)	1 (1)		
Retired	1 (2)	1 (1)		
Unemployed	2 (3)	1 (1)		
Family status, n (%)			3.07[f]	.38
Single	30 (47)	26 (38)		
In a relationship	26 (41)	29 (42)		
Married	8 (13)	12 (17)		
Other	—	2 (3)		
Current psychotherapeutic or psychiatric treatment at baseline, n (%)	15 (23)	17 (25)	0.03[a]	.87

[a] χ^2_1.

[b] t_{131}.

[c] χ^2_4.

[d] Not applicable.

[e] χ^2_6.

[f] χ^2_3.

Table 2. Clinical variables at baseline (N=133).

Characteristic	Intervention group (n=64), mean (SD)	Control group (n=69), mean (SD)	Test statistic	*P* value
SPS-R [a] total score	16.62 (4.33)	15.68 (4.04)	–1.30[b]	.20
SPS-R Symptom Severity	9.94 (1.97)	9.20 (2.26)	–1.99[b]	.048
SPS-R Impairment	6.69 (3.08)	6.48 (2.58)	–0.43[b]	.67
KPD-38[c] total score	2.47 (0.52)	2.42 (0.57)	–0.43[b]	.67
SPIS[d]	17.73 (10.59)	18.30 (10.00)	0.32[b]	.75
MIDAS[e] focused skin picking	18.92 (5.15)	18.48 (4.54)	–0.53[b]	.60
MIDAS automatic skin picking	17.91 (4.94)	17.99 (3.88)	0.10[f]	.92

[a] SPS-R: Skin Picking Scale–Revised.

[b] t_{131}.

[c] KPD-38: Clinical Psychological Diagnosis System–38.

[d] SPIS: Skin Picking Impact Scale.

[e] MIDAS: Milwaukee Inventory for the Dimensions of Adult Skin Picking.

[f] $t_{119.42}$.

Attitudes and Expectations

Prior to randomization, almost all participants expected the program to be generally helpful and beneficial for them. They expected to feel supported and gain a positive effect on their well-being. More than two-thirds (83/133, 62.4%) indicated that they would like to participate due to insufficient support options within the regular health care system (Table 3).

Common reasons for participation were the flexibility of the internet-based setting regarding time (106/133, 79.7%) and location (109/133, 82.0%), expertise for skin picking (98/133, 73.7%), free counseling (97/133, 72.9%), anonymity (83/133, 62.4%), the option to contact somebody in a quick and easy way (75/133, 56.4%), and the possibility to ask questions in written form (49/133, 36.8%).

Efficacy

In comparison to the control group, the intervention group yielded considerably higher reductions in skin picking severity and symptom severity resulting in moderate effect sizes of *d*=0.67 and *d*=0.79 (Cohen *d*, Table 4). The intervention and control group both showed improved scores on the impairment scale, but the analyses did not reveal a significant difference between the groups.

Table 3. Attitudes and expectations toward SaveMySkin (N=133). Answers on the 4-point Likert scale were dichotomized (disagree: “does not apply” and “applies somewhat”; agree: “applies mostly” and “totally applies”).

Statement	Agreement, n (%)
I believe that my participation in SaveMySkin will have a positive effect on my well-being.	118 (88.7)
I believe that I will feel supported by SaveMySkin.	119 (89.5)
The effort for the participation in SaveMySkin seems low to me.	86 (64.7)
My motivation to participate in SaveMySkin is high.	121 (91.0)
In general, I have a positive attitude toward communication technologies (eg, computer, mobile phone, internet).	125 (94.0)
Without SaveMySkin, I am sufficiently supplied with health care offers.	32 (24.1)
I want to participate in SaveMySkin because I believe participation would be helpful for me.	130 (97.7)
I believe that I would benefit from participation in SaveMySkin.	131 (98.5)
I would like to participate in SaveMySkin because health care services are insufficient.	83 (62.4)
Other reasons for your participation.	32 (24.1)

Table 4. Efficacy (linear mixed-effects models) of the intervention.

Variable	Estimate	95% CI	*t* test (*df*)	*P* value	Cohen *d*
SPS-R[a] total score					
Time	–1.17	–2.02, –0.33	–2.73 (204.50)	.007	—[b]
Group	0.77	–0.58, 2.13	1.13 (161.78)	.26	—
Time × group	–1.59	–2.83, –0.34	–2.50 (203.88)	.01	0.67
SPS-R Symptom Severity					
Time	–0.73	–1.17, –0.29	–3.24 (198.83)	.001	—
Group	0.60	–0.09, 1.30	1.71 (155.82)	.09	—
Time × group	–1.01	–1.67, –0.36	–3.06 (197.60)	.003	0.79
SPS-R Impairment					
Time	–0.44	–0.97, 0.09	–1.64 (206.48)	.10	—
Group	0.19	–0.71, 1.09	0.42 (172.50)	.67	—
Time × group	–0.57	–1.35, 0.21	–1.44 (206.31)	.15	0.4
KPD-38[c] total score					
Time	0.02	–0.08, 0.11	0.31 (202.04)	.76	—
Group	0.02	–0.16, 0.20	0.22 (166.86)	.83	—
Time × group	–0.05	–0.20, 0.09	–0.72 (200.85)	.47	0.2
Skin Picking Impact Scale					
Time	–0.16	–1.99, 1.68	–0.17 (204.78)	.87	—
Group	–0.87	–4.14, 2.40	–0.52 (171.56)	.60	—
Time × group	–1.82	–4.52, 0.87	–1.34 (203.91)	.18	0.37
MIDAS [d] focused skin picking					
Time	0.20	–0.68, 1.07	0.44 (205.10)	.66	—
Group	0.15	–1.38, 1.69	0.20 (174.80)	.84	—
Time × group	–0.60	–1.89, 0.68	–0.93 (204.56)	.35	0.26
MIDAS automatic skin picking					
Time	0.06	–0.71, 0.84	0.17 (203.64)	.87	—
Group	–0.32	–1.75, 1.10	–0.45 (164.29)	.65	—
Time × group	–0.41	–1.55, 0.73	–0.71 (202.19)	.48	0.19

[a] SPS-R: Skin Picking Scale–Revised.

[b] Not applicable.

[c] KPD-38: Clinical Psychological Diagnosis System–38.

[d] MIDAS: Milwaukee Inventory for the Dimensions of Adult Skin Picking.

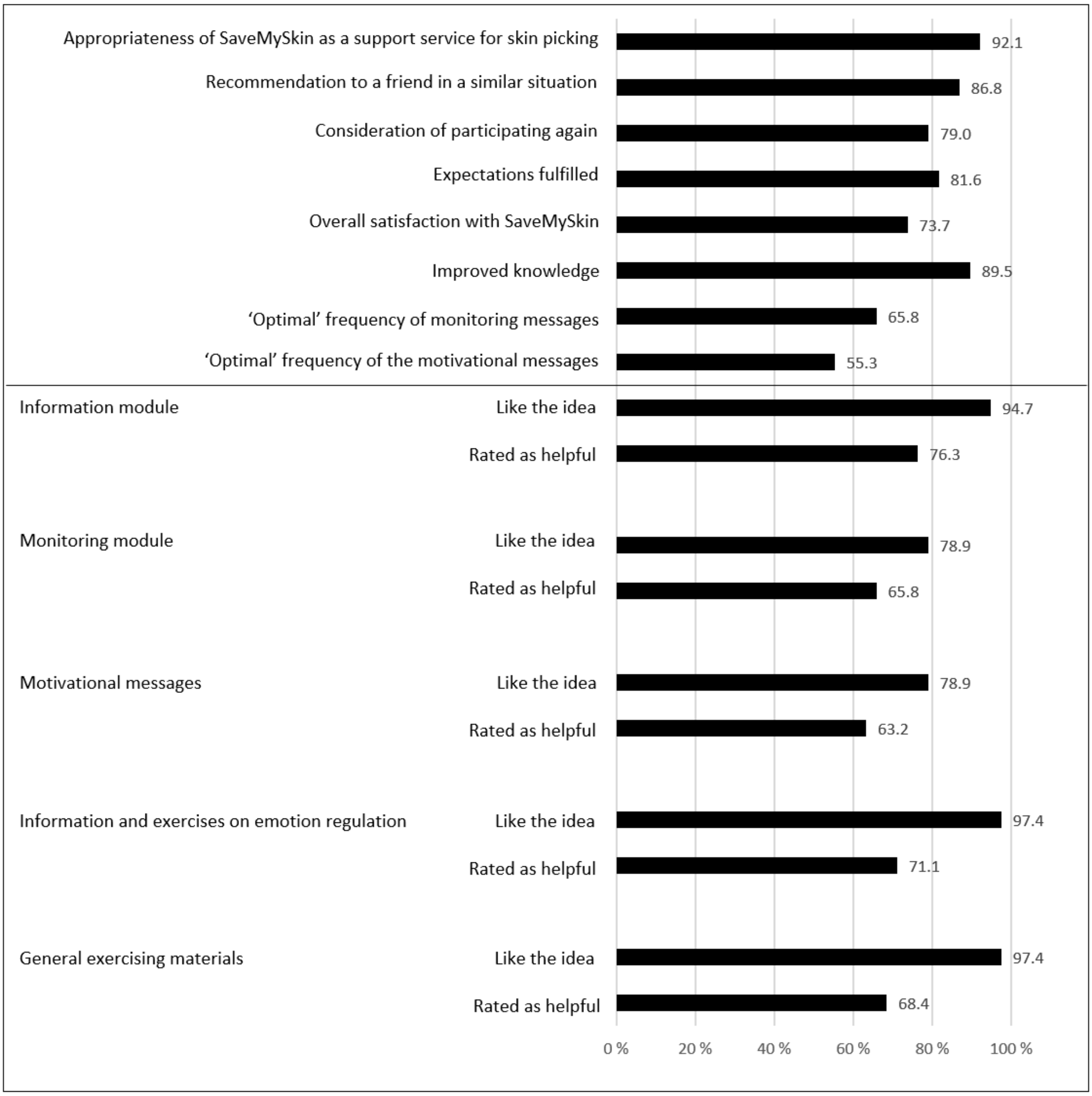
User Satisfaction

All results regarding user satisfaction refer to the subsample of the intervention group that completed the entire post-assessment questionnaire (38/64).

Overall Satisfaction

Almost all participants agreed that the program is an appropriate and supportive offering for individuals with skin picking and that they would recommend it to a friend in a similar situation. Most participants also agreed that the program met their expectations and improved their knowledge on skin picking (Figure 2). The length of the intervention was rated as optimal by 39% (15/38) and as too short by 42% (16/38). Only 18% (7/38) thought it was too long.

Figure 2. User satisfaction and evaluation of modules.



Evaluation of Modules

More than 94% of the t3 completers in the intervention group (36/38, 95%) liked the idea that information materials, information and exercises about emotion regulation, and additional exercises on several topics were part of the program. The idea of a daily motivational message and monitoring, and inclusion of all three chat types (psychological individual chat, psychological group chat, dermatological group chat) were also positively evaluated by more than three-quarters of the sample (29-34/38, 76% to 89%). Most participants evaluated the different SaveMySkin modules as helpful. The helpfulness of the chat module was not evaluated by most of the participants (82% to 87% for the different chats), since only a very small percentage used the chats. It should be noted that evaluation of the single modules contains up to n=4 values in the category "not able to evaluate" (4/38, 11%). Details on the evaluation of program modules are presented in Figure 2.

Adherence and Program Use

Monitoring

On average, participants completed 25.81 (SD 26.96; range 0-81) out of 84 daily monitoring assessments. Three-quarters of the sample (48/64, 75%) completed more than 3 monitorings, 50% (32/64) answered more than 9, and 25% (16/64) completed more than 45 monitoring questionnaires. In sum, 47% (30/64) participated in the monitoring at least once a week and 42% (27/64) participated at least twice a week.

Chat

Five individual chat sessions were booked and carried out (4/64 participants, 6%). The psychological group chat was used by 3 participants (5%) and the dermatological group chat by 6 individuals (9%), who participated up to 3 times in one of the chat types. Participants were asked why they had not used the chat (33/64). The most common answers were “I don’t really know why I didn’t use the chat” (20/33, 61%), “I had no need because I could seek advice somewhere else (eg, psychotherapy)” (13/33, 39%), and “I had no need because I felt good” (12/33, 36%).

Other response options were “I didn’t really know what to expect” (10/33, 30%), “I couldn’t imagine that an internet chat would be helpful in this topic” (8/33, 24%), “It was too much effort for me to book an appointment” (7/33, 21%), “I didn’t know about this option” (5/33, 15%), “I was scared that I am technically not fit enough (eg, cannot type fast enough)” (3/33, 9%), and “Other reasons” (11/33, 33%; most often time conflicts or no time [6/11]).

Log-Ins and Views

On average, participants logged in 6.42 (SD 10.15) times (median 3; range 1-67). The average number of page views per user across the following 4 basic modules was 29.31 (SD 42.04; Information: mean 8.17 [SD 11.92]; Skills: mean 11.52 [SD 20.84]; Tools: mean 5.59 [SD 8.09], Emotions: mean 4.03 [SD 6.16]). These modules contain 35 pages in total. Answering the daily monitoring questionnaire was not counted as a log-in.

Effects on Secondary Outcomes

No statistically significant time × group interactions occurred for secondary outcome variables (Table 4).

Help-Seeking Behavior During Participation

At the time of post-assessment, 17% (8/48) of the control group and 18% (7/38) of the intervention group who completed the t3 questionnaire indicated that they had used professional help due to skin picking in the last 12 weeks. In the intervention group, 24% (9/38) indicated that they were planning to use professional help; in the control group, 15% (7/48) were planning to seek professional help.

Discussion

Principal Findings

Skin picking disorder is associated with psychological distress, impairment in social life, and medical risks, but currently individuals with skin picking rarely receive the required professional support due to an insufficient health care supply. Our study investigated attitudes and expectations toward an internet-based self-help program for skin picking as well as user satisfaction and effects on skin picking severity and impairment.

Full recruitment of the initially targeted sample size for this pilot trial was rapidly achieved, indicating a high willingness to participate in an internet-based intervention for skin picking. Randomization resulted in two comparable study groups that only differed marginally in the SPS-R subscale Symptom Severity.

Participants were highly motivated and expected increased well-being and a feeling of support provided by the program. Such positive expectations are known to be an important factor contributing to intervention effects [40]. Flexibility in terms of time and location, expertise related to skin picking, and a lack of other health care options were further reasons for participating. Almost all t3 completers in the control group requested access to the intervention (96%) after the waiting period of 12 weeks, indicating a persistent motivation to use the intervention.

The majority of those in the intervention group who completed the post-assessment reported a high satisfaction with the modules included in SaveMySkin and the program in general (eg, appropriateness, length, recommendation to others). On average, participants completed more than 2 monitorings per week, suggesting that daily monitorings might be too frequent.

The intervention group yielded substantial reductions in skin picking severity (SPS-R total score) and specifically in the subscale Symptom Severity compared with the control group. The size of these effects (d=0.67 and d=0.79) is comparable to the overall effect of behavioral treatments for skin picking disorder reported in a meta-analysis (standardized mean difference 0.68) [20]. The analyses did not confirm meaningful differences between groups regarding improvements in skin picking–related impairment measured via the SPS-R and SPIS. Given the rather short time period covered in the trial, this result is not surprising since skin picking–related impairment (eg, impaired self-esteem, avoidance, skin damage) may only improve slowly, even if skin picking frequency and intensity are improved. Furthermore, some medical consequences, especially scars, often need to be considered as permanent. The short study period may also be responsible for the lack of effects on general psychological impairment and different dimensions of skin picking (focused vs automatic) in our study. Dimensions of skin picking were assessed with regard to habitual but not necessarily current patterns (eg, “I am usually not aware of picking my skin during the picking episode”), so potential changes might not be reflected properly. Also, sensitivity to change has yet to be explored for this assessment instrument (MIDAS). Apart from the MIDAS, the applied instruments proved appropriate for interventional studies. As the primary outcome measure, the SPS-R proved to be sensitive to change. This is of special importance for subsequent studies, given the lack of interventional studies on skin picking disorder and the associated uncertainty about the adequate measurement of intervention effects.

Concerning use of the self-help program, it turned out that the chat module was used only rarely, even though most participants (more than 76%) liked the idea that different chat modules were included in the program. Given that chatting is not an obligatory key element of the program but an optional offer for those who feel the need for personal counseling, the low chat use is not concerning. Rather, it is in accordance with previous research suggesting that a considerable number of users in online communities do not actively produce content (eg, posting in a forum) but rather read and browse through the platform [41]. More than half of the participants who did not use the chat indicated that they were not sure why. Potential underlying reasons could, for example, be insecurity and shyness when talking with others about this very personal topic, even if it is online, or a reluctance to commit to chat participation at a certain date and time. Even though various reasons led to a rather infrequent use of the module, it seems to be important to keep the chat as part of the program because most participants liked the idea that the different chats were included. The module may not be used by many participants, but it could nevertheless be very important for some individuals, especially for severely impaired individuals without access to other support.

Limitations and Implications for Future Research

This study included several limitations, which should be considered in the interpretation of results. First, all assessments were conducted online, and inclusion relied exclusively on self-reported data, so internal validity and generalizability to a larger clinical population might be compromised. Another shortcoming results from the rather low response rate at post-assessment. The lack of data from approximately one-third of the sample limits the validity of our findings, since it remains unclear how satisfied the nonresponders were and how they changed in symptomatology. Additionally, our study did not investigate the stability of effects, as it focused on the classic aims of a pilot study and therefore did not include extended follow-up assessments. However, the results of this study clearly demonstrated the feasibility of an internet-based intervention in the target group. Furthermore, the study provided preliminary evidence for the efficacy of the intervention. Subsequent research should therefore investigate the efficacy and cost effectiveness of SaveMySkin in a fully powered RCT. This trial should apply a sequential enrollment procedure including a clinician-rated assessment of psychiatric conditions and a dermatological assessment and documentation of the skin status. Moreover, the study period should be expanded to assess the stability of intervention effects and explore potential long-term effects on impairment.

The program use in our study and user feedback suggest minor adaptions of the program, which will be implemented to ensure that users' needs are met even better by the provided support. The dermatological group chat will be replaced by a forum where participants can ask dermatologists for advice. This asynchronous form of communications seems to fit user behavior better, since the forum can be accessed any time and does not require meeting a certain appointment. Additionally, the self-management module will be reorganized into topical sessions that will guide participants through the process in a more structured manner. This adjustment was recommended by several users. It may also improve program use, since users receive more guidance and make visible progress as they move from one session to the next. Moreover, visibility of progress will be improved by including a new module providing visual feedback on daily monitoring data (eg, graph displaying changes in symptom severity) and completed exercises within sessions. Additionally, following user recommendations, the daily monitoring messages will no longer be delivered via email but via mobile messaging technology (eg, short message service) in order to facilitate more immediate participation. These adjustments are expected to improve user satisfaction and the extent of program use and may ultimately further enhance the efficacy of the intervention. In a subsequent efficacy study, the impact of certain program modules should be analyzed in more detail by exploring associations between program use (especially within the monitoring module) and changes in skin picking symptomatology.

The pilot study derived important knowledge on the feasibility and adequacy of an internet-based intervention for skin picking disorder and therefore provides essential information for subsequent research. We are aware that in general the intervention effects observed in a pilot trial have only limited value for guiding the preparation of a subsequent RCT (eg, in terms of power calculations) [42]. Therefore, the pilot results should be considered as highly promising but preliminary evidence that should be interpreted with adequate caution.

Conclusions

To our knowledge, this is the first study to comprehensively investigate an internet-based self-help intervention for skin picking disorder in an RCT. This pilot study showed that SaveMySkin seems to be a welcomed and suitable program for individuals with skin picking that can induce substantial improvements in symptomatology. The results of the pilot trial will be used to design a subsequent study on the efficacy and cost effectiveness of SaveMySkin, which may be a beneficial complement to conventional health care for skin picking disorder.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT‐EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File)2577 KB - jmir_v21i9e15011_app1.pdf]

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th Edition. Washington: American Psychiatric Publishing; 2013.
2. Wilhelm S, Keuthen NJ, Deckersbach T, Engelhard IM, Forker AE, Baer L, et al. Self-injurious skin picking: clinical characteristics and comorbidity. J Clin Psychiatry 1999 Jul;60(7):454-459. [Medline: 10453800]
3. Hayes SL, Storch EA, Berlanga L. Skin picking behaviors: an examination of the prevalence and severity in a community sample. J Anxiety Disord 2009 Apr;23(3):314-319 [FREE Full text] [doi: 10.1016/j.janxdis.2009.01.008] [Medline: 19223150]

4. Odlaug BL, Lust K, Schreiber LRN, Christenson G, Derbyshire K, Grant JE. Skin picking disorder in university students: health correlates and gender differences. Gen Hosp Psychiatry 2013;35(2):168-173 [FREE Full text] [doi: 10.1016/j.genhosppsych.2012.08.006] [Medline: 23123103]
5. Monzani B, Rijsdijk F, Cherkas L, Harris J, Keuthen N, Mataix-Cols D. Prevalence and heritability of skin picking in an adult community sample: a twin study. Am J Med Genet B Neuropsychiatr Genet 2012 Jul;159B(5):605-610 [FREE Full text] [doi: 10.1002/ajmg.b.32067] [Medline: 22619132]
6. Keuthen NJ, Koran LM, Aboujaoude E, Large MD, Serpe RT. The prevalence of pathologic skin picking in US adults. Compr Psychiatry 2010;51(2):183-186 [FREE Full text] [doi: 10.1016/j.comppsych.2009.04.003] [Medline: 20152300]
7. Neziroglu F, Rabinowitz D, Breytman A, Jacofsky M. Skin picking phenomenology and severity comparison. Prim Care Companion J Clin Psychiatry 2008;10(4):306-312 [FREE Full text] [doi: 10.4088/pcc.v10n0406] [Medline: 18787665]
8. Odlaug BL, Grant JE. Clinical characteristics and medical complications of pathologic skin picking. Gen Hosp Psychiatry 2008;30(1):61-66. [doi: 10.1016/j.genhosppsych.2007.07.009] [Medline: 18164942]
9. Tucker BTP, Woods DW, Flessner CA, Franklin SA, Franklin ME. The Skin Picking Impact Project: phenomenology, interference, and treatment utilization of pathological skin picking in a population-based sample. J Anxiety Disord 2011 Jan;25(1):88-95. [doi: 10.1016/j.janxdis.2010.08.007] [Medline: 20810239]
10. Flessner CA, Woods DW. Phenomenological characteristics, social problems, and the economic impact associated with chronic skin picking. Behav Modif 2006 Nov;30(6):944-963. [doi: 10.1177/0145445506294083] [Medline: 17050772]
11. Grant JE, Odlaug BL, Chamberlain SR, Keuthen NJ, Lochner C, Stein DJ. Skin picking disorder. Am J Psychiatry 2012 Nov;169(11):1143-1149. [doi: 10.1176/appi.ajp.2012.12040508] [Medline: 23128921]
12. Lochner C, Roos A, Stein DJ. Excoriation (skin-picking) disorder: a systematic review of treatment options. Neuropsychiatr Dis Treat 2017;13:1867-1872 [FREE Full text] [doi: 10.2147/NDT.S121138] [Medline: 28761349]
13. Moritz S, Fricke S, Treszl A, Wittekind CE. Do it yourself! Evaluation of self-help habit reversal training versus decoupling in pathological skin picking: a pilot study. J Obsessive Compuls Relat Disord 2012 Jan;1(1):41-47. [doi: 10.1016/j.jocrd.2011.11.001]
14. Teng EJ, Woods DW, Twohig MP. Habit reversal as a treatment for chronic skin picking: a pilot investigation. Behav Modif 2006 Jul;30(4):411-422. [doi: 10.1177/0145445504265707] [Medline: 16723422]
15. Twohig MP, Hayes SC, Masuda A. A preliminary investigation of acceptance and commitment therapy as a treatment for chronic skin picking. Behav Res Ther 2006 Oct;44(10):1513-1522. [doi: 10.1016/j.brat.2005.10.002] [Medline: 16368072]
16. Schuck K, Keijsers GPJ, Rinck M. The effects of brief cognitive-behaviour therapy for pathological skin picking: a randomized comparison to wait-list control. Behav Res Ther 2011 Jan;49(1):11-17 [FREE Full text] [doi: 10.1016/j.brat.2010.09.005] [Medline: 20934685]
17. Flessner CA, Mouton-Odum S, Stocker AJ, Keuthen NJ. StopPicking.com: Internet-based treatment for self-injurious skin picking. Dermatol Online J 2007 Oct 13;13(4):3. [Medline: 18319000]
18. Capriotti MR, Ely LJ, Snorrason I, Woods DW. Acceptance-enhanced behavior therapy for excoriation (skin-picking) disorder in adults: a clinical case series. Cogn Behav Pract 2015 May;22(2):230-239. [doi: 10.1016/j.cbpra.2014.01.008]
19. Flessner CA, Busch AM, Heideman PW, Woods DW. Acceptance-enhanced behavior therapy (AEBT) for trichotillomania and chronic skin picking: exploring the effects of component sequencing. Behav Modif 2008 Sep;32(5):579-594. [doi: 10.1177/0145445507313800] [Medline: 18334614]
20. Schumer MC, Bartley CA, Bloch MH. Systematic review of pharmacological and behavioral treatments for skin picking disorder. J Clin Psychopharmacol 2016 Apr;36(2):147-152 [FREE Full text] [doi: 10.1097/JCP.0000000000000462] [Medline: 26872117]
21. Selles RR, McGuire JF, Small BJ, Storch EA. A systematic review and meta-analysis of psychiatric treatments for excoriation (skin-picking) disorder. Gen Hosp Psychiatry 2016;41:29-37 [FREE Full text] [doi: 10.1016/j.genhosppsych.2016.04.001] [Medline: 27143352]
22. Bloch MH, Landeros-Weisenberger A, Dombrowski P, Kelmendi B, Wegner R, Nudel J, et al. Systematic review: pharmacological and behavioral treatment for trichotillomania. Biol Psychiatry 2007 Oct 15;62(8):839-846. [doi: 10.1016/j.biopsych.2007.05.019] [Medline: 17727824]
23. McGuire JF, Ung D, Selles RR, Rahman O, Lewin AB, Murphy TK, et al. Treating trichotillomania: a meta-analysis of treatment effects and moderators for behavior therapy and serotonin reuptake inhibitors. J Psychiatr Res 2014 Nov;58:76-83 [FREE Full text] [doi: 10.1016/j.jpsychires.2014.07.015] [Medline: 25108618]
24. Flessner C, Penzel F, Keuthen N. Current treatment practices for children and adults with trichotillomania: consensus among experts. Cogn Behav Pract 2010 Aug;17(3):290-300. [doi: 10.1016/j.cbpra.2009.10.006]
25. Snorrason I, Belleau EL, Woods DW. How related are hair pulling disorder (trichotillomania) and skin picking disorder? A review of evidence for comorbidity, similarities and shared etiology. Clin Psychol Rev 2012 Nov;32(7):618-629. [doi: 10.1016/j.cpr.2012.05.008] [Medline: 22917741]
26. Ebert DD, Zarski A, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: a meta-analysis of randomized controlled outcome trials. PLoS One 2015;10(3):e0119895 [FREE Full text] [doi: 10.1371/journal.pone.0119895] [Medline: 25786025]

27. Königbauer J, Letsch J, Doebler P, Ebert D, Baumeister H. Internet- and mobile-based depression interventions for people with diagnosed depression: a systematic review and meta-analysis. J Affect Disord 2017 Dec 01;223:28-40. [doi: 10.1016/j.jad.2017.07.021] [Medline: 28715726]
28. Watson HJ, Joyce T, French E, Willan V, Kane RT, Tanner-Smith EE, et al. Prevention of eating disorders: a systematic review of randomized, controlled trials. Int J Eat Disord 2016 Sep;49(9):833-862. [doi: 10.1002/eat.22577] [Medline: 27425572]
29. Moessner M, Bauer S, Özer F, Wolf M, Zimmer B, Kordy H. Cost-effectiveness of an internet-based aftercare intervention after inpatient treatment in a psychosomatic hospital. Psychother Res 2014;24(4):496-503 [FREE Full text] [doi: 10.1080/10503307.2013.845919] [Medline: 24188127]
30. Gallinat C, Moessner M, Haenssle HA, Winkler JK, Backenstrass M, Bauer S. SaveMySkin: an Internet-based self-help intervention for skin picking. Study protocol for a randomized pilot study. Contemp Clin Trials Commun 2019 Mar;13:100315 [FREE Full text] [doi: 10.1016/j.conctc.2018.100315] [Medline: 31011654]
31. Gallinat C, Keuthen NJ, Backenstrass M. Ein Selbstbeurteilungsinstrument zur Erfassung von Dermatillomanie: Reliabilität und Validität der deutschsprachigen Version der Skin Picking Scale-Revised [A self-report instrument for the assessment of dermatillomania: reliability and validity of the German Skin Picking Scale–Revised]. Psychother Psychosom Med Psychol 2016 Jun;66(6):249-255. [doi: 10.1055/s-0042-107255] [Medline: 27286530]
32. Snorrason I, Ólafsson RP, Flessner CA, Keuthen NJ, Franklin ME, Woods DW. The Skin Picking Scale-Revised: factor structure and psychometric properties. J Obsessive Compuls Relat Disord 2012 Apr;1(2):133-137. [doi: 10.1016/j.jocrd.2012.03.001]
33. Keuthen NJ, Deckersbach T, Wilhelm S, Engelhard I, Forker A, O'Sullivan RL, et al. The Skin Picking Impact Scale (SPIS): scale development and psychometric analyses. Psychosomatics 2001;42(5):397-403. [doi: 10.1176/appi.psy.42.5.397] [Medline: 11739906]
34. Mehrmann LM, Hunger A, Gerlach AL. Pathologisches Hautzupfen/-quetschen (Skin Picking). Erste Ergebnisse zur Psychometrie störungsspezifischer Messinstrumente [Excoriation disorder (skin picking): an initial analysis of disorder-specific assessment instruments]. Zeitschrift für Klinische Psychologie und Psychotherapie 2017 Jan;46(1):23-31. [doi: 10.1026/1616-3443/a000386]
35. Walther MR, Flessner CA, Conelea CA, Woods DW. The Milwaukee Inventory for the Dimensions of Adult Skin Picking (MIDAS): initial development and psychometric properties. J Behav Ther Exp Psychiatry 2009 Mar;40(1):127-135. [doi: 10.1016/j.jbtep.2008.07.002] [Medline: 18725154]
36. Percevic R, Gallas C, Wolf M, Haug S, Hünerfauth T, Schwarz M, et al. Das Klinisch Psychologische Diagnosesystem 38 (KPD-38). Entwicklung, Normierung und Validierung eines Selbstbeurteilungsbogens für den Einsatz in Qualitätssicherung und Ergebnismonitoring in der Psychotherapie und psychosomatischen Medizin [The Klinisch Psychologische Diagnosesystem 38 (KPD-38): development, standardization, and validation of a self-report questionnaire for use in quality assurance and outcome monitoring in psychotherapy and psychosomatic medicine]. Diagnostica 2005 Jul;51(3):134-144. [doi: 10.1026/0012-1924.51.3.134]
37. Moessner M, Gallas C, Haug S, Kordy H. The clinical psychological diagnostic system (KPD-38): sensitivity to change and validity of a self-report instrument for outcome monitoring and quality assurance. Clin Psychol Psychother 2011;18(4):331-338 [FREE Full text] [doi: 10.1002/cpp.717] [Medline: 20812367]
38. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
39. Lorah J. Effect size measures for multilevel models: definition, interpretation, and TIMSS example. Large-scale Assess Educ 2018 Jul 23;6(1):1 [FREE Full text] [doi: 10.1186/s40536-018-0061-2]
40. Greenberg RP, Constantino MJ, Bruce N. Are patient expectations still relevant for psychotherapy process and outcome? Clin Psychol Rev 2006 Oct;26(6):657-678. [doi: 10.1016/j.cpr.2005.03.002] [Medline: 15908088]
41. Nonnecke B, Andrews D, Preece J. Non-public and public online community participation: needs, attitudes and behavior. Electron Commerce Res 2006 Jan;6(1):7-20. [doi: 10.1007/s10660-006-5985-x]
42. Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch Gen Psychiatry 2006 May;63(5):484-489. [doi: 10.1001/archpsyc.63.5.484] [Medline: 16651505]

Abbreviations

CBT: cognitive-behavioral therapy
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
KPD-38: Clinical Psychological Diagnosis System–38
MIDAS: Milwaukee Inventory for the Dimensions of Adult Skin Picking
RCT: randomized controlled trial
SPIS: Skin Picking Impact Scale
SPS-R: Skin Picking Scale–Revised

Edited by G Eysenbach; submitted 12.06.19; peer-reviewed by M Yap, KP Wong; comments to author 05.07.19; revised version received 19.07.19; accepted 19.07.19; published 20.09.19.

Please cite as:
Gallinat C, Moessner M, Haenssle HA, Winkler JK, Backenstrass M, Bauer S
An Internet-Based Self-Help Intervention for Skin Picking (SaveMySkin): Pilot Randomized Controlled Trial
J Med Internet Res 2019;21(9):e15011
URL: https://www.jmir.org/2019/9/e15011
doi:10.2196/15011
PMID:31586368

Original Paper

A Web-Based Photo-Alteration Intervention to Promote Sleep: Randomized Controlled Trial

Isabel Perucho[1*], BA; Kamalakannan M Vijayakumar[1*], BA; Sean N Talamas[2], PhD; Michael Wei-Liang Chee[3], MBBS; David I Perrett[4], PhD; Jean C J Liu[1,3], PhD

[1]Division of Social Sciences, Yale-NUS College, Singapore, Singapore
[2]School of Psychology, University of Pennsylvania, Philadelphia, PA, United States
[3]Neuroscience and Behavioral Disorders Programme, Duke-NUS Medical School, Singapore, Singapore
[4]School of Psychology and Neuroscience, University of St Andrews, St Andrews, United Kingdom
[*]these authors contributed equally

Corresponding Author:
Jean C J Liu, PhD
Division of Social Sciences
Yale-NUS College
16 College Ave West #02-221
Singapore, 138527
Singapore
Phone: 65 66013694
Email: jeanliu@yale-nus.edu.sg

Abstract

Background: Receiving insufficient sleep has wide-ranging consequences for health and well-being. Although educational programs have been developed to promote sleep, these have had limited success in extending sleep duration. To address this gap, we developed a Web-based program emphasizing how physical appearances change with varying amounts of sleep.

Objective: The aims of this study were to evaluate (1) whether participants can detect changes in appearances as a function of sleep and (2) whether this intervention can alter habitual sleep patterns.

Methods: We conducted a 5-week, parallel-group, randomized controlled trial among 70 habitual short sleepers (healthy adults who reported having <7 hours of sleep routinely). Upon study enrollment, participants were randomly assigned (1:1) to receive either standard information or an appearance-based intervention. Both groups received educational materials about sleep, but those in the appearance group also viewed a website containing digitally edited photographs that showed how they would look with varying amounts of sleep. As the outcome variables, sleep duration was monitored objectively via actigraphy (at baseline and at postintervention weeks 1 and 4), and participants completed a measure of sleep hygiene (at baseline and at postintervention weeks 2, 4, and 5). For each outcome, we ran intention-to-treat analyses using linear mixed-effects models.

Results: In total, 35 participants were assigned to each group. Validating the intervention, participants in the appearance group (1) were able to identify what they looked like at baseline and (2) judged that they would look more attractive with a longer sleep duration (t_{26}=10.35, P<.001). In turn, this translated to changes in sleep hygiene. Whereas participants in the appearance group showed improvements following the intervention ($F_{1,107.99}$=9.05, P=.003), those in the information group did not ($F_{1,84.7}$=0.19, P=.66). Finally, there was no significant effect of group nor interaction of group and time on actigraphy-measured sleep duration (smallest P=.26).

Conclusions: Our findings suggest that an appearance-based intervention, while not sufficient as a stand-alone, could have an adjunctive role in sleep promotion.

Trial Registration: ClinicalTrials.gov NCT02491138; https://clinicaltrials.gov/ct2/show/study/NCT02491138.

(J Med Internet Res 2019;21(9):e12500) doi:10.2196/12500

KEYWORDS

sleep; public health; physical appearance; outward appearance

Introduction

Background

Sleep is often described as 1 of 3 pillars of health, ranked alongside nutrition and exercise as modifiable targets of well-being. Underscoring this point, a large body of evidence suggests that the habitual curtailment of sleep increases the risk of obesity [1-3], coronary heart disease [4], stroke [4], and all-cause mortality [5,6]. Despite the centrality of sleep, 1 in 3 adults routinely obtain less than the recommended 7 hours of sleep for healthy adults [7,8]. Accordingly, the US Department of Health and Human Services has outlined a nation-wide goal to reduce the number of habitual short sleepers over a 10-year period [9].

Despite this goal, sleep promotion campaigns have met with limited success. In a typical program, participants are briefed about the mechanics of sleep (eg, sleep architecture), the importance of sleep, and sleep hygiene—lifestyle habits that facilitate sleep (eg, abstaining from excessive caffeine intake during the night [10]) [11-13]. Delivery of these programs range from simply providing a pamphlet [14] to hosting multiple sessions (eg, 8 sessions over 5 weeks [15]). Although sleep promotion programs may raise awareness, they fare poorly in changing actual behavior [11,12]. Even among programs that have reported increased sleep duration, gains have been short-lived, observed only immediately after the intervention [13].

Reviewing this evidence, Cassoff et al suggested that merely providing information may not be sufficient and that greater emphasis should be placed on motivational aspects of modifying sleep [12]. This concurs with surveys suggesting that—despite knowing the importance of sleep—many forsake it for immediate priorities such as meeting deadlines [16,17]. That is, they engage in *temporal discounting*, weighting short-term rewards (eg, completing one’s assignment) over long-term outcomes (eg, reduced risk of all-cause mortality) [18,19]. This aligns with research in other health domains (eg, appetite regulation), where interventions emphasizing short-term gains (eg, financial incentives for weight loss or exercise) were found to encourage healthy behaviors [20,21]. Correspondingly, one way to motivate better sleep patterns may be to highlight short-term outcomes [22,23]—the proximate incentives of increased sleep duration or the proximate costs of sleep curtailment. To this end, we evaluated a sleep promotion program emphasizing 1 immediate outcome—how physical appearances change as a function of sleep duration.

Developing an Appearance-Based Intervention for Sleep Promotion

Physical appearance is highly valued across cultures [24], and the desire to be attractive has motivated healthy behaviors in both men and women [25]. For example, in the prevention of skin cancers, showing participants how they would look with continued sun exposure has been found repeatedly to increase sun-protection behaviors and reduce indoor tanning frequency [25]. In smoking cessation campaigns, showing the photoaging effects of tobacco has likewise increased readiness to change and actual quit attempts [26,27]. Finally, a recent study found that when participants were shown how their skin color would vary as a function of nutrition, their intake of fruit and vegetable consumption increased, with effects lasting for 10 weeks postintervention [28].

In the context of sleep, physical appearances track sleep duration within a short time window (eg, after 31 hours of total sleep deprivation [29,30] or following 2 nights with 4 hours of sleep opportunity [31]). Relative to adequate sleep (≥7 hours), sleep loss results in hanging eyelids, redder and swollen eyes, darker eye circles, wrinkles and fine lines around the eyes, a droopy mouth, and poorer skin quality (eg, worsened skin hydration and elasticity and paler skin) [30,32]. In turn, these alterations are associated with looking sadder [30], less intelligent [33], less attractive [29], less healthy [29], and less desirable as a social companion [31]. Given the immediate and salient nature of these changes, physical appearances could serve as a motivator to overcome temporal discounting. Accordingly, we developed a novel Web-based intervention to capitalize on these changes.

To summarize, we sought in this study to highlight the immediate impact of sleep loss on physical appearances. Drawing from other health domains (sun protection, smoking cessation, and diet), we hypothesized that the appearance-based intervention would augment a standard informational campaign in increasing sleep duration. Furthermore, because poor sleep hygiene predisposes a person to impaired nocturnal sleep and is targeted in sleep education programs [10,11], we tracked changes in sleep hygiene habits as an additional indicator of behavioral change [34].

Methods

Study Design

This study involved a randomized controlled open-label trial with 2 arms: a standard information intervention and an appearance-based intervention. The trial design and outcomes were preregistered in ClinicalTrials.gov (NCT02491138) and remained unchanged throughout the study. All procedures were approved by the National University of Singapore’s Institutional Review Board (A-15-083).

Participants

A total of 70 young adults were recruited from the National University of Singapore between July 2015 and December 2016. Participants responded to advertisements within the university and registered their interest via an online website. Thereafter, they were included if they were (1) aged between 18 and 24 years; (2) had no history of psychiatric, sleep, neurological, or medical disorders (including insomnia); (3) had no history of substance abuse; and (4) reported habitual short sleep (as defined by current sleep guidelines: sleep duration of <7 hours [8]).

Following written informed consent (see Multimedia Appendix 1), participants were randomized to the 2 groups with a 1:1 allocation. The allocation sequence was prepared by the study coordinator before trial commencement and involved a computer-generated list of random numbers. Both participants and research assistants became aware of the allocation on the intervention day; however, research staff involved in data entry and data cleaning were blinded to participant grouping. Participants were not told what the intervention of interest was.

Procedures

As a baseline measure, participants completed 1 week of sleep monitoring where they kept a sleep diary [35] and wore a wrist actigraph on the nondominant hand (Actiwatch; Philips Respironics, Inc, Pittsburgh, PA). Activity was recorded in epochs of 2 min, and actigraphy measures were calculated using Actiware 6.0.5 (Philips Respironics, Inc, Pittsburgh, PA) after photosensor data and sleep diaries were reviewed.

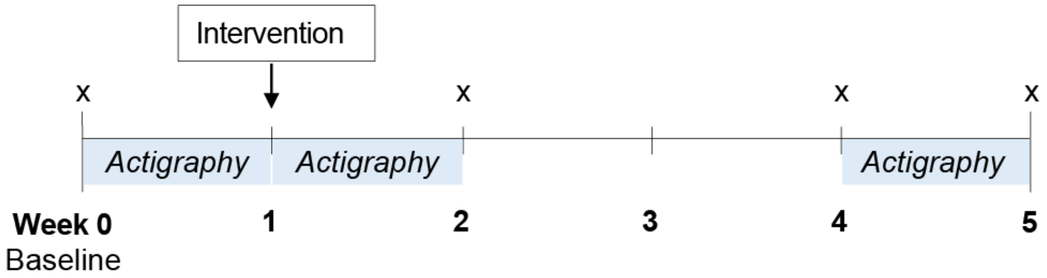
During the baseline phase, participants also attended the laboratory for a face-to-face visit where they completed the Sleep Hygiene scale [36]. This scale involved 19 items assessing the number of days each week participants engaged in poor sleep hygiene practices (eg, *slept in a room that was too bright*; *worried, planned, or thought about important matters in bed*). The 19 items were then averaged to form a composite score, with higher scores indicating poorer sleep hygiene (Cronbach α=.72). As this component is targeted in typical sleep education programs, any improvements from baseline would represent a preliminary step in behavioral change.

As additional baseline measures, participants were also characterized with 2 questionnaires assessing sleep quality and attitudes: (1) the Dysfunctional Beliefs and Attitudes about Sleep scale (DBAS-16) [37], a 16-item measure of insomnia-related cognition (Cronbach α=.87); and (2) the Pittsburgh Sleep Quality Index (PSQI) [38], a 19-item measure of sleep quality over the past week (Cronbach α=.52). Finally, facial photographs were obtained from all participants (regardless of their intervention group).

For the photoshoot, 3 facial photographs were taken in a standardized room under constant lighting (resolution: 3648 × 5472 pixels, Canon Powershot G7X; Canon Inc, Tokyo, Japan). Before taking photographs, participants removed any spectacles and jewelry, combed their hair backwards, were clean shaven, and applied a facial wipe. They were then instructed to look into the camera with a neutral expression, with photographs taken at a fixed distance from the camera.

After the baseline phase (week 0), participants visited the laboratory individually where they received either the information or appearance-based intervention. Following this visit, participants repeated the sleep hygiene scale (weeks 2, 4, and 5) and completed 2 weeks of wrist actigraphy (weeks 1 and 4). Finally, they repeated the DBAS-16 and PSQI questionnaires (in weeks 2, 4, and 5). Upon study completion, participants were reimbursed SGD $50. All study visits took place at the behavioral laboratories of Yale-NUS College in Singapore, with procedures and measures administered in English—the lingua franca of the country (see Figure 1 for a schematic of study procedures).

Figure 1. Schematic of study procedures. All participants received the intervention in week 1 and had their sleep patterns monitored through actigraphy in weeks 0 (baseline), 1, and 4. The sleep hygiene questionnaire was also administered in weeks 0 (baseline), 2, 4, and 5 (marked in the figure with an “x”).



Interventions

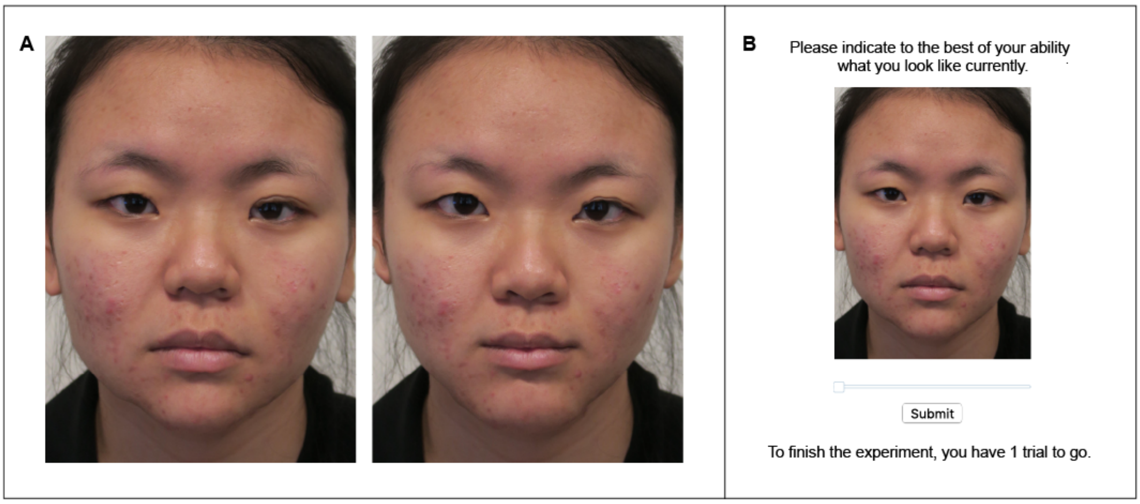
Standard Information

In the standard information group, participants were shown a slideshow containing materials based on public health websites [39,40] and previous sleep education studies [11-13]. A trained research assistant walked participants through the slide contents, introducing them to: the functions of sleep, the consequences of sleep curtailment, behavioral signs of insufficient sleep, and sleep hygiene principles to promote sleep (eg, noise management in the sleeping environment). Participants were then given a take-home pamphlet containing this same information and were not given the opportunity to ask questions. Each session was conducted individually and spanned approximately 20 min.

Appearance-Based Intervention

In the appearance group, transformed versions of baseline photographs were presented through an internet platform to show how appearances would change with sleep duration. Sleep-related changes were ascertained from an earlier study where 25 healthy adults were photographed following rested wakefulness (7-9 hours of sleep) and following 2 nights of sleep restriction (where they were given 4 hours of sleep opportunity) [31]. Using Psychomorph 6 [41] — a free software for photo transformation—these faces were blended to create 2 masks for each state (rested vs sleep restricted). Transformations were then applied by first delineating participants’ photographs with 175 feature points aligned based on interpupillary distance. Using the masks, participants’ photographs were transformed to create a continuum of 13 images, with Image 0 showing what participants would look like if they had insufficient sleep (300% of the difference between the rested and sleep restricted masks; Figure 2). This progressed incrementally with Image 6 as the original image, and Image 12 showing what participants would look like if they received more sleep (300% of the difference in the opposite direction; Figure 2).

Figure 2. In the appearance-based intervention, participants were shown what they would look like as a function of varying sleep duration. (A) Each participant's face was digitally edited to show them their sleep restricted (left image) and well-rested selves (right image). (B) A continuum of 13 images was created and presented to participants through an online website.



Please indicate to the best of your ability what you look like currently.

To finish the experiment, you have 1 trial to go.

During the intervention visit, participants were shown the same slideshow presentation as the information group. However, after hearing about the signs of insufficient sleep, they were shown a website where the 13 transformed images were presented on a continuum, with a sliding scale to manipulate appearance changes as a function of sleep duration (Figure 2). Participants scrolled through the website in a self-paced manner and were asked to (1) identify what they looked like at present and (2) to adjust the scale to select their most attractive self. Upon making their choices, participants were again presented with Images 0 (their sleep-restricted selves) and 12 (their well-rested selves). Referring to these images, the trained research assistant highlighted how sleep curtailment affected their appearances (eg, hanging eyelids and droopy mouth), and how that could be perceived by others (eg, as less attractive). Thereafter, participants were reminded of the costs many incur for physical attractiveness (eg, paying a gym membership for $100/month), relative to the benefits accrued from extending sleep duration. Finally, participants were presented information about sleep hygiene practices and were given a take-home pamphlet (similar to the information group). Each session was conducted individually and spanned approximately 30 min, with neither an invitation for participants to ask questions nor further access to the website (following the intervention).

Statistical Analysis

For the primary outcome measures (actigraphy-measured sleep duration and sleep hygiene scores), we ran intention-to-treat analyses using linear mixed-effects models, with parameters estimated using maximum likelihood estimation for a first-order autoregressive covariance structure. This examined individual change in each measure over time, with group (information or appearance), time (sleep duration: weeks 0, 2, 4, and 5; sleep hygiene: 0, 1, and 4), and the group × time interaction entered as fixed effects. Random intercepts accounted for correlated data due to repeated measures. Although the reliability of the PSQI was low in our sample (Cronbach α=.52), baseline scores differed between the groups and were entered as covariates [42]; however, the conclusions did not change with and without this adjustment. Finally, for sleep duration, separate models were run for weekdays and for weekends to account for differences in sleep patterns across the week [43].

For the secondary outcome measures (global PSQI and DBAS-16 scores), we ran linear mixed-effects models identical to those used for the primary outcome measures, except that the model with global PSQI scores did not include covariates.

All analyses were conducted using SPSS 25 (IBM Corp, Armonk, NY) and R 3.4.0 (R Core Team, Vienna, Austria), with the type 1 decision-wise error rate controlled at α=.05. Power calculations for the main intention-to-treat analyses showed that there was statistical power at the recommended .80 level to detect a medium effect size (computed through simulations, based on a meta-analysis evaluating appearance-based interventions for sun protection behaviors [25]). Accordingly, data collection was scheduled to cease when 70 participants had been recruited.

Results

Participant Characterstics

Full details on participant flow are shown in Figure 3. At baseline, actigraphy monitoring showed that participants slept an average of 5.9 hours on weekdays (SD 0.90 hours) and 6.4 hours on weekends (SD 1.32 hours). There were no significant group differences in baseline: gender, ethnicity, age, DBAS, sleep hygiene scores, and actigraphy-measured sleep variables

(Tables 1 and 2). However, baseline global PSQI differed between groups and was included as a covariate.

Figure 3. Consolidated Standards of Reporting Trials flow diagram.

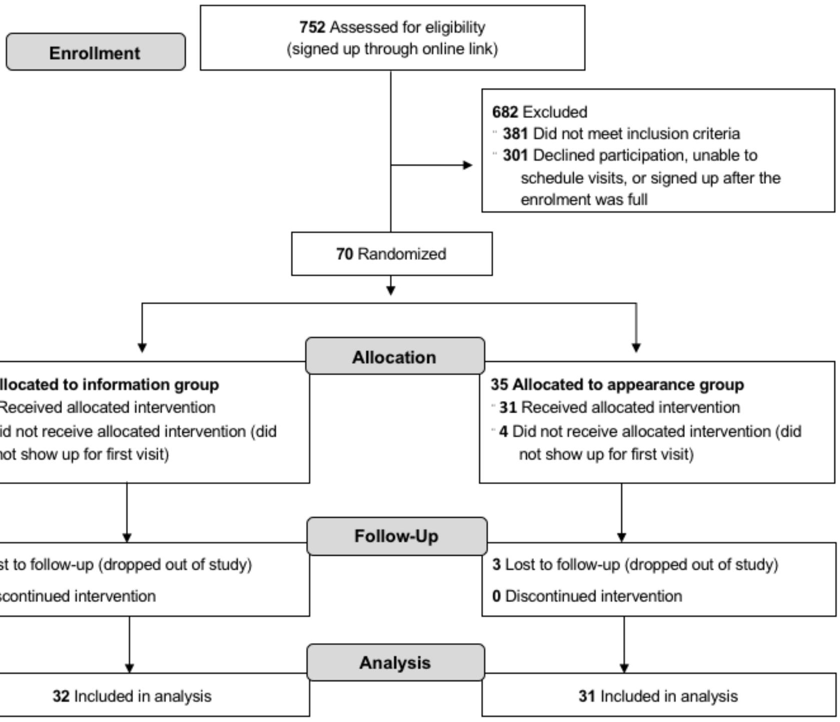


Table 1. Baseline characteristics of participants in the information and appearance groups.

Characteristics	Information (n=32), n (%)	Appearance (n=31), n (%)	Chi-square (*df*)	*P* value
Gender				
Female	25 (78)	18 (58)	2.92 (1)	.09
Ethnicity			7.83 (8)	.45
Chinese	23 (72)	22 (71)	—[a]	—
Indian	4 (12)	7 (22)	—	—
Others	5 (16)	2 (6)	—	—

[a] Not applicable.

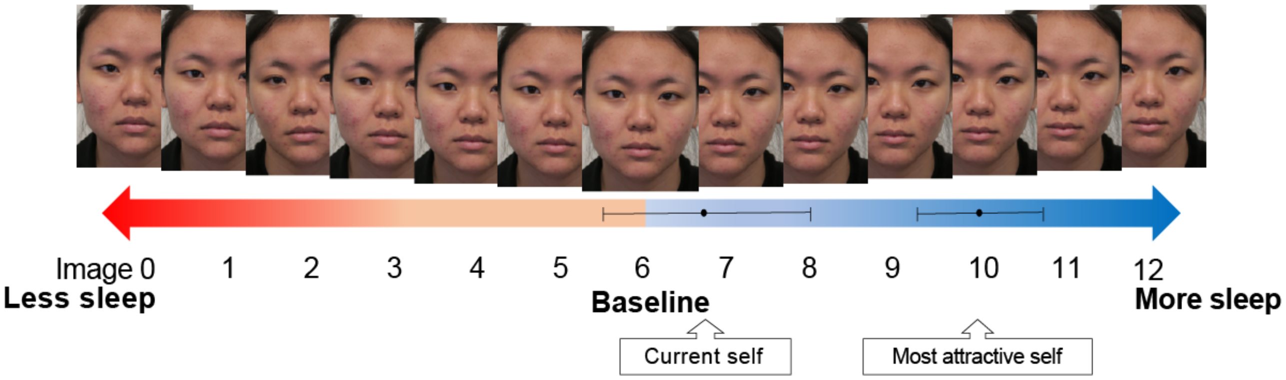
Table 2. Baseline characteristics of participants by sleep score and actigraphy metrics.

Characteristics	Information (n=32), mean (SD)	Appearance (n=31), mean (SD)	*t* test	*P* value
Age (years)	20.66 (1.91)	22.13 (2.14)	−0.93 (61)	.36
Mean sleep questionnaire scores				
Pittsburgh Sleep Quality Index (global score)	10.72 (2.70)	9.57 (1.65)	1.99 (57)	.05
Dysfunctional Beliefs and Attitudes about Sleep	4.42 (1.32)	4.67 (1.44)	−0.71 (61)	.48
Sleep Hygiene Score	2.61 (0.74)	2.52 (0.65)	0.51 (61)	.61
Mean actigraphy metrics (averaged across the week)				
Bed time (hours:min)	2:18 (1:02)	1:57 (1:14)	1.19 (56)	.24
Wake time (hours:min)	8:57 (1:05)	8:47 (1:12)	0.57 (56)	.57
Weekday sleep duration (hours:min)	5:55 (0:52)	5:54 (0:56)	0.04 (56)	.97
Weekend sleep duration (hours:min)	6:07 (1:18)	6:42 (1:18)	−1.71 (55)	.09
Sleep latency (mins)	12.77 (8.01)	14.00 (8.81)	0.55 (56)	.58
Wake after sleep onset (mins)	30.78 (17.24)	29.99 (17.36)	0.17 (56)	.86
Sleep efficiency (%)	85.16 (5.14)	86.12 (5.51)	−0.69 (56)	.49

Validating the Appearance-Based Intervention

When participants in the appearance group were asked to select what they looked like, they identified Image 6.81 on average (SD 3.32). A one-sample *t* test found no evidence that this differed from the actual baseline image (Image 6; t_{26}=1.28, *P*=.21, 95% CI for the mean difference: −0.50 to 2.13); in other words, participants correctly identified their current appearances. On the other hand, participants selected Image 9.96 (SD 1.99) as the most attractive version of themselves. This differed from the baseline by an average increment of 3.96 steps (95% CI for the mean difference: 3.18-4.75; t_{26}=10.35, *P*<.001; *d*=1.99). Together, the pattern of website clicks validates the appearance-based intervention, showing how participants (1) were sensitive to the photo transformations and (2) judged that they would look more attractive after a longer sleep duration (Figure 4).

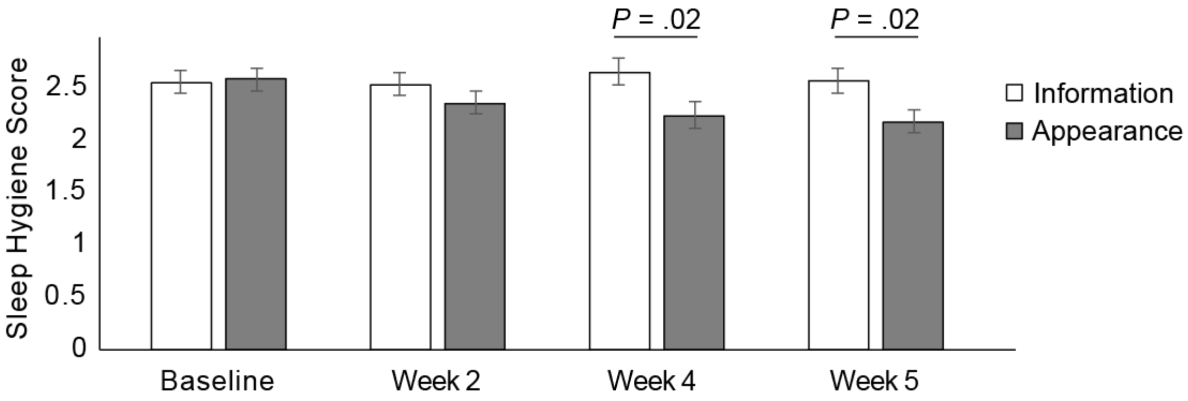
Figure 4. Participants viewed a continuum of 13 images transformed to show what they would look like with less or more sleep. When asked to identify their current selves, participants accurately identified an image close to the baseline (Image 6). They further judged that they would look more attractive following extended sleep. (Along the continuum, dots represent the mean images participants chose, and horizontal lines depict 95% CIs for the mean.).



Effect of the Appearance-Based Intervention: Sleep Hygiene

Compared with a standard information program, the appearance-based approach was more successful in changing behaviors that could result in better sleep. Namely, on the sleep hygiene measure, participants in the appearance group showed improved sleep hygiene after the intervention (effect of time: $F_{1,107.99}$=9.05, *P*=.003), whereas those in the information group did not show significant changes (effect of time: $F_{1,84.7}$=0.19, *P*=.66; Figure 5). Comparing across time points, no group differences were observed at baseline or at week 2 postintervention (smallest *P*=.27). However, the appearance group reported better sleep hygiene than the information group in week 4 ($F_{1,38}$=6.14, *P*=.02; 95% CI: −0.78 to −0.08; η_p^2=0.14) and week 5 ($F_{1,49}$=6.23, *P*=.02; 95% CI: −0.68 to −0.07; η_p^2=0.11).

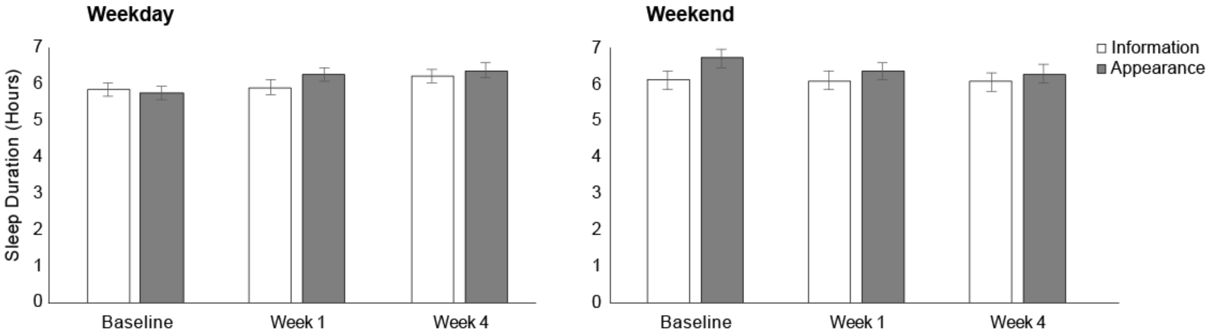
Figure 5. Mean sleep hygiene scores for participants in the information and appearance groups, plotted as a function of time (baseline and postintervention weeks 2, 4, and 5). A higher score corresponds to poorer sleep hygiene, and vertical lines represent 1 standard error of the mean.



Effect of the Appearance-Based Intervention: Actigraphy-Measured Sleep Duration

As shown in Figure 6, there was neither any significant effect of time on weekday sleep duration ($F_{2,96.71}$=0.19, P=.83) nor a significant interaction between time and group ($F_{2,96.71}$=1.38, P=.26). For weekend sleep duration, there was again no significant effect of time ($F_{2,103.77}$=0.50), nor a time × group interaction ($F_{2,103.77}$=0.16, P=.85; Figure 5). In Multimedia Appendix 2, Table S1 shows adjusted means for the primary mixed-effects analyses, and Multimedia Appendix 2, Table S2 shows the corresponding unadjusted means. Finally, Multimedia Appendix 2, Table S3 shows how sleep quality and insomnia-related cognition remained stable throughout the trial (smallest P for time or the time-group interaction=.48).

Figure 6. Mean sleep duration of participants in the information and appearance groups, plotted as a function of time (baseline and postintervention weeks 1 and 4). Vertical lines represent 1 standard error of the mean.



Discussion

Principal Findings

In this study, we developed a Web-based appearance-based intervention to promote sleep among habitual short sleepers. As a proof of concept, we first confirmed that participants (1) were sensitive to appearance changes and (2) judged their well-rested selves as more attractive. Correspondingly, this approach was more effective than a standard information program in promoting sleep hygiene, with benefits sustained for 1 month after the intervention.

Although these results are promising, the appearance-based intervention did not alter sleep duration itself. One possible explanation is that many paths lead to short sleep, of which a motivational component is only one. Thus, while appealing to attractiveness likely motivated change in presleep habits, it was not sufficient to alter sleep patterns. This account would suggest the need for a multipronged approach by public health agencies: for example, if sleep curtailment occurs because of immediate priorities such as pressing deadlines [16,17], an appearance-based intervention needs to be paired with techniques that help participants address their immediate priorities [44]. Alternatively, the intervention could be paired with structural changes such as a delayed school start time or reduced work hours, both of which have been found to increase sleep duration [45-47]. Further research will need to explore whether these pairings will ultimately address short sleep and how exactly motivation plays a role.

More broadly, our findings add to the growing body of research showing how an emphasis on appearances—though not sufficient as a stand-alone—may supplement traditional education in public health [25,26,28]. In recent years, appearance-based interventions have gained traction through the technological advances of photo-transformation software. These have been implemented in mobile- and Web-based applications [27,48,49] and have been applied across a range of health domains (sun protection [25], smoking cessation [26], nutrition [28], and—with this study—sleep). Given the potential reach of internet dissemination, one way forward may be to create a single appearance-based intervention that targets multiple health domains. If introduced in a timely manner (eg, incorporated into photo uploading features on a mobile dating application), this could boost existing public health campaigns, motivating healthy behaviors in a cost-effective manner.

Study Limitations

In presenting our study, we note several limitations in participant selection and randomization. First, we targeted participants with habitual short sleep, departing from earlier sleep education studies where participants were recruited regardless of sleep history [11]. Accordingly, we could not assess how the appearance-based intervention would affect the broader population. Second, we recruited university students, a group whose schedules are driven by high-stake deadlines within a short 13-week semester. It remains possible that stronger treatment effects would be observed in groups without similar deadlines (eg, adolescents or young working adults). Third, in choosing university students, we sought—in a proof-of-concept trial—to maximize the potential impact of our intervention. Across adulthood, although the likelihood of being a short sleeper is fairly even across age groups [50], concerns about physical appearances peak in young adults (≤24 years) and decrease linearly across the life span [51]. We thus reasoned that the selected age group (18-24 years) would be most amenable to an appearance-based intervention. However, in choosing these age cutoffs, we did not assess directly whether participants valued their appearances and recognize that our findings may not generalize to other age groups—for example, to older participants whose appearances may be affected by normal aging. Moving forward, selection of the appearance-based strategy can thus be customized based on the importance of physical attractiveness to the individual. Finally, as an eHealth trial, we were unable to blind the identity of which intervention participants were randomized to.

Conclusions

In conclusion, we designed in this study a novel intervention emphasizing physical attractiveness as a function of sleep patterns. This was assessed through a rigorous randomized controlled trial design with the following [52]: (1) the key intervention personalized (following best practices in appearance-based interventions [28]), (2) sleep duration objectively assessed (via actigraphy), and (3) participants monitored for 5 weeks. Using this design, we found that the appearance-based intervention was more effective than standard education in promoting sleep hygiene, a precursor for healthier sleep. At the same time, these effects did not translate to sleep extension, suggesting the need to assess how appearance-based strategies can be paired with other interventions. In summary, while urging further replication and extension of this work, our preliminary results suggest that beauty—the driving force that launched a thousand ships—may have an adjunctive role in promoting sleep.

Acknowledgments

This work was supported by grants awarded to JCJL by the Singapore Ministry of Education (start-up grant number: R-607-264-057-121 and AcRF Tier 1: IG15-B052). The authors gratefully acknowledge Benjamin Low, Pramesh S/O Prabakaran, Eveline Tan, Cheslie dela Cruz, Jim Seet, and Roger Liew for assistance with study preparation, data collection, and data entry.

Authors' Contributions

IP and SNT participated in research design, data collection, and the writing of the paper. KMV participated in research design, data collection, data analysis, and the writing of the paper. MWLC, DIP, and JCJL participated in research design, data analysis, and the writing of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Informed consent documentation.
[PDF File (Adobe PDF File)132 KB - jmir_v21i8e12500_app1.pdf]

Multimedia Appendix 2
Tables S1-3: Additional analyses.
[PDF File (Adobe PDF File)22 KB - jmir_v21i8e12500_app2.pdf]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File)2368 KB - jmir_v21i9e12500_app3.pdf]

References

1. Patel SR, Hu FB. Short sleep duration and weight gain: a systematic review. Obesity (Silver Spring) 2008 Mar;16(3):643-653 [FREE Full text] [doi: 10.1038/oby.2007.118] [Medline: 18239586]
2. Chen X, Beydoun MA, Wang Y. Is sleep duration associated with childhood obesity? A systematic review and meta-analysis. Obesity (Silver Spring) 2008 Feb;16(2):265-274 [FREE Full text] [doi: 10.1038/oby.2007.63] [Medline: 18239632]
3. Fatima Y, Doi SA, Mamun AA. Longitudinal impact of sleep on overweight and obesity in children and adolescents: a systematic review and bias-adjusted meta-analysis. Obes Rev 2015 Feb;16(2):137-149. [doi: 10.1111/obr.12245] [Medline: 25589359]
4. Cappuccio FP, Cooper D, D'Elia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. Eur Heart J 2011 Jun;32(12):1484-1492. [doi: 10.1093/eurheartj/ehr007] [Medline: 21300732]
5. Cappuccio FP, D'Elia L, Strazzullo P, Miller MA. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep 2010 May;33(5):585-592 [FREE Full text] [doi: 10.1093/sleep/33.5.585] [Medline: 20469800]
6. Gallicchio L, Kalesan B. Sleep duration and mortality: a systematic review and meta-analysis. J Sleep Res 2009 Jun;18(2):148-158 [FREE Full text] [doi: 10.1111/j.1365-2869.2008.00732.x] [Medline: 19645960]
7. Liu Y, Wheaton AG, Chapman DP, Cunningham TJ, Lu H, Croft JB. Prevalence of healthy sleep duration among adults--United States, 2014. MMWR Morb Mortal Wkly Rep 2016 Feb 19;65(6):137-141 [FREE Full text] [doi: 10.15585/mmwr.mm6506a1] [Medline: 26890214]
8. Consensus Conference Panel, Watson NF, Badr MS, Belenky G, Bliwise DL, Buxton OM, Non-Participating Observers, American Academy of Sleep Medicine Staff, et al. Recommended amount of sleep for a healthy adult: a joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. J Clin Sleep Med 2015 Jun 15;11(6):591-592 [FREE Full text] [doi: 10.5664/jcsm.4758] [Medline: 25979105]
9. Healthy People 2020. SH-4: Increase the Proportion of Adults Who Get Sufficient Sleep URL: https://www.healthypeople.gov/node/5261/data_details [accessed 2018-10-13] [WebCite Cache ID 738paPcWv]
10. Irish LA, Kline CE, Gunn HE, Buysse DJ, Hall MH. The role of sleep hygiene in promoting public health: a review of empirical evidence. Sleep Med Rev 2015 Aug;22:23-36 [FREE Full text] [doi: 10.1016/j.smrv.2014.10.001] [Medline: 25454674]
11. Blunden SL, Chapman J, Rigney GA. Are sleep education programs successful? The case for improved and consistent research efforts. Sleep Med Rev 2012 Aug;16(4):355-370. [doi: 10.1016/j.smrv.2011.08.002] [Medline: 22104441]
12. Cassoff J, Knäuper B, Michaelsen S, Gruber R. School-based sleep promotion programs: effectiveness, feasibility and insights for future research. Sleep Med Rev 2013 Jun;17(3):207-214. [doi: 10.1016/j.smrv.2012.07.001] [Medline: 23063417]
13. Chung KF, Chan MS, Lam YY, Lai CS, Yeung WF. School-based sleep education programs for short sleep duration in adolescents: a systematic review and meta-analysis. J Sch Health 2017 Jun;87(6):401-408. [doi: 10.1111/josh.12509] [Medline: 28463450]
14. Bakotić M, Radosević-Vidacek B, Koscec A. Educating adolescents about healthy sleep: experimental study of effectiveness of educational leaflet. Croat Med J 2009 Apr;50(2):174-181 [FREE Full text] [doi: 10.3325/cmj.2009.50.174] [Medline: 19399951]
15. Wolfson AR, Harkins E, Johnson M, Marco C. Effects of the Young Adolescent Sleep Smart Program on sleep hygiene practices, sleep health efficacy, and behavioral well-being. Sleep Health 2015 Sep;1(3):197-204 [FREE Full text] [doi: 10.1016/j.sleh.2015.07.002] [Medline: 29073440]
16. Orzech KM. A qualitative exploration of adolescent perceptions of healthy sleep in Tucson, Arizona, USA. Soc Sci Med 2013 Feb;79:109-116. [doi: 10.1016/j.socscimed.2012.05.001] [Medline: 22694989]
17. Suen LK, Hon KL, Tam WW. Association between sleep behavior and sleep-related factors among university students in Hong Kong. Chronobiol Int 2008 Sep;25(5):760-775. [doi: 10.1080/07420520802397186] [Medline: 18780202]
18. Kirby KN, Herrnstein RJ. Preference reversals due to myopic discounting of delayed reward. Psychol Sci 2016 May 6;6(2):83-89. [doi: 10.1111/j.1467-9280.1995.tb00311.x]
19. Green L, Myerson J. A discounting framework for choice with delayed and probabilistic rewards. Psychol Bull 2004 Sep;130(5):769-792 [FREE Full text] [doi: 10.1037/0033-2909.130.5.769] [Medline: 15367080]
20. Aggarwal S, Dizon-Ross R, Zucker A. Faculty Directory | The University of Chicago Booth School of Business. 2019. Incentivizing Behavioral Change: The Role of Time Preferences URL: https://faculty.chicagobooth.edu/rebecca.dizon-ross/research/papers/Incentive_Design.pdf
21. Purnell JQ, Gernes R, Stein R, Sherraden MS, Knoblock-Hahn A. A systematic review of financial incentives for dietary behavior change. J Acad Nutr Diet 2014 Jul;114(7):1023-1035 [FREE Full text] [doi: 10.1016/j.jand.2014.03.011] [Medline: 24836967]
22. Gerend MA, Cullen M. Effects of message framing and temporal context on college student drinking behavior. J Exp Soc Psychol 2008 Jul;44(4):1167-1173. [doi: 10.1016/j.jesp.2008.02.007]
23. de Bruijn GJ, Budding J. Temporal consequences, message framing, and consideration of future consequences: persuasion effects on adult fruit intake intention and resolve. J Health Commun 2016 Aug;21(8):944-953. [doi: 10.1080/10810730.2016.1179366] [Medline: 27442225]

24. Langlois JH, Kalakanis L, Rubenstein AJ, Larson A, Hallam M, Smoot M. Maxims or myths of beauty? A meta-analytic and theoretical review. Psychol Bull 2000 May;126(3):390-423. [doi: 10.1037//0033-2909.126.3.390] [Medline: 10825783]
25. Williams AL, Grogan S, Clark-Carter D, Buckley E. Appearance-based interventions to reduce ultraviolet exposure and/or increase sun protection intentions and behaviours: a systematic review and meta-analyses. Br J Health Psychol 2013 Feb;18(1):182-217. [doi: 10.1111/j.2044-8287.2012.02089.x] [Medline: 22989352]
26. Flett K, Clark-Carter D, Grogan S, Davey R. How effective are physical appearance interventions in changing smoking perceptions, attitudes and behaviours? A systematic review. Tob Control 2013 Mar;22(2):74-79. [doi: 10.1136/tobaccocontrol-2011-050236] [Medline: 22573734]
27. Burford O, Jiwa M, Carter O, Parsons R, Hendrie D. Internet-based photoaging within Australian pharmacies to promote smoking cessation: randomized controlled trial. J Med Internet Res 2013 Mar 26;15(3):e64 [FREE Full text] [doi: 10.2196/jmir.2337] [Medline: 23531984]
28. Whitehead RD, Ozakinci G, Perrett DI. A randomized controlled trial of an appearance-based dietary intervention. Health Psychol 2014 Jan;33(1):99-102. [doi: 10.1037/a0032322] [Medline: 23527517]
29. Axelsson J, Sundelin T, Ingre M, van Someren EJ, Olsson A, Lekander M. Beauty sleep: experimental study on the perceived health and attractiveness of sleep deprived people. Br Med J 2010 Dec 14;341:c6614 [FREE Full text] [doi: 10.1136/bmj.c6614] [Medline: 21156746]
30. Sundelin T, Lekander M, Kecklund G, van Someren EJ, Olsson A, Axelsson J. Cues of fatigue: effects of sleep deprivation on facial appearance. Sleep 2013 Sep 1;36(9):1355-1360 [FREE Full text] [doi: 10.5665/sleep.2964] [Medline: 23997369]
31. Sundelin T, Lekander M, Sorjonen K, Axelsson J. Negative effects of restricted sleep on facial appearance and social appeal. R Soc Open Sci 2017 May;4(5):160918 [FREE Full text] [doi: 10.1098/rsos.160918] [Medline: 28572989]
32. Kim MA, Kim EJ, Kang BY, Lee HK. The effects of sleep deprivation on the biophysical properties of facial skin. J Cosmet Dermatol Sci Appl 2017;7(1):34-47. [doi: 10.4236/jcdsa.2017.71004]
33. Talamas SN, Mavor KI, Axelsson J, Sundelin T, Perrett DI. Eyelid-openness and mouth curvature influence perceived intelligence beyond attractiveness. J Exp Psychol Gen 2016 May;145(5):603-620. [doi: 10.1037/xge0000152] [Medline: 26913618]
34. Brown FC, Buboltz WJ, Soper B. Development and evaluation of the sleep treatment and education program for students (STEPS). J Am Coll Health 2006;54(4):231-237. [doi: 10.3200/JACH.54.4.231-237] [Medline: 16450848]
35. 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 1;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
36. Gellis LA, Lichstein KL. Sleep hygiene practices of good and poor sleepers in the United States: an internet-based study. Behav Ther 2009 Mar;40(1):1-9. [doi: 10.1016/j.beth.2008.02.001] [Medline: 19187812]
37. Morin CM, Vallières A, Ivers H. Dysfunctional beliefs and attitudes about sleep (DBAS): validation of a brief version (DBAS-16). Sleep 2007 Nov;30(11):1547-1554 [FREE Full text] [doi: 10.1093/sleep/30.11.1547] [Medline: 18041487]
38. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
39. National Heart, Lung, and Blood Institute (NHLBI). 2011. In Brief: Your Guide to Healthy Sleep URL: http://www.nhlbi.nih.gov/files/docs/public/sleep/healthysleepfs.pdf [accessed 2015-08-01] [WebCite Cache ID 738q8K9P6]
40. Health Promotion Board, Singapore. 2015. Corporate Website URL: http://www.hpb.gov.sg/HOPPortal/content/conn/HOPUCM/path/ContributionFolders/uploadedFiles/HPB_Online/Educational_Materials/sleep.pdf
41. Tiddeman BP, Burt DM, Perrett DI. Prototyping and transforming facial textures for perception research. IEEE Comput Graph Appl 2001;21(5):42-50. [doi: 10.1109/38.946630]
42. Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. Trials 2014 Apr 23;15:139 [FREE Full text] [doi: 10.1186/1745-6215-15-139] [Medline: 24755011]
43. Wittmann M, Dinich J, Merrow M, Roenneberg T. Social jetlag: misalignment of biological and social time. Chronobiol Int 2006;23(1-2):497-509. [doi: 10.1080/07420520500545979] [Medline: 16687322]
44. Ariely D, Wertenbroch K. Procrastination, deadlines, and performance: self-control by precommitment. Psychol Sci 2002 May;13(3):219-224. [doi: 10.1111/1467-9280.00441] [Medline: 12009041]
45. Wheaton AG, Chapman DP, Croft JB. School start times, sleep, behavioral, health, and academic outcomes: a review of the literature. J Sch Health 2016 May;86(5):363-381 [FREE Full text] [doi: 10.1111/josh.12388] [Medline: 27040474]
46. Schiller H, Lekander M, Rajaleid K, Hellgren C, Åkerstedt T, Barck-Holst P, et al. The impact of reduced worktime on sleep and perceived stress - a group randomized intervention study using diary data. Scand J Work Environ Health 2017 Mar 1;43(2):109-116 [FREE Full text] [doi: 10.5271/sjweh.3610] [Medline: 27942734]
47. Lockley SW, Cronin JW, Evans EE, Cade BE, Lee CJ, Landrigan CP, Harvard Work Hours, Health and Safety Group. Effect of reducing interns' weekly work hours on sleep and attentional failures. N Engl J Med 2004 Oct 28;351(18):1829-1837. [doi: 10.1056/NEJMoa041404] [Medline: 15509816]

48. Brinker TJ, Seeger W. Photoaging mobile apps: a novel opportunity for smoking cessation? J Med Internet Res 2015 Jul 27;17(7):e186 [FREE Full text] [doi: 10.2196/jmir.4792] [Medline: 26215210]
49. Brinker TJ, Seeger W, Buslaff F. Photoaging mobile apps in school-based tobacco prevention: the mirroring approach. J Med Internet Res 2016 Jun 28;18(6):e183 [FREE Full text] [doi: 10.2196/jmir.6016] [Medline: 27352819]
50. Knutson KL, van Cauter E, Rathouz PJ, DeLeire T, Lauderdale DS. Trends in the prevalence of short sleepers in the USA: 1975-2006. Sleep 2010 Jan;33(1):37-45 [FREE Full text] [doi: 10.1093/sleep/33.1.37] [Medline: 20120619]
51. Pliner P, Chaiken S, Flett GL. Gender differences in concern with body weight and physical appearance over the life span. Pers Soc Psychol Bull 2016 Jul 2;16(2):263-273. [doi: 10.1177/0146167290162007]
52. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Br Med J 2010 Mar 23;340:c332 [FREE Full text] [doi: 10.1136/bmj.c332] [Medline: 20332509]

Abbreviations

DBAS-16: Dysfunctional Beliefs and Attitudes about Sleep scale
PSQI: Pittsburgh Sleep Quality Index

Edited by G Eysenbach; submitted 14.10.18; peer-reviewed by I Mircheva, K Miller, V Rocío; comments to author 04.04.19; revised version received 28.06.19; accepted 23.07.19; published 26.09.19.

Please cite as:
Perucho I, Vijayakumar KM, Talamas SN, Chee MWL, Perrett DI, Liu JCJ
A Web-Based Photo-Alteration Intervention to Promote Sleep: Randomized Controlled Trial
J Med Internet Res 2019;21(9):e12500
URL: https://www.jmir.org/2019/9/e12500
doi:10.2196/12500
PMID:31573913

Review

Information and Communication Systems to Tackle Barriers to Breastfeeding: Systematic Search and Review

Kymeng Tang[1], MSc; Kathrin Gerling[1], PhD; Wei Chen[2], PhD; Luc Geurts[1], PhD

[1]e-Media Research Lab, KU Leuven, Leuven, Belgium

[2]Center for Intelligent Medical Electronics, Fudan University, Shanghai, China

Corresponding Author:
Kymeng Tang, MSc
e-Media Research Lab
KU Leuven
Andreas Vesaliusstraat 13
Leuven, 3000
Belgium
Phone: 32 16 32 24 97
Email: kymeng.tang@kuleuven.be

Abstract

Background: Breastfeeding has many benefits for newborns, mothers, and the wider society. The World Health Organization recommends mothers to feed newborns exclusively with breastmilk for the first 6 months after birth, but breastfeeding rates in many countries fail to align with the recommendations because of various barriers. Breastfeeding success is associated with a number of determinants, such as self-efficacy, intention to breastfeed, and attitudes toward breastfeeding. Information and communication technology (ICT) has been leveraged to support breastfeeding by means of improving knowledge or providing practical supports in different maternal stages. Previous reviews have examined and summarized the effectiveness and credibility of interventions; however, no review has been done from a human-computer interaction perspective that is concerned with novel interaction techniques and the perspective of end users.

Objective: The objective of this review was to provide a comprehensive overview of existing digital interventions that support breastfeeding by investigating systems' objective, technology design, validation process, and quality attributes, both in terms of clinical parameters as well as usability and user experience.

Methods: A systematic search was conducted in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines in the following libraries: PubMed, Science Direct, Association for Computing Machinery Digital Library (ACM Digital Library), and Institute of Electrical and Electronics Engineers Xplore (IEEE Xplore).

Results: A total of 35 papers discussing 30 interventions were included. The main goals of these interventions were organized into 4 categories: breastfeeding education (n=12), breastfeeding promotion (n=8), communication support (n=6), and daily practical support (n=4). Of the interventions, 13 target mothers in the postnatal period. Most interventions come in forms of client communication systems (n=18), which frequently leverage Web technologies, text message, and mobile apps to provide breastfeeding support. Systems predominantly focus on mothers; validation strategies were rather heterogeneous, with 12 user studies concerning usability and user experience and 18 clinical validation studies focusing on the effects of the interventions on breastfeeding determinants; 5 papers did not report results. Generally, straightforward systems (eg, communication tools or Web-based solutions) seem to be more effective than complex interventions (eg, games).

Conclusions: Existing information and communication systems offer effective means of improving breastfeeding outcomes, but they do not address all relevant periods in parenthood (eg, the antenatal period) and often do not involve important stakeholders, such as partners. There is an opportunity to leverage more complex technical systems to open up avenues for the broader design of ICT to support breastfeeding; however, considering evaluation outcomes of existing support systems of higher complexity, such systems need to be designed with care.

(J Med Internet Res 2019;21(9):e13947) doi:10.2196/13947

KEYWORDS

breastfeeding; technology; review

Introduction

Background

Breastfeeding has many benefits for infants, mothers, and the wider society. A meta-analysis by Victora et al [1] suggests that infants who are breastfed show better immunity to child infections, are less likely to have oral occlusion misalignment and diabetes, and have higher intelligence [1,2]. Faster uterus recovery after delivery, shorter weight stabilization period [3], and natural contraception [1] are some of the immediate benefits for breastfeeding mothers. Other positive lifelong effects for breastfeeding mothers include protection against breast cancer and other diseases [1]. Beyond improved health outcomes, breastfeeding has far-reaching economic implications, such as a huge saving on health care costs [4,5]. However, breastfeeding rates in many countries fall short of achieving the World Health Organization (WHO) directive that recommends mothers to continue to exclusively breastfeed for 6 months [6].

Breastfeeding success is associated with a number of barriers and facilitators. Studies [7,8] have shown that birth complications, mode of delivery, medical conditions of mothers and infants, and physical availability negatively affect exclusiveness and initiation of breastfeeding. Socioeconomic parameters including age, marital status, income, education, and getting back to work also affect breastfeeding duration [7,9-11].

Self-efficacy [12], the perception of milk supply [9,13-17], and initiation and strength of intention to breastfeed [18-20] are reported as modifiable breastfeeding determinants. Self-efficacy, for example, is associated with the perception of being supported [9], exposure to breastfeeding activities [9], early breastfeeding practice [5], and past experience [9]. Perception of milk supply, on the other hand, depends on mothers' self-efficacy and level of knowledge and skills [12]. Sufficient breastfeeding knowledge and skills may help mothers avoid physical discomfort [21]. Breastfeeding initiation and intention are influenced by subjective norms [5], acknowledgment of the benefits [5], attitudes toward breastfeeding, and perception of being supported from mothers' social network [5,9], for example, family, partners, and health care professionals. Partners, in particular, potentially contribute to breastfeeding maintenance and feeding plan decision [22].

Prior Works

Information and communication technology (ICT) has been leveraged for breastfeeding support, for example, to provide breastfeeding education [23], through persuasive systems designed to encourage breastfeeding [24] or provide advice throughout the process [25]. Existing review papers [26-29] and meta-analyses [30] have addressed the credibility and effectiveness of specific technology-based interventions (eg, phone calls [28], websites [26], or mobile app use in China [27]). However, there is no comprehensive analysis of digital solutions to support breastfeeding from the perspective of technology design, taking into account the end users who systems are designed for, what experience they provide for the end users, and how they relate to the barriers and facilitators of breastfeeding. Here, we address this issue to map the landscape of the currently available solutions to support breastfeeding and identify challenges and opportunities for future studies in this area.

Objectives

This review aims to give an overview of the currently available digital interventions to address barriers and facilitators to breastfeeding by investigating the trend in technology design through the lens of human-computer interaction that focuses on the design, development, and evaluation of technology to solve real-world challenges that involve end users. Through this review, we reveal strengths and weaknesses in existing systems and underlying technologies, thereby identifying future opportunities for researchers in human-computer interaction, digital health, and public health, and hope to inform the work of health care professionals.

Toward these goals, we seek to explore this space via the following research questions:

RQ1. What type of ICT-based breastfeeding-supporting systems are available? Who are they intended for?

RQ2. How did the systems integrate into health care provisions to support breastfeeding? What technology platforms were used to achieve their goals?

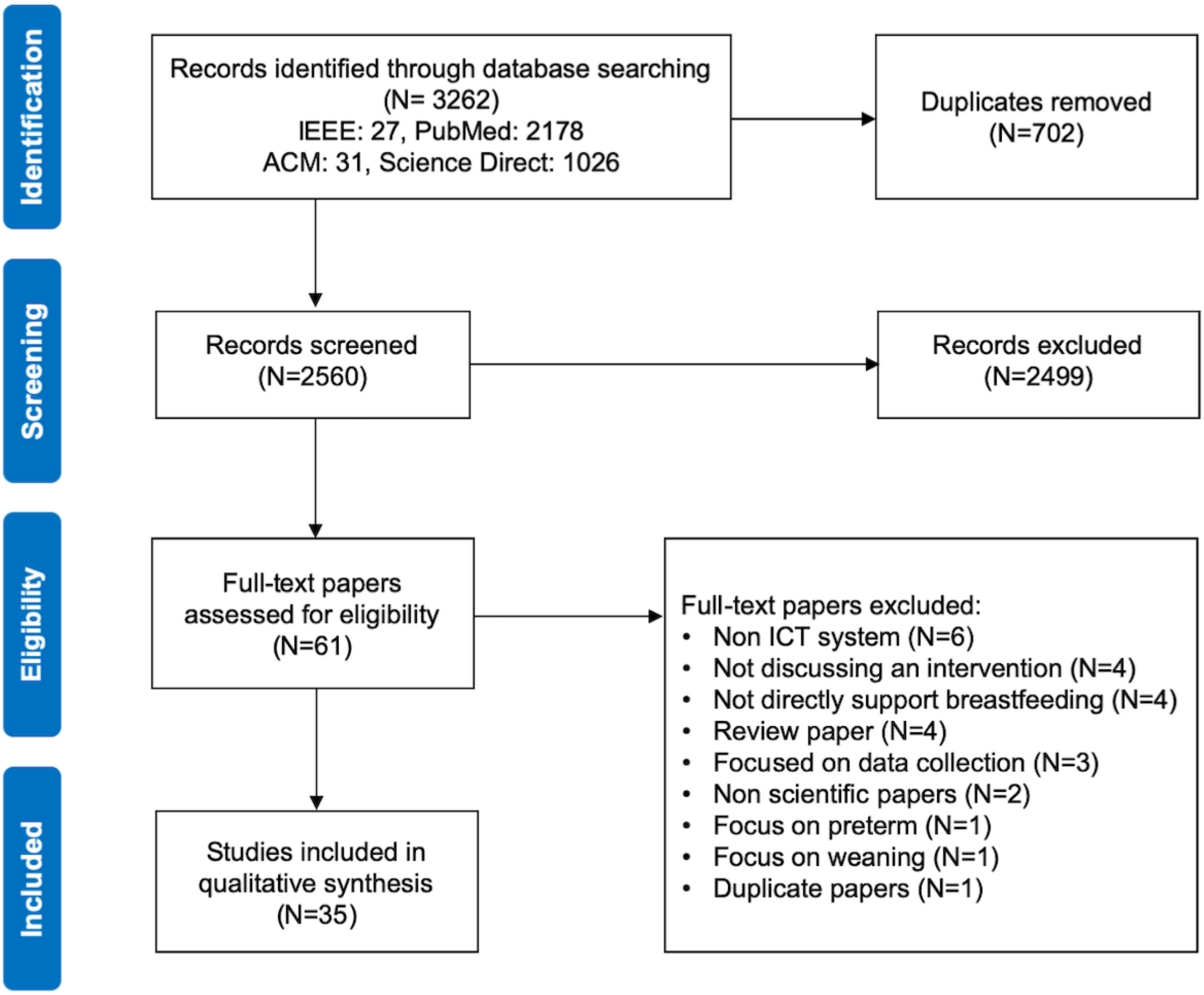
RQ3. Were the existing systems validated in terms of the experience they provide for users and their effectiveness in clinical terms?

Methods

Search Criteria and Procedure

This review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses [31] guidelines (see Figure 1) to select and process papers. Reflecting on our goals that entail a survey of interventions across disciplines, we queried papers in technical and medical libraries: Institute of Electrical and Electronics Engineers Xplore (IEEE Xplore), Association for Computing Machinery Digital Library (ACM Digital Library), Science Direct, and PubMed. We searched for papers that mentioned the terms: *breastfeeding* and *technology* and their variations in Title, Abstract, and Keywords fields. Combinations of search terms in the search query were modified for every database to preserve the intent of the query.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart of the paper selection process. IEEE: Institute of Electrical and Electronics Engineers; ACM: Association for Computing Machinery; ICT: information and communication technology.



Paper Selection Process

We queried papers published before November 2018 using a search strategy specified in Table 1. A total of 3262 resulting items were imported to Elsevier's Mendeley Desktop 1.9 for duplicates removal. At this stage, 702 duplicates were removed. On the basis of predefined inclusion and exclusion criteria, we performed a 2-phase eligibility scan: (1) title and abstract screening and (2) full-text screening for eligibility.

Inclusion Criteria

We included papers discussing digital interventions that satisfied all the following criteria: (1) full text in English, (2) targeted healthy human subjects regardless of breastfeeding role, and (3) provided direct or partial breastfeeding support.

We also included publications that discussed a system from different perspectives, for example, from experience-centric and clinical aspects. Work-in-progress papers were also included if their fully implemented system was not available.

Exclusion Criteria

We excluded papers that met any of the following criteria: (1) contained only abstract, (2) targeted subjects with special medical conditions, for example, preterm infants or severely sick mothers, and (3) reviews, books, book chapters, or reports or papers from scientific magazines.

If a paper of a fully implemented system is available, the work-in-progress papers of that system will be excluded.

Table 1. Literature search.

Terms	Search string
Breastfeeding	"breastfeeding" OR "breastfeed" OR "infant feeding" OR "bottle feeding"
	AND
Technology	"technology" OR "app" OR "application" OR "e-technology" OR "electronic health" OR "e-health" OR "ehealth" OR "mobile" OR "mobile health" OR "mhealth" OR "m-health" OR "computer" OR "internet" OR "web" OR "game" OR "play"

In the full-text assessment phase, we eliminated 26 publications that did not fit our criteria (see Figure 1). Finally, we included 35 eligible papers that described 30 digital interventions. Of the papers published, 31 were in medical or health informatics journals. The other 4 papers were presented in human-computer interaction and computer science conferences.

Data Extraction and Analysis

As a preprocessing step, papers that discussed the same systems were grouped. We defined an extraction scheme that comprised 3 corresponding perspectives.

We first coded the goals and methods of breastfeeding support of each system to extract necessary data to answer *RQ1*. Here, systems with similar goals were grouped to craft categories from the data. For systems that fell into multiple categories, we coded the highest priority goal. We then extracted excerpts of breastfeeding-supporting methods for each system. Finally, we coded each system by its intended audience and context of use, if made explicit by the authors.

To identify the technology platforms used by the systems (*RQ2*), we coded the technology used in the systems and the motivation behind the choice of platform for each system. In this step, we classify the systems according to the system classification framework for digital health intervention 1.0 of WHO [32]. After that, we further classify the systems by their technology platforms and rationale behind the choice of technology.

The last data extraction step is to identify the validation process and the reported effectiveness of the systems (*RQ3*). The papers were first examined to identify whether a user study or a clinical study had been conducted; a paper may be coded twice if it discusses both types of studies. In case of user studies, we coded their methods and results (eg, controlled lab study or in-the-wild deployment). For clinical studies, we extracted outcome measures, the statistical significance of results, the number of participants in the study, and possible limitations of results, if explicitly stated.

Results

Here, we present our results in line with the research questions.

Goals of Information and Communication Systems to Support Breastfeeding and Target Audience

This section aims to address *RQ1* by showing an overview of the identified interventions (summarized in Table 2) supplemented by a summary of their intended users and the context of use in Table 3. The interventions were organized based on their main objectives and the methods of breastfeeding support. Here, 4 main clusters emerged: (1) breastfeeding education for mothers and training tools for health care professionals, (2) breastfeeding promotion using persuasive techniques, (3) communication tools for mothers, partners, and health care professionals, and (4) daily practical breastfeeding support.

Breastfeeding Education

In total, 11 of the 30 systems aimed to provide breastfeeding education. Here, we separate educational interventions for mothers and training tools for health care practitioners.

A total of 9 papers [25,33-40] studied 8 educational interventions for mothers. The majority of these interventions focused on digitizing existing knowledge into generic learning modules with multimedia to aid explanation [25,34,35,38-40]. Geoghegan–Morphet et al [25] offered a Web-based forum in addition to educational resources. Besides these generic learning programs, some educational interventions were tailored. For instance, Abbass et al [33] involved an indigenous community to craft culturally relevant educational resources. Similarly, Joshi et al [35,39] customized educational content to suit Hispanic mothers.

Other modified generic systems for mothers emphasized learning via interactive exploration. For example, an interactive agent [36] guided mothers to explore different aspects of the breastfeeding process through simulated conversations. Grassley et al [37] proposed a quest-based game for mothers to build up their breastfeeding knowledge by completing playful quests. The game engaged users in various Web-based learning activities, such as reading and watching breastfeeding-related multimedia contents.

We identified 6 papers [23,41,42,44-46] that discussed 4 training tools for health care professionals. These systems aimed to improve breastfeeding knowledge and support skills. A total of 3 papers [42,45,46] evaluated Breastfeeding Basics [43], a Web intervention that provided modular breastfeeding educational resources and training materials for health care practitioners. A total of 2 papers [23,41] focused on designing course contents to be published on off-the-shelf e-learning platforms, such as Moodle and Blackboard Learn. Meanwhile, 1 study [44] made use of a website with 2 separated forums: one for pediatric residents and the other for breastfeeding mothers. This intervention allowed the pediatric residents to learn from their peers on one forum and apply their new skills when supporting mothers on the other forum.

Table 2. Summary of the included papers.

Reference	Intervention description	Intended user	Usage context
Breastfeeding education for mothers and training tools for health care professionals			
Abbass–Dick et al, 2018 [33]	Breastfeeding Web resources for indigenous audiences	Mothers	Unspecified
Cheng et al, 2003 [34]	General breastfeeding Web resources	Parents and parents-to-be	Prenatal
Joshi et al, 2015 [35]	Bilingual breastfeeding education on touch screen kiosks	Mothers	Postnatal
Edwards et al, 2013 [36]	Breastfeeding consultation with a computer agent	Mothers	Pre- and postnatal
Grassley et al, 2017 [37]	Breastfeeding quest game	Mothers	Prenatal
Huang et al, 2007 [38]	Web-based breastfeeding education program	Mothers	Prenatal
Joshi et al, 2016 [39]	Bilingual breastfeeding education on touch screen kiosks	Mothers	Postnatal
Labarere et al, 2011 [40]	Computer-based breastfeeding lessons on CD-ROM	Mothers	Postnatal
Geoghegan–Morphet et al, 2014 [25]	Web-based breastfeeding resource and virtual infant feeding support clinic	Parents and parents-to-be	Postnatal
Cianelli et al, 2015 [23]	Breastfeeding electronic learning program on Blackboard Learn platform	Caregivers	Nursing school
Colaceci et al, 2017 [41]	Breastfeeding electronic learning program on Moodle platform	Caregivers	Career training
Deloian et al, 2015 [42]	Publicly available breastfeeding Web education—Breastfeeding Basics [43]	Caregivers	Unspecified
Lasarte Velillas et al, 2007 [44]	Breastfeeding education via Web forums	Mothers and caregivers	Postnatal
Lewin and O’Connor, 2012 [45]	Publicly available breastfeeding Web education—Breastfeeding Basics [43]	Caregivers	Unspecified
O'Connor et al, 2011 [46]	Publicly available breastfeeding Web education—Breastfeeding Basics [43]	Caregivers	Unspecified
Breastfeeding encouragement			
Wardle et al, 2018 [47]	Milk Matters app to facilitate breastmilk donation	Mothers	Postnatal
Gallegos et al, 2014 [48]	Weekly 2-way SMS[a] to tackle breastfeeding challenges and encourage positive feeding practices	Mothers	Postnatal
Hmone et al, 2017 [24]	Breastfeeding promotion via SMS	Mothers	Pre- and postnatal
Jiang et al, 2014 [49]	SMS to improve breastfeeding practice	Mothers	Pre- and postnatal
Litterbach et al, 2017 [50]	Multichannel infant feeding support and motivation	Mothers	Postnatal
Maslowsky et al, 2016 [51]	Breastfeeding education and support through mobile phone calls	Mothers	Postnatal
Zunza et al, 2017 [52]	2-way SMS and motivational interview to promote breastfeeding among HIV positive mothers	Mothers	Postnatal
Unger et al, 2018 [53]	SMS to improve breastfeeding practices and contraception use	Mothers	Pre- and postnatal
Communication tools			
White et al, 2018 [54]	Milk Man app to engage fathers to support breastfeeding mothers	Fathers	Unspecified
Rojjanasrirat et al, 2012 [55]	Breastfeeding support via video conference	Mothers	Postnatal
Friesen et al, 2015 [56]	Breastfeeding support via video conference	Mothers	Pre- and postnatal
Demirci et al, 2018 [57]	Breastfeeding support via video conference on mobile	Mothers	Postnatal
Thomas and Shaikh, 2012 [58]	Breastfeeding support over the internet (email, phone call, and Web search)	Mothers	Postnatal
Giglia et al, 2015 [59]	Web-based resource and multichannel breastfeeding support	Mothers	Postnatal

Reference	Intervention description	Intended user	Usage context
Ahmed et al, 2016 [60]	Web-based breastfeeding diary	Mothers	Postnatal
White et al, 2016 [61]	Milk Man app to engage fathers to support breastfeeding mothers	Fathers	Unspecified
Daily practical supports			
Balaam et al, 2015 [62]	FeedFinder app to facilitate breastfeeding location search	Mothers	Postnatal
Wang et al, 2018 [63]	MoomMae app, breastfeeding diary and breastfeeding place finder	Mothers	Postnatal
dela Cruz and Mendoza, 2017 [64]	MilkTrack app to facilitate breastmilk donation process	Mothers	Unspecified
Chaovalit and Pongnumkul, 2017 [65]	MoomMae app, breastfeeding diary and breastfeeding place finder	Mothers	Postnatal

[a] SMS: short message service.

Table 3. Target population of the systems and context of use.

Target users and context	Reference
Mothers	
Unspecified	[33,64]
Prenatal	[37,38]
Pre- and postnatal	[24,36,49,53,56]
Postnatal	[35,39,44,47,48,51,52,55,57-60,62,63,65]
Fathers	
Unspecified	[54,61]
Parents and parents-to-be	
Prenatal	[34]
Postnatal	[25]
Mothers and practitioners	
Unspecified	[44]
Health care professionals	
Unspecified	[42,45,46]
Nursing school	[23]
Career training	[41]

Breastfeeding Promotion

In the second cluster, we identified 8 publications [24,47-53] that described systems that encouraged breastfeeding. The majority [24,48,49,52,53] of the systems utilized text messages to send out personalized breastfeeding tips and encouragement messages and sample breastfeeding experiences via short text message (SMS) responses. The content of the text messages was based on focus group discussions [53], WHO guidelines, expert inputs, and literature studies [49]. A total of 2 interventions [48,52] did not specify the source of message content and 1 system [24] used the Health Belief Model [66] to frame promotion strategies. Maslowsky et al [51] proposed a two-fold intervention that required nurses to call mothers 48 hours after hospital discharge. The idea is to deliver a maternal education session and a follow-up phone call to sample experience and provide support if needed. To keep users engaged with the intervention, Wardle et al [47] encouraged mothers to donate breastmilk using positive reinforcement techniques in their mobile app. The Growing Healthy Program [50] encouraged breastfeeding with weekly personalized motivational messages and on-demand breastfeeding resources aimed at parents with low socioeconomic status.

Communication Support

The third cluster comprises 7 papers [54-59,61] discussing 6 interventions that facilitate communication between peers and professionals. Of the systems, 3 [55-57] provided tele-lactation consultations via video conference for mothers with limited access to maternal care. Giglia et al [59] designed a multichannel support intervention through focus group; the intervention provides lactation supports through Web-based resources, email, and video call, all at the same time. For mothers who were not satisfied with their current caregivers, Thomas and Shaikh [58] described how mothers sought breastfeeding supports on the Web, using various information and communication tools. Besides interventions for mothers, White et al [54,61,67] used the social cognitive theory to design a mobile app that fostered peer support among expecting and new fathers by engaging them in discussions. This mobile app grouped fathers based on their baby's age and maternal stage of his partner.

Daily and Practical Support

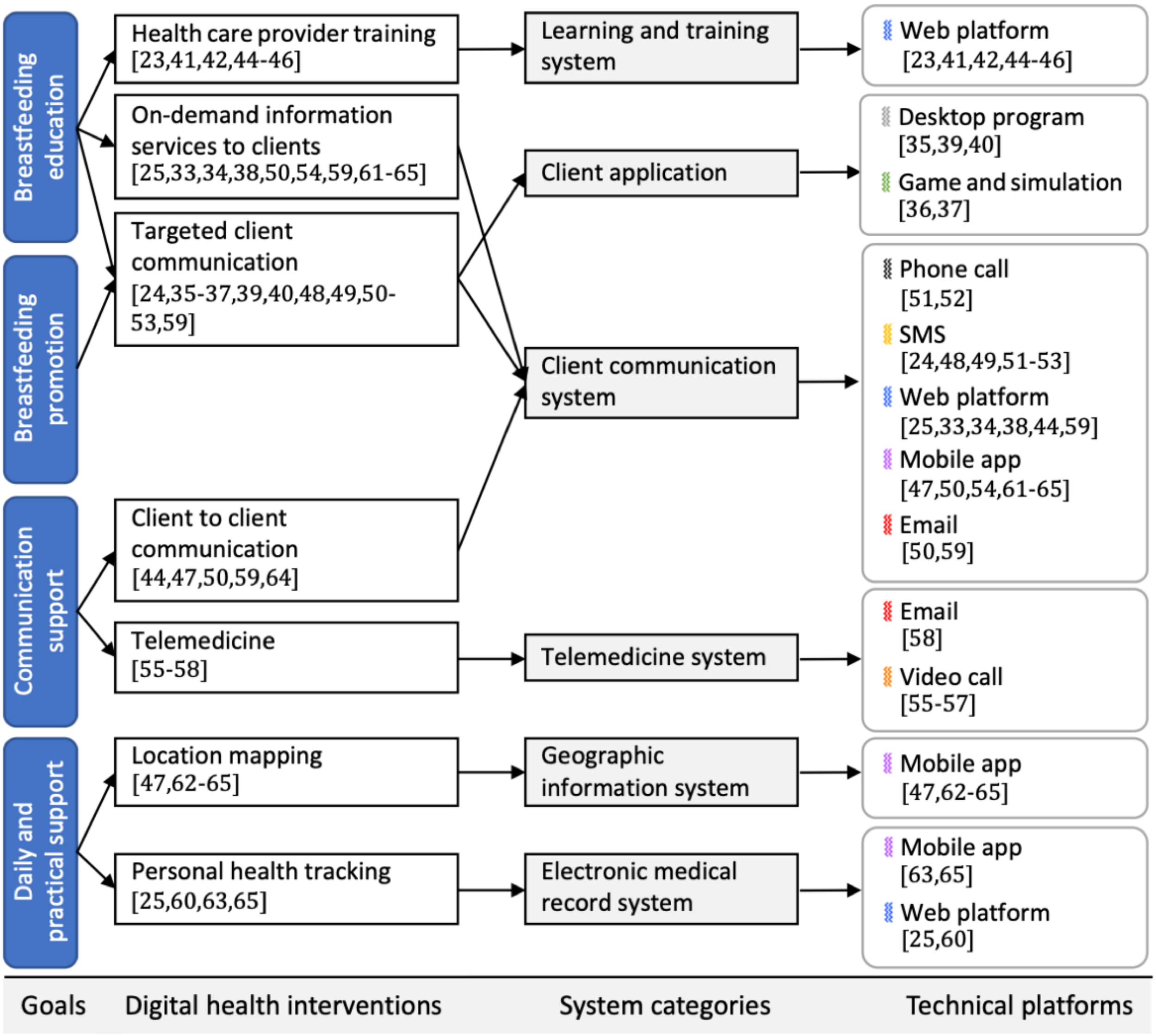
In the fourth cluster, we found 3 mobile apps and a Web intervention that practically supported breastfeeding. FeedFinder [62], MoomMae [63,65], and MilkTrack [64] are mobile apps that allow mothers to find, rate, and review suitable breastfeeding locations in public spaces. Besides breastfeeding location crowdsourcing, other features were incorporated as well, for example, MilkTrack [64] provides practical breastfeeding resources and a milk donation platform, whereas MoomMae [63,65] provides a personal breastfeeding diary. Ahmed et al [60], in contrary, only provided a Web breastfeeding diary. The diary allows health care practitioners to gain insights into mothers' breastfeeding experience and personalize their support.

On a general level, this shows that ICT to support breastfeeding addresses a broad range of goals; however, the strongest categories that emerged throughout our analysis were systems to provide breastfeeding education and promote breastfeeding.

Technical Platforms

To answer the second research question (*RQ2*), we characterized the digital interventions by their technology platform and system purpose. We structured our findings around the WHO system classification framework for digital health interventions [32] to maintain a mutual taxonomy across digital health domains. Here, the framework acts as connective tissues to relate system objectives to the used technology platform as shown in Figure 2. On the basis of the categorization scheme, the systems fall into the following 6 overlapping categories.

Figure 2. Taxonomy of the identified systems regarding intervention purpose and technology. SMS: standard message service.



Client Communication Systems

Client communication systems were used in 18 digital interventions [24,25,33,34,38,44,47-54,59,61-65] to communicate breastfeeding information to targeted clients (parents), provide on-demand information services to clients, and facilitate client-to-client communication. These systems leveraged 5 technology platforms to reach their intended audiences.

SMS appeared in 6 interventions [24,48,49,51-53] that aimed to deliver breastfeeding information and encouraging messages to mothers. Most of them [24,48,49,52,53] were inspired by the trend of leveraging SMSs to promote healthy behavior. Affordability and high availability also contributed to the popularity of this technology [48,49,52], particularly in developing countries, as 5 SMS-based breastfeeding interventions [24,49,51-53] were studied in Myanmar, China, Ecuador, and African countries, whereas only 1 [48] was introduced in Australia. The degree of user interactions within these SMS-based interventions is rather limited: 2 systems [24,48] required users to reply with a predefined code or a syntax, whereas the other 4 [49,51-53] invited free text response, but manual human processing was required to extract the content.

Phone calls, which operate on the same infrastructure as SMS, were used together with SMS in 2 interventions [51,52] to compensate for the lack of in-depth communication.

A total of 2 interventions [50,59] used *email* to initiate conversations with peers [59] and deliver notifications and weekly messages in a mobile health intervention [50] that also provided personalized infant-feeding information through a mobile app.

A total of 6 *mobile apps* [47,50,54,61-65] provide on-demand information service and client-to-client communication via platforms for online discussion and information sharing. Mobile app development gained popularity among intervention designers following the ubiquity of smartphones [47,50,54,61-65]. Mobile apps can be used to establish social connectivity for parents to connect to peers and foster information sharing, for instance, review public breastfeeding spaces [62,63,65], facilitate milk donation logistic [64], and share parenting experiences [50,54,61]. In addition to social connectivity, breastfeeding information can also be provided through mobile apps [47,50,64].

A total of 6 *Web-based interventions* [25,33,34,38,44,59] were designed to provide on-demand information services and facilitate client-to-client communication. Web technologies are ubiquitous, flexible, and easy to deploy and have low operational costs [33,34,38,41,46]. They can offer a rich user-interaction through the intuitive presentation of information (eg, using multimedia or interactive content) [34,38]. On top of that, they can be tailor-made to fit the exact requirements of an intervention [33,34,38,50,60]. Intervention designers can use the Web as a supporting channel for parents to look up breastfeeding information, for example, a Web-based virtual maternity clinic [25], or as a Web-based discussion platform to connect mothers with care providers and peers [25,44,59].

Learning and Training Systems and Client Applications

In this group, 4 learning and training systems [23,41,42,44-46] and 4 client applications [35-37,39,40] were identified.

A total of 4 learning and training interventions [23,41,42,44-46] use *Web platforms* to improve breastfeeding knowledge of health care providers. In this context, Web-based systems become a potential learning platform because they can provide comprehensive sets of educational tools, such as lessons, knowledge evaluations [23,41], and Web-based discussion [44], and be adapted to the busy schedule of health care practitioners [41,46].

A total of 4 client applications [35-37,39,40] are self-service *desktop programs* to communicate breastfeeding information. These systems came in various shapes: interactive software running on touchscreen kiosks in clinics [39,68], software provided via CD-ROM [40], *interactive agent* that simulated lactation consultation [36], and breastfeeding quest *game* [37]. These experimental interventions aimed to explore the potential of interactive learning as an alternative mode of communicating breastfeeding information.

Geographic Information and Electronic Medical Record System

A total of 2 mobile apps [47,62,64] are categorized as geographic information systems, 2 Web interventions [25,60] are electronic medical record systems, and 1 mobile app [63,65] situates in both system categories.

FeedFinder [62] and MilkTrack [64], both of which also belong to the client communication system category, made use of positioning capability of smartphones to help mothers find and map suitable breastfeeding locations in public. In addition to breastfeeding location services, the MoomMae app [63,65] offers a breastfeeding diary function to keep track of mothers' feeding habits, so they can accurately report feeding behaviors to health care professionals.

Both electronic medical record systems [25,60] are based on Web technologies. These interventions have an experience-sampling function that allows mothers to log their experience on a Web diary. The diary can be used as a personal record [25] or be accessed by health care providers to personalize their breastfeeding support tactics [60].

Telemedicine

Telemedicine came in forms of remote lactation consultation via video calls [55-57] and email communication [58]. All 3 identified video-call interventions [55-57] took place in the United States, which has a geographically dispersed population. Video calls break distance barriers for parents who live in rural areas [55,57] and facilitate the delivery of professional breastfeeding support when health care facilities are limited [56]. More importantly, the technology allows rich communication in real time (eg, allows health care practitioners to observe breastfeeding parents and demonstrate the feeding process [55]). Email was also used by mothers to initiate communication with health care professionals and inquire for breastfeeding information [58].

This illustrates the breadth of information and communication systems that are available to support breastfeeding, both in terms of audience and underlying technology. Here, it is interesting to note that interventions predominantly rely on existing communication channels, such as SMS or phone calls, or implement straightforward and well-explored technologies, such as Web portals.

Validation Process and Effectiveness of Existing Systems

To understand the validation process of the systems (*RQ3*), we coded the papers into 2 categories: user studies (focused on usability, ie, determining whether systems can be used effectively and efficiently, and user experience, ie, studies that explore whether systems engage users through a positive overall experience) and clinical validation studies. A total of 5 papers [24,25,52,61,65] without discussion of results were excluded from this section.

User Studies

A total of 20 systems were evaluated in terms of usability, user satisfaction, and user experience. Usability and user experience of 12 systems were evaluated through user studies, whereas 8 clinical validations included measures of user satisfaction and discussions of usability flaws as a part of their evaluation process. Here, we organize the results of the studies according to their objectives.

Among the reported user studies, a breadth of quantitative and qualitative research methods was employed; participant samples included interaction design experts, health care experts, and prospective end users.

Usability was assessed with different methods, depending on the objectives of the evaluation, including expert reviews and user studies. The evaluation of MoomMae [63] involved interviews and surveys with 21 breastfeeding mothers who used the app for 4 weeks. Although the app was rated as useful, postexposure interviews revealed a negative usability trend; common flaws were confusing user interface elements, ease of use, and screen readability. The bilingual desktop breastfeeding education [35] was evaluated using Nielsen usability heuristics [69] by 2 usability experts, and the paper reported 91 usability flaws across 271 screens. Cheng et al [34] involved 20 participants in different places to evaluate a Web-based breastfeeding education system with graphical content in terms of user satisfaction and learning outcomes. The study shows that graphical content shortened view time, increased user satisfaction, and did not negatively affect learning.

Besides usability, factors contributing to user experience were also explored. The Milk Matters [47], Milk Man [61], and FeedFinder [62] apps, for example, involved end users to conduct formative tests throughout their design and development cycles to minimize usability issues and adapt to user preferences. Formative user studies with stakeholders were also conducted in the development of a culturally adapted Web-based intervention for indigenous mothers [33]. The evaluation of these systems stressed on other quality attributes, such as engagement, emotion, and ease of use in specific circumstances. The analysis of comments collected after an in-the-wild release of FeedFinder [62] app showed improvement of mothers' confidence to breastfeed in public places. Engagement factors of the Milk Man app [61] included connectivity for fathers to seek and offer support, share personal experiences, and seek peer social connection. Likewise, the Growing Healthy Program [50] was found to be engaging because of high perceived usefulness, content that suits users' parental beliefs, and— from a technical perspective—the use of push notifications.

Other quality measures concerned user satisfaction and perceived usefulness of the systems, although some of them did not report full-fledged user studies. In general, most of the interventions were perceived positively, but some unexpected usability flaws were discovered. For example, the CD-ROM desktop program [40] reported good reviews from participants; however, only 119 of the 240 participants in the intervention unit (a total of 993 mothers from both intervention and control groups combined) actually accessed the intervention. Likewise, participants in a 2-way SMS intervention [48] reported not being able to remember the predefined SMS response codes, suggesting that systems would need to be adapted for effective in-the-wild deployment.

Clinical Validation Studies

A total of 16 systems (18 papers, see Multimedia Appendix 1) were validated in terms of clinical outcomes. Of those, we found 13 systems that fully or partially achieved their objectives, 1 system validation led to inconclusive results (not statistically significant), and 2 could not demonstrate a measurable effect. In Multimedia Appendix 1, we present a summary of the clinically validated systems and indicate reported effects on outcome parameters. Here, we organize systems based on their reported outcome measures.

The breastfeeding learning and training systems for health care professionals were effective with statistical significance in contributing to the knowledge of health care professionals and supporting skills, although selection bias might exist in the studies. The 3 in-the-wild evaluations [42,45,46] of the Breastfeeding Basics [43] educational program over the course of 9, 10, and 12 years with respective numbers of 15374, 18522, and 19671 participants using pre- and posttest consistently reported improvement in all aspects of breastfeeding knowledge among the included participants. It is worth mentioning that most of the participants included in these 3 studies were required to take part in the intervention for education or professional purpose. The Blackboard-based system [23] and the forum-based intervention [44] that were evaluated based on pre- and posttests with 86 nursing students and 42 pediatrics (recruited based on self-selection) were found to be effective in improving breastfeeding knowledge. Selection bias is also found in an evaluation of the Moodle-based intervention [41], with 15004 participants that showed positive effects on breastfeeding attitude and support practice; only participants who completed and passed the first-round evaluations were included in the study.

Some of the client communication systems were also effective and statistically significant in improving breastfeeding practice and determinants: an assessment of the clinic touchscreen kiosk breastfeeding intervention [39] with 46 mothers indicated an improvement in knowledge, self-efficacy, and intention to breastfeeding. A Web-based breastfeeding education [38] evaluated with 65 primigravids at their 29 to 36 weeks of pregnancy succeeded in improving breastfeeding rates, knowledge, and attitudes. A total of 2 [49,53] of the 5 SMS-based systems have reported statistically significant improvement of the duration of exclusive breastfeeding at the sixth month, with study cohorts of 582 and 298 participants. The evaluation of the postnatal phone-call support [51] with 178 new mothers through a follow-up phone call 3 months after the intervention commencement has reported a higher rate of exclusive breastfeeding and lower formula feeding rate among mothers in the intervention group, when compared with the control group. From a yearlong study with 414 mothers on multichannel lactation support [59], a positive long-term effect on exclusive breastfeeding in the intervention group was reported.

Not all effective systems fully achieved their goals. Despite showing significant improvements in breastfeeding duration and exclusivity, a study of the breastfeeding diary [60] with 96 participants did not show statistical significance in a second objective, decreasing postpartum depression. Similarly, a study with 24 participants on the Growing Healthy Program [50] has indicated an improvement in mothers' confidence in their choice of feeding method (either breastfeeding or bottle feeding) and perception of milk supply. However, feeding plans and intentions were not influenced as the decisions were taken before exposure to the intervention. A study on an SMS-based breastfeeding encouragement system [48] with 234 participants was shown effective in improving breastfeeding exclusivity but showed no statistically significant impact on self-efficacy, which might offer an explanation for its failure to increase breastfeeding rates.

Finally, client applications did not produce measurable effects. A 1-month controlled trial with a total of 993 mothers (from intervention and control groups) on a CD-ROM software [40] and a 25-participant pilot study of the breastfeeding game [37] failed to improve knowledge [37,40], breastfeeding rates [40], intention [37], and self-efficacy [37]. A pilot controlled study of the interactive agent [36] with 15 women showed no significant improvement in breastfeeding intention and self-efficacy and failed to alter attitudes toward breastfeeding. It is worth noting that these complex systems were neither custom-built to suit the end users nor involved the primary stakeholders in the development process but rather a modification of existing technology platforms.

Here, the heterogeneous nature of interventions and validation processes and outcome parameters limit the opportunity to conduct a meaningful comparison between interventions. On a general level, our results suggest that breastfeeding learning and training systems for health care professionals tend to succeed in improving breastfeeding knowledge among their intended users, whereas some client communication systems showed a positive effect on improving breastfeeding adherence. Other breastfeeding determinants were rarely influenced, and validation studies evaluating complex technology (eg, games) suggested that such systems were ineffective in their current designs.

Discussion

Principal Findings

This review gives an overview of digital interventions to support breastfeeding. We draw from research in computer science, engineering, and medical field to provide an overview of systems with a focus on technology, users, and outcomes. Our results show that the majority of systems were designed to address mothers in the postnatal period, either promoting or educating them about breastfeeding (*RQ1*). From a technical perspective, client communication systems were the most commonly used systems, with Web technologies, mobile apps, and SMS being the dominant platforms and only a small number of studies exploring more complex technologies, such as games (*RQ2*). System effectiveness was predominantly demonstrated in terms of improved breastfeeding knowledge, although improvement in behavioral outcomes might be because of systems that provided continuous proactive support. In terms of usability and user experience, results were mixed, with some systems failing to engage users (*RQ3*). Generally, our analysis suggests that straightforward technology fared better than complex systems, leaving room for an interesting debate. Here, we discuss these results; drawing from our findings, we further reflect on available systems through the lens of barriers and facilitators toward breastfeeding, and we outline opportunities for future research into the development of engaging technology interventions to promote breastfeeding.

Current Trends in Information and Communication Technology to Support Breastfeeding

Our review suggests that, based on the WHO system categorization, client communication systems are the predominant group of ICT to support breastfeeding. Systems in the group use SMS, Web platforms, and mobile apps to educate, support, and encourage mothers [24,25,33,34,38,44,47-54,59,61-65]. In terms of effectiveness, results show that educational interventions [23,25,33-42,44-46] for mothers and health care professionals tend to focus on improving knowledge, self-efficacy, intention, and attitude. In contrast, systems that facilitate communication [54-59,61] and encourage breastfeeding [24,47-53] can be successful in helping mothers to maintain breastfeeding practice, possibly because of the improvement of perceived support [9]. Interestingly, more complex interventions (eg, the breastfeeding game [37]) did not lead to a significant improvement in breastfeeding knowledge. Details about the contents and technical design process of the interventions are rather limited, which make a precise conclusion about factors contributing to the success, or failure, of these systems difficult to draw; however, what is known from other studies exploring the design of games for health is that this is a complex process and requires careful consideration to produce the desired outcomes. Here, future studies should explore how more complex breastfeeding interventions can be designed in an effective manner. Finally, although educational interventions were effective in improving knowledge, very few of them improved self-efficacy, intention, and attitude. This suggests that stakeholders currently need to engage with multiple systems to obtain the full benefits of ICT in this setting. Hence, it may be beneficial to consider pathways toward holistic system development to provide digital solutions that not only improve breastfeeding knowledge but also are designed in a way that also allows them to affect other breastfeeding determinants.

Opportunities for Future Research: Reflecting on Gaps in Existing Technology Through the Lens of Breastfeeding Determinants

When aligning the results of our analysis with breastfeeding determinants [5,7,9,12-17,21,22], it becomes clear that existing systems predominantly address knowledge and practical support, whereas other factors are omitted, particularly elements such as practical breastfeeding skills that affect the perception of milk supply and physical comfort [12]. Considering the use of ICT in this space, such a system would need to integrate more complex technology than those currently used (ie, Web-based systems and smartphone apps). Here, sensing systems offer an interesting design opportunity, and first attempts have been made to develop systems that can help track infant milk intake while breastfeeding [70,71]. Likewise, important determinants such as attitudes toward breastfeeding, intention, and initiation are not addressed by existing studies that predominantly target the postnatal period (13 of 30 systems), although the decision to breastfeed is usually taken in the third trimester of pregnancy [5,50]. Here, there would be potential for the development of technology solutions that address parents-to-be in the antenatal period, slowly introducing them to the topic, not only through educational but possibly also through experiential systems that give a glimpse at the breastfeeding process [72]. Finally, and perhaps most surprisingly, only 1 intervention [54,61] was specifically aimed at partners, despite research showing that partners in general, and fathers in particular, play a major role in a mother’s decision to breastfeed and success throughout breastfeeding journey [22]. Here, there is a large potential for future studies to be more inclusive, improving partners’ interest in breastfeeding, their knowledge, and their ability to support the breastfeeding mother. Generally speaking, future studies should explore the potential of emerging technologies, such as cross-platform interventions with multiple elements (eg, an educational app combined with channels for practical support), or immersive and tangible computing systems that can provide realistic insights into breastfeeding not just for mothers but also their partners (eg, a virtual reality breastfeeding simulation).

Comparison With Previous Studies

To the best of our knowledge, this review is the first of its kind to offer a broad overview of the available systems by considering aspects (goals, methods of support, target audiences, usage context, technology platforms, and their rational and validation process) other than clinical effectiveness. Other systematic reviews [29,30] analyzed the effectiveness of the intervention from the medical aspect, whereas some other reviews focused on specific technologies (eg, phone calls [28], websites [26], or mobile apps [27]) rather than surveying all available systems. Unlike previous approaches, we investigated different characteristics of the interventions instead of summarizing their effectiveness (eg, via meta-analysis). Here, our intent is to inspire researchers from other disciplines (eg, computer science) to contribute to this space.

Conclusions

There are various attempts to leverage ICT to encourage breastfeeding, aiming to improve breastfeeding education, persuade mothers to initiate and follow through with breastfeeding, and provide practical support. Our results show that although certain groups of systems are effective, they often only target 1 domain (eg, improving knowledge), requiring end users to engage with a multitude of systems to achieve good coverage. In addition, the majority of interventions exclusively targeted mothers and failed to consider other relevant stakeholders, most strikingly partners. Therefore, further study is necessary to explore how innovative concepts in ICT can be fully leveraged to provide comprehensive breastfeeding support, starting in the antenatal period and extending beyond the birth of the child, while engaging both parents. Through reflection on the way that existing systems (fail to) address determinants of breastfeeding, our review provides a first step toward outlining research opportunities for future study in this space.

Acknowledgments

This study was supported by KU Leuven Campus Group T and the Center for Intelligent Medical Electronics at Fudan University via a scholarship offered to the first author (KT).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example of raw data retrieved from the Polar H10, Polar A370, and Tempo HR.
[PDF File (Adobe PDF File), 70KB - jmir_v21i8e13947_app1.pdf]

References

1. Victora CG, Bahl R, Barros AJ, França GV, Horton S, Krasevec J, Lancet Breastfeeding Series Group. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. Lancet 2016 Jan 30;387(10017):475-490. [doi: 10.1016/S0140-6736(15)01024-7] [Medline: 26869575]

2. Horta BL, de Mola CL, Victora CG. Breastfeeding and intelligence: a systematic review and meta-analysis. Acta Paediatr 2015 Dec;104(467):14-19. [doi: 10.1111/apa.13139] [Medline: 26211556]
3. Heinig MJ, Dewey KG. Health effects of breast feeding for mothers: a critical review. Nutr Res Rev 1997 Jan;10(1):35-56. [doi: 10.1079/NRR19970004] [Medline: 19094257]
4. Bartick M, Reinhold A. The burden of suboptimal breastfeeding in the United States: a pediatric cost analysis. Pediatrics 2010 May;125(5):e1048-e1056. [doi: 10.1542/peds.2009-1616] [Medline: 20368314]
5. Rollins NC, Bhandari N, Hajeebhoy N, Horton S, Lutter CK, Martines JC, Lancet Breastfeeding Series Group. Why invest, and what it will take to improve breastfeeding practices? Lancet 2016 Jan 30;387(10017):491-504. [doi: 10.1016/S0140-6736(15)01044-2] [Medline: 26869576]
6. World Health Organization. Breastfeeding URL: https://www.who.int/nutrition/topics/exclusive_breastfeeding/en [accessed 2019-03-06] [WebCite Cache ID 76IvUP0MD]
7. Sayres S, Visentin L. Breastfeeding: uncovering barriers and offering solutions. Curr Opin Pediatr 2018 Aug;30(4):591-596. [doi: 10.1097/MOP.0000000000000647] [Medline: 29782384]
8. Leung GM, Lam TH, Ho LM. Breast-feeding and its relation to smoking and mode of delivery. Obstet Gynecol 2002 May;99(5 Pt 1):785-794. [doi: 10.1016/S0029-7844(02)01940-3] [Medline: 11978288]
9. Meedya S, Fahy K, Kable A. Factors that positively influence breastfeeding duration to 6 months: a literature review. Women Birth 2010 Dec;23(4):135-145. [doi: 10.1016/j.wombi.2010.02.002] [Medline: 20299299]
10. Ahluwalia IB, Morrow B, Hsia J. Why do women stop breastfeeding? Findings from the Pregnancy Risk Assessment and Monitoring System. Pediatrics 2005 Dec;116(6):1408-1412. [doi: 10.1542/peds.2005-0013] [Medline: 16322165]
11. Ong G, Yap M, Li FL, Choo TB. Impact of working status on breastfeeding in Singapore: evidence from the national breastfeeding survey 2001. Eur J Public Health 2005 Aug;15(4):424-430. [doi: 10.1093/eurpub/cki030] [Medline: 16030134]
12. Blixt I, Mårtensson LB, Ekström AC. Process-oriented training in breastfeeding for health professionals decreases women's experiences of breastfeeding challenges. Int Breastfeed J 2014;9:15 [FREE Full text] [doi: 10.1186/1746-4358-9-15] [Medline: 25221613]
13. Blyth R, Creedy DK, Dennis CL, Moyle W, Pratt J, de Vries SM. Effect of maternal confidence on breastfeeding duration: an application of breastfeeding self-efficacy theory. Birth 2002 Dec;29(4):278-284. [doi: 10.1046/j.1523-536X.2002.00202.x] [Medline: 12484390]
14. McCarter-Spaulding DE, Kearney MH. Parenting self-efficacy and perception of insufficient breast milk. J Obstet Gynecol Neonatal Nurs 2001;30(5):515-522. [doi: 10.1111/j.1552-6909.2001.tb01571.x] [Medline: 11572532]
15. Otsuka K, Dennis CL, Tatsuoka H, Jimba M. The relationship between breastfeeding self-efficacy and perceived insufficient milk among Japanese mothers. J Obstet Gynecol Neonatal Nurs 2008;37(5):546-555. [doi: 10.1111/j.1552-6909.2008.00277.x] [Medline: 18811774]
16. Arora S, McJunkin C, Wehrer J, Kuhn P. Major factors influencing breastfeeding rates: mother's perception of father's attitude and milk supply. Pediatrics 2000 Nov;106(5):E67. [doi: 10.1542/peds.106.5.e67] [Medline: 11061804]
17. Gatti L. Maternal perceptions of insufficient milk supply in breastfeeding. J Nurs Scholarsh 2008;40(4):355-363 [FREE Full text] [doi: 10.1111/j.1547-5069.2008.00234.x] [Medline: 19094151]
18. Mitra AK, Khoury AJ, Hinton AW, Carothers C. Predictors of breastfeeding intention among low-income women. Matern Child Health J 2004 Jun;8(2):65-70. [doi: 10.1023/B:MACI.0000025728.54271.27] [Medline: 15198173]
19. Riva E, Banderali G, Agostoni C, Silano M, Radaelli G, Giovannini M. Factors associated with initiation and duration of breastfeeding in Italy. Acta Paediatr 1999 Apr;88(4):411-415. [doi: 10.1111/j.1651-2227.1999.tb01132.x] [Medline: 10342540]
20. Donath SM, Amir LH, ALSPAC Study Team. Relationship between prenatal infant feeding intention and initiation and duration of breastfeeding: a cohort study. Acta Paediatr 2003;92(3):352-356. [doi: 10.1080/08035250310009293] [Medline: 12725552]
21. Kent JC, Ashton E, Hardwick CM, Rowan MK, Chia ES, Fairclough KA, et al. Nipple pain in breastfeeding mothers: incidence, causes and treatments. Int J Environ Res Public Health 2015 Sep 29;12(10):12247-12263 [FREE Full text] [doi: 10.3390/ijerph121012247] [Medline: 26426034]
22. Scott JA, Landers MC, Hughes RM, Binns CW. Factors associated with breastfeeding at discharge and duration of breastfeeding. J Paediatr Child Health 2001 Jun;37(3):254-261. [doi: 10.1046/j.1440-1754.2001.00646.x] [Medline: 11468040]
23. Cianelli R, Villegas N, Azaiza K, Henderson S, Hooshmand M, Peragallo N. Developing and testing an online breastfeeding training among undergraduate nursing students. Clin Nurs Stud 2015;3(1):82-88 [FREE Full text] [doi: 10.5430/cns.v3n1p82] [Medline: 27766266]
24. Hmone MP, Li M, Alam A, Dibley MJ. Mobile phone short messages to improve exclusive breastfeeding and reduce adverse infant feeding practices: protocol for a randomized controlled trial in Yangon, Myanmar. JMIR Res Protoc 2017 Jun 28;6(6):e126 [FREE Full text] [doi: 10.2196/resprot.7679] [Medline: 28659252]
25. Geoghegan-Morphet N, Yuen D, Rai E, Angelini M, Christmas M, da Silva O. Development and implementation of a novel online breastfeeding support resource: the maternal virtual infant nutrition support clinic. Breastfeed Med 2014 Dec;9(10):520-523. [doi: 10.1089/bfm.2014.0051] [Medline: 25244066]

26. Shaikh U, Scott BJ. Extent, accuracy, and credibility of breastfeeding information on the internet. J Hum Lact 2005 May;21(2):175-183. [doi: 10.1177/0890334405275824] [Medline: 15886343]
27. Zhao J, Freeman B, Li M. How do infant feeding apps in China measure up? A content quality assessment. JMIR Mhealth Uhealth 2017 Dec 6;5(12):e186 [FREE Full text] [doi: 10.2196/mhealth.8764] [Medline: 29212627]
28. Oriá MO, Dodou HD, Chaves AF, Santos LM, Ximenes LB, Vasconcelos CT. Effectiveness of educational interventions conducted by telephone to promote breastfeeding: a systematic review of the literature. Rev Esc Enferm USP 2018 Jul 23;52:e03333 [FREE Full text] [doi: 10.1590/S1980-220X2017024303333] [Medline: 30043928]
29. McArthur L, Otosen M, Picarella L. Technology for breastfeeding support: a systematic review. J Inform Nu 2018;3(1):21-26,32 [FREE Full text]
30. Lau Y, Htun TP, Tam WS, Klainin-Yobas P. Efficacy of e-technologies in improving breastfeeding outcomes among perinatal women: a meta-analysis. Matern Child Nutr 2016 Jul;12(3):381-401. [doi: 10.1111/mcn.12202] [Medline: 26194599]
31. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
32. World Health Organization. 2018. Classification of Digital Health Interventions v1.0: A Shared Language to Describe the Uses of Digital Technology for Health URL: https://apps.who.int/iris/handle/10665/260480 [accessed 2019-05-15]
33. Abbass-Dick J, Brolly M, Huizinga J, Newport A, Xie F, George S, et al. Designing an ehealth breastfeeding resource with indigenous families using a participatory design. J Transcult Nurs 2018 Sep;29(5):480-488. [doi: 10.1177/1043659617731818] [Medline: 29308703]
34. Cheng WC, Thompson CB, Smith JA, Pugh L, Stanley C. A web-based breastfeeding education program. J Perinat Educ 2003;12(1):29-41 [FREE Full text] [doi: 10.1624/105812403X106702] [Medline: 17273328]
35. Joshi A, Perin DM, Amadi C, Trout K. Evaluating the usability of an interactive, bi-lingual, touchscreen-enabled breastfeeding educational programme: application of Nielson's heuristics. J Innov Health Inform 2015 Mar 5;22(2):265-274 [FREE Full text] [doi: 10.14236/jhi.v22i2.71] [Medline: 26245240]
36. Edwards RA, Bickmore T, Jenkins L, Foley M, Manjourides J. Use of an interactive computer agent to support breastfeeding. Matern Child Health J 2013 Dec;17(10):1961-1968. [doi: 10.1007/s10995-013-1222-0] [Medline: 23329167]
37. Grassley JS, Connor KC, Bond L. Game-based online antenatal breastfeeding education: a pilot. Appl Nurs Res 2017 Feb;33:93-95 [FREE Full text] [doi: 10.1016/j.apnr.2016.10.011] [Medline: 28096030]
38. Huang MZ, Kuo SC, Avery MD, Chen W, Lin KC, Gau ML. Evaluating effects of a prenatal web-based breastfeeding education programme in Taiwan. J Clin Nurs 2007 Aug;16(8):1571-1579. [doi: 10.1111/j.1365-2702.2006.01843.x] [Medline: 17655546]
39. Joshi A, Amadi C, Meza J, Aguire T, Wilhelm S. Evaluation of a computer-based bilingual breastfeeding educational program on breastfeeding knowledge, self-efficacy and intent to breastfeed among rural Hispanic women. Int J Med Inform 2016 Jul;91:10-19. [doi: 10.1016/j.ijmedinf.2016.04.001] [Medline: 27185505]
40. Labarère J, Gelbert-Baudino N, Laborde L, Arragain D, Schelstraete C, François P. CD-ROM-based program for breastfeeding mothers. Matern Child Nutr 2011 Jul;7(3):263-272. [doi: 10.1111/j.1740-8709.2009.00235.x] [Medline: 21689269]
41. Colaceci S, Giusti A, Chapin EM, Bettinelli ME, de Angelis A, Zambri F, et al. E-learning to improve healthcare professionals' attitudes and practices on breastfeeding. Breastfeed Med 2017 Dec;12(10):629-636. [doi: 10.1089/bfm.2017.0060] [Medline: 28926281]
42. Deloian BJ, Lewin LO, O'Connor ME. Use of a web-based education program improves nurses' knowledge of breastfeeding. J Obstet Gynecol Neonatal Nurs 2015;44(1):77-86. [doi: 10.1111/1552-6909.12534] [Medline: 25580851]
43. Breastfeeding Basics. URL: https://www.breastfeedingbasics.com [accessed 2019-03-06] [WebCite Cache ID 76fMknbsq]
44. Velillas JJ, Hernández-Aguilar MT, Alonso CR, Díaz-Gómez NM, Papí AG, Espuelas CF, et al. A breastfeeding e-learning project based on a web forum. Breastfeed Med 2007 Dec;2(4):219-228. [doi: 10.1089/bfm.2007.0008] [Medline: 18081459]
45. Lewin LO, O'Connor ME. 'BreastfeedingBasics': web-based education that meets current knowledge competencies. J Hum Lact 2012 Aug;28(3):407-413. [doi: 10.1177/0890334411435990] [Medline: 22550095]
46. O'Connor ME, Brown EW, Lewin LO. An internet-based education program improves breastfeeding knowledge of maternal-child healthcare providers. Breastfeed Med 2011 Dec;6(6):421-427. [doi: 10.1089/bfm.2010.0061] [Medline: 21029021]
47. Wardle CJ, Green M, Mburu CW, Densmore M. Exploring Co-Design with Breastfeeding Mothers. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI'18; April 21-26, 2018; Montreal, Canada. [doi: 10.1145/3173574.3174056]
48. Gallegos D, Russell-Bennett R, Previte J, Parkinson J. Can a text message a week improve breastfeeding? BMC Pregnancy Childbirth 2014 Nov 6;14:374 [FREE Full text] [doi: 10.1186/s12884-014-0374-2] [Medline: 25369808]
49. Jiang H, Li M, Wen LM, Hu Q, Yang D, He G, et al. Effect of short message service on infant feeding practice: findings from a community-based study in Shanghai, China. JAMA Pediatr 2014 May;168(5):471-478. [doi: 10.1001/jamapediatrics.2014.58] [Medline: 24639004]

50. Litterbach EK, Russell CG, Taki S, Denney-Wilson E, Campbell KJ, Laws RA. Factors influencing engagement and behavioral determinants of infant feeding in an mhealth program: qualitative evaluation of the growing healthy program. JMIR Mhealth Uhealth 2017 Dec 18;5(12):e196 [FREE Full text] [doi: 10.2196/mhealth.8515] [Medline: 29254908]
51. Maslowsky J, Frost S, Hendrick CE, Cruz FO, Merajver SD. Effects of postpartum mobile phone-based education on maternal and infant health in Ecuador. Int J Gynaecol Obstet 2016 Jul;134(1):93-98 [FREE Full text] [doi: 10.1016/j.ijgo.2015.12.008] [Medline: 27126905]
52. Zunza M, Cotton MF, Mbuagbaw L, Lester R, Thabane L. Interactive weekly mobile phone text messaging plus motivational interviewing in promotion of breastfeeding among women living with HIV in South Africa: study protocol for a randomized controlled trial. Trials 2017 Jul 17;18(1):331 [FREE Full text] [doi: 10.1186/s13063-017-2079-0] [Medline: 28716146]
53. Unger JA, Ronen K, Perrier T, DeRenzi B, Slyker J, Drake AL, et al. Short message service communication improves exclusive breastfeeding and early postpartum contraception in a low- to middle-income country setting: a randomised trial. BJOG 2018 Nov;125(12):1620-1629. [doi: 10.1111/1471-0528.15337] [Medline: 29924912]
54. White BK, Giglia RC, Scott JA, Burns SK. How new and expecting fathers engage with an app-based online forum: qualitative analysis. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e144 [FREE Full text] [doi: 10.2196/mhealth.9999] [Medline: 29914862]
55. Rojjanasrirat W, Nelson EL, Wambach KA. A pilot study of home-based videoconferencing for breastfeeding support. J Hum Lact 2012 Nov;28(4):464-467. [doi: 10.1177/0890334412449071] [Medline: 22802355]
56. Friesen CA, Hormuth LJ, Petersen D, Babbitt T. Using videoconferencing technology to provide breastfeeding support to low-income women: connecting hospital-based lactation consultants with clients receiving care at a community health center. J Hum Lact 2015 Nov;31(4):595-599. [doi: 10.1177/0890334415601088] [Medline: 26297347]
57. Demirci J, Kotzias V, Bogen DL, Ray KN, Uscher-Pines L. Telelactation via mobile app: perspectives of rural mothers, their care providers, and lactation consultants. Telemed J E Health 2018 Sep 13:- (epub ahead of print). [doi: 10.1089/tmj.2018.0113] [Medline: 30212280]
58. Thomas JR, Shaikh U. Use of electronic communication by physician breastfeeding experts for support of the breastfeeding mother. Breastfeed Med 2012 Dec;7(6):393-396. [doi: 10.1089/bfm.2011.0133] [Medline: 22612624]
59. Giglia R, Cox K, Zhao Y, Binns CW. Exclusive breastfeeding increased by an internet intervention. Breastfeed Med 2015;10(1):20-25. [doi: 10.1089/bfm.2014.0093] [Medline: 25358119]
60. Ahmed AH, Roumani AM, Szucs K, Zhang L, King D. The effect of interactive web-based monitoring on breastfeeding exclusivity, intensity, and duration in healthy, term infants after hospital discharge. J Obstet Gynecol Neonatal Nurs 2016;45(2):143-154 [FREE Full text] [doi: 10.1016/j.jogn.2015.12.001] [Medline: 26779838]
61. White BK, Burns SK, Giglia RC, Scott JA. Designing evaluation plans for health promotion mHealth interventions: a case study of the Milk Man mobile app. Health Promot J Austr 2016 Feb;27(3):198-203. [doi: 10.1071/HE16041] [Medline: 27784506]
62. Balaam M, Comber R, Jenkins E, Sutton S, Garbett A. FeedFinder: A Location-Mapping Mobile Application for Breastfeeding Women. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI'15; April 18-23, 2015; Seoul, Republic of Korea p. 1709-1718. [doi: 10.1145/2702123.2702328]
63. Wang CJ, Chaovalit P, Pongnumkul S. A breastfeed-promoting mobile app intervention: usability and usefulness study. JMIR Mhealth Uhealth 2018 Jan 26;6(1):e27 [FREE Full text] [doi: 10.2196/mhealth.8337] [Medline: 29374000]
64. dela Cruz DR, Mendoza DM. Milktrack: Design and Development of Mobile Application and Logistics System in Empowering Breastfeeding Practice in the Philippines. In: Proceedings of the 2017 IEEE Region 10 Conference. 2017 Presented at: TENCON'17; November 5-8, 2017; Penang, Malaysia p. 2242-2246. [doi: 10.1109/TENCON.2017.8228234]
65. Chaovalit P, Pongnumkul S. Moommae: A Location-Aware Mobile Information System for Breastfeeding Mothers in Thailand. In: Proceedings of the 10th International Conference on Ubi-Media Computing and Workshops. 2017 Presented at: Ubi-Media'17; August 1-4, 2017; Pattaya, Thailand. [doi: 10.1109/UMEDIA.2017.8074104]
66. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q 1988;15(2):175-183. [doi: 10.1177/109019818801500203] [Medline: 3378902]
67. White BK, Martin A, White JA, Burns SK, Maycock BR, Giglia RC, et al. Theory-based design and development of a socially connected, gamified mobile app for men about breastfeeding (Milk Man). JMIR Mhealth Uhealth 2016 Jun 27;4(2):e81 [FREE Full text] [doi: 10.2196/mhealth.5652] [Medline: 27349756]
68. Joshi A, Amadi C, Meza J, Aguirre T, Wilhelm S. Comparison of socio-demographic characteristics of a computer based breastfeeding educational intervention among rural Hispanic women. J Community Health 2015 Oct;40(5):993-1001. [doi: 10.1007/s10900-015-0023-3] [Medline: 25868495]
69. Nielsen J. Usability Inspection Methods. In: Proceedings of the Conference on Companion on Human Factors in Computing Systems. 1995 Presented at: CHI'95; May 7-11, 1995; Denver, Colorado, USA p. 413-414. [doi: 10.1145/223355.223730]
70. Farooq M, Chandler-Laney P, Hernandez-Reif M, Sazonov E. A Wireless Sensor System for Quantification of Infant Feeding Behavior. In: Proceedings of the Conference on Wireless Health. 2015 Presented at: WH'15; October 14-16, 2015; Bethesda, Maryland. [doi: 10.1145/2811780.2811934]

71. Tamilia E, Formica D, Visco AM, Scaini A, Taffoni F. An automated system for quantitative analysis of newborns' oral-motor behavior and coordination during bottle feeding. Conf Proc IEEE Eng Med Biol Soc 2015;2015:7386-7389. [doi: 10.1109/EMBC.2015.7320098] [Medline: 26737998]
72. Gerling K, Hicks K, Buttrick L, Headleand C, Williams B, Hall J, et al. Potential and Limitations of Playful Technology to Support Infant Feeding. In: Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. 2018 Presented at: CHI EA'18; October 28-31, 2018; Melbourne, VIC, Australia p. 431-437. [doi: 10.1145/3270316.3271519]

Abbreviations

CD-ROM: Compact Disk Read Only Memory
ICT: information and communication technology
SMS: short message service
WHO: World Health Organization

Edited by G Eysenbach; submitted 11.03.19; peer-reviewed by Y Lau, P Chaovalit; comments to author 06.05.19; revised version received 18.06.19; accepted 19.07.19; published 27.09.19.

Please cite as:
Tang K, Gerling K, Chen W, Geurts L
Information and Communication Systems to Tackle Barriers to Breastfeeding: Systematic Search and Review
J Med Internet Res 2019;21(9):e13947
URL: https://www.jmir.org/2019/9/e13947/
doi:10.2196/13947
PMID:31573903

Review

Literature on Wearable Technology for Connected Health: Scoping Review of Research Trends, Advances, and Barriers

Tatjana Loncar-Turukalo[1], PhD; Eftim Zdravevski[2], PhD; José Machado da Silva[3], PhD; Ioanna Chouvarda[4], PhD; Vladimir Trajkovik[2], PhD

[1]Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia
[2]Faculty of Computer Science and Engineering, Saints Cyril and Methodius University, Skopje, North Macedonia
[3]Institute for Systems and Computer Engineering, Technology and Science, Faculty of Engineering, University of Porto, Porto, Portugal
[4]Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece

Corresponding Author:
Tatjana Loncar-Turukalo, PhD
Faculty of Technical Sciences
University of Novi Sad
Trg Dositeja Obradovica 6
Novi Sad, 21000
Serbia
Phone: 381 691463297
Email: turukalo@uns.ac.rs

Abstract

Background: Wearable sensing and information and communication technologies are key enablers driving the transformation of health care delivery toward a new model of connected health (CH) care. The advances in wearable technologies in the last decade are evidenced in a plethora of original articles, patent documentation, and focused systematic reviews. Although technological innovations continuously respond to emerging challenges and technology availability further supports the evolution of CH solutions, the widespread adoption of wearables remains hindered.

Objective: This study aimed to scope the scientific literature in the field of pervasive wearable health monitoring in the time interval from January 2010 to February 2019 with respect to four important pillars: technology, safety and security, prescriptive insight, and user-related concerns. The purpose of this study was multifold: identification of (1) trends and milestones that have driven research in wearable technology in the last decade, (2) concerns and barriers from technology and user perspective, and (3) trends in the research literature addressing these issues.

Methods: This study followed the scoping review methodology to identify and process the available literature. As the scope surpasses the possibilities of manual search, we relied on the natural language processing tool kit to ensure an efficient and exhaustive search of the literature corpus in three large digital libraries: Institute of Electrical and Electronics Engineers, PubMed, and Springer. The search was based on the keywords and properties to be found in articles using the search engines of the digital libraries.

Results: The annual number of publications in all segments of research on wearable technology shows an increasing trend from 2010 to February 2019. The technology-related topics dominated in the number of contributions, followed by research on information delivery, safety, and security, whereas user-related concerns were the topic least addressed. The literature corpus evidences milestones in sensor technology (miniaturization and placement), communication architectures and fifth generation (5G) cellular network technology, data analytics, and evolution of cloud and edge computing architectures. The research lag in battery technology makes energy efficiency a relevant consideration in the design of both sensors and network architectures with computational offloading. The most addressed user-related concerns were (technology) acceptance and privacy, whereas research gaps indicate that more efforts should be invested into formalizing clear use cases with timely and valuable feedback and prescriptive recommendations.

Conclusions: This study confirms that applications of wearable technology in the CH domain are becoming mature and established as a scientific domain. The current research should bring progress to sustainable delivery of valuable recommendations, enforcement of privacy by design, energy-efficient pervasive sensing, seamless monitoring, and low-latency 5G communications. To complement technology achievements, future work involving all stakeholders providing research evidence on improved care pathways and cost-effectiveness of the CH model is needed.

(J Med Internet Res 2019;21(9):e14017) doi:10.2196/14017

KEYWORDS

wearable technology; telemedicine; assisted living facilities; review

Introduction

Background

As the worldwide population grows and the access to health care is increasingly being demanded, real-time monitoring of various physiological signals has driven the research and development of diverse wearable and implantable systems. Connected health (CH) describes the new paradigm of a technology-enabled model of health and lifestyle management [1]. It is implicitly a multidisciplinary technology domain set up to provide preventive and remote treatments. CH relies on a digital information structure based on the internet, sensing, communications, and intelligent techniques, in support of health-related applications, systems, and engineering.

Wearables, as well as hearables (in-ear devices) and nearables (neighboring devices that interact with wearables) integrated into the wider concept of Internet of Things (IoT), are being considered the most likely technologies to transform future health care and lifestyles [2,3]. This revolution began with the smartphone, which is now becoming a widespread intrusive and ubiquitous technology. Most current wearables and nearables are equipped with different types of sophisticated sensors. Different types of sensors powered by advanced analytics are being explored to develop functionalities of truly portable medical laboratories. Seamless integration of these measurements in smartphone apps permits for targeted information to be delivered on time, enhancing the user experience in typical assisted living scenarios. The general acceptance, ease of use, and reliability of smartphones facilitates user adherence to different added value apps that allow filling a gap in the area of self-physiological sensing and fitness monitoring [4]. Wearable technology has become mainstream, with the most significant influence on fitness and health care industries [2].

The importance gained by wearables among consumer devices can be tracked by their increasing share in consumer electronics shows promoting self-care and health management. According to the International Data Corporation, 172.2 million wearable units were shipped in 2018 [4], and this number is expected to grow, contributing significantly to the revolution of the IoT market [5]. Advances in wearable technologies and user acceptance of available consumer wearable devices pave the pathway toward seamless physiological monitoring.

The first body area networks (BANs) and wearable units comprised a number of sensors with a processing unit and wireless nodes assembled on printed circuit boards [6,7]. The design was bulky and uncomfortable, accompanied by large batteries, and had numerous issues associated with frequent recharge and loss of data communication. Since then, tremendous progress has been made in sensing technologies. The bulky design is being rescaled to a system on chip. Lowered power consumption, reliable communications, distributed processing, and data analytics improved the potential of wearables and made a significant impact on technology acceptance [7]. The technology innovations directly responded to user-related concerns (sensor miniaturization, seamless monitoring, secured communications, lower power consumption, energy harvesting, and plug-and-play functionalities) as well as safety and security (reliable sensing and data preprocessing, secured data communication, and reliable analytics).

However, the user feedback reviews report that initial user enthusiasm on wearables is often lost because of unclear use cases (unclear end user need), price, and associated complexities in device pairing with a smartphone [8]. The translation to long-term commitment to wearables requires clear use scenarios, valuable feedback, and constructive recommendations [8,9]. The inevitable transformation from a traditional, reactive health care model to a proactive and preventive model will bring clear use cases of CH solutions for early diagnostic or chronic condition monitoring [1]. Innovative CH scenarios are strongly motivated, exact, and economically beneficial [3,10].

The role that sensing, and information and communication technologies have gained as essentials in digital health has been summarized and elaborated in numerous research articles on sensors, data analytics, and secure and reliable communication platforms for CH solutions [3,10-16]. To stimulate and facilitate knowledge transfer and dissemination among policymakers and stakeholders, it is equally important to summarize those original findings with respect to specific application scenarios and specific user groups. Systematic review studies deliver such overviews based on an exhaustive manual screening of available digital libraries, providing a qualitative analysis of included studies, and unbiased performance comparison of the corresponding CH solutions [17,18]. The examples of such review studies offering a useful insight into the spectra of the related wearable technologies, target user groups, and application domains are plentiful. Wilde et al [19] reviewed the usage of apps or wearables for monitoring physical activity and sedentary behavior and emphasized the barriers and facilitators for their adoption. A scoping review [20] summarized the practices and recommendations for designing, implementing, and evaluating mobile health (mHealth) technologies to support the management of chronic conditions of older adults, considering articles published from 2005 till 2015. Kvedar et al [10] focused on the concept of CH as an overarching structure for telemedicine and telehealth and provided examples of its value to professionals and patients. In the study by Liu et al [21], materials, design strategies, and powering systems applied in soft electronics were reviewed. It also summarizes the application of these devices in cardiology, dermatology, electrophysiology, and sweat diagnostics and discusses the possibilities for replacement of the corresponding traditional clinical tools.

The transformation of the wearable landscape in the last decade is thus evidenced in a plethora of original articles and patent documentation and summarized and compared in numerous focused systematic reviews [3,10-16,19-21]. In this paper, we scoped the wearable technology field over the decade, starting from 2010 to February 2019, to identify trends in literature with respect to 4 important pillars: technology, safety and security, prescriptive insight, and user concerns. The collected literature reflects on the achieved progress, open issues, perspectives, and gaps in the development of wearable systems for future CH domain. The covered topics mainly relate to enabling technology: sensing, data aggregation and processing, communication protocols, power supply, data protection, and data analytics. However, the results of numerous pilots and experience gained with consumer wearables provide an insight into different user-related concerns. After exploring the literature published over the last decade, we have summarized state-of-the-art technologies, future research focus, and paper statistics related to the following key issues: enabling technology topics, application of wearable sensors in CH, and different user concerns.

With the more general, high-level perspective on the research on wearable technology, user-related concerns and challenges experienced over broad application area, this scoping review aimed at overlooking research trends unconstrained to a particular user group, health condition, or lifestyle scenario and including both mHealth and smart living environments. The extensive search scope is supported by automated search procedures relying on natural language processing (NLP) algorithms. The trends over the last decade were analyzed using a set of identified articles from 3 large digital libraries.

Purpose of This Review

Many studies elaborating on the use of sensors and wearables in assisted living environments, CH, and wellness and fitness apps were published in the last decade [3,10-16,19-22]. Those studies provide significant input for designing future CH systems, indicating benefits, but also shortcomings, barriers, and user feedback [19,23-29]. Nevertheless, there is a lack of studies with a general overview of the nature and extent of published research in that context.

This study aimed to identify and scope the scientific literature related to wearables in health monitoring, as measured by trends in the research evidence available in 3 large digital libraries: Institute of Electrical and Electronics Engineers (IEEE), PubMed, and Springer. The study scoped the field from several perspectives aiming to capture key drivers and major constraints in the deployment of wearable technology for health. The enabling technology relies on advances in sensing, processing, communications, and data protection. Conversely, multiple user perspectives imply privacy, utility, complexity, price, relevance, reliability, and significance of delivered feedback.

The objective of this study was to scope the research on wearable technology for health with regard to the following research questions:

- What are the most significant research trends and milestones on wearables seen as an enabling technology and as a key driver facilitating CH solutions?
- What are the most critical identified barriers and concerns from the technology and user perspectives and what trends are reflected in the research literature relating to these issues?

As an added value, this review can help identify the topics that need more detailed research in terms of elaboration of the obstacles and potential breakthroughs. The list of relevant articles resulting from this study can be filtered with respect to different fields (eg, keywords) to identify articles of interest for a systematic review in a specific subfield. The details in the list facilitate fast manual screening and selection of the subset of articles for further qualitative analysis. This type of preliminary search in planning a systematic review provides valuable answers on the feasibility (ie, does any evidence in literature exist), relevance (ie, has a similar systematic review already been done), and amount of time needed (ie, volume of the found evidence) to conduct a systematic review.

Methods

Scoping Review Methodology

This study adopted a scoping review methodology to identify and process the literature on wearables published from January 2010 to February 2019. Using a scoping technique, we aimed to examine the research evidence in the broad field of wearables, analyzing technology trends, including the resolved and emerging issues. The lack of a qualitative analysis of identified papers, the broad topic range, and the number of studies involved defined our approach as a scoping review and differentiated it from a systematic review [30,31]. The purpose of this study fully complies with the aims of a scoping review "to search, select and synthesize the knowledge addressing an exploratory question to map key concepts, types of evidence, and gaps in research," as defined by Colquhoun et al [32]. Systematic reviews in the field of wearables, for its breadth and depth, have to focus more narrowly on wearable solutions and user concerns in a prespecified application scenario to facilitate qualitative analysis of included studies.

All emerging review types share their basis in scientific methodology, that is, they rely on formal and explicit methods for search and assessment of published studies and synthesizing of research evidence in conclusions on a well-defined research question [17]. One of the protocols for systematic reviews in health care, the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) [18], provides a good example of thorough and rigorous checklist guidance. The corresponding PRISMA flow diagram illustrates the information flow reflecting the number of studies in different systematic review stages: study collection, study scanning, eligibility evaluation, thorough qualitative synthesis, and quantitative synthesis in meta-analysis [18]. The methodological framework for scoping reviews is underpinned by this exact and transparent way systematic reviews are conducted [17], providing sufficient details to reproduce the results. The workflow for a scoping review proposed by Arksey and O'Malley [30], and adopted in this study, includes 5 stages:

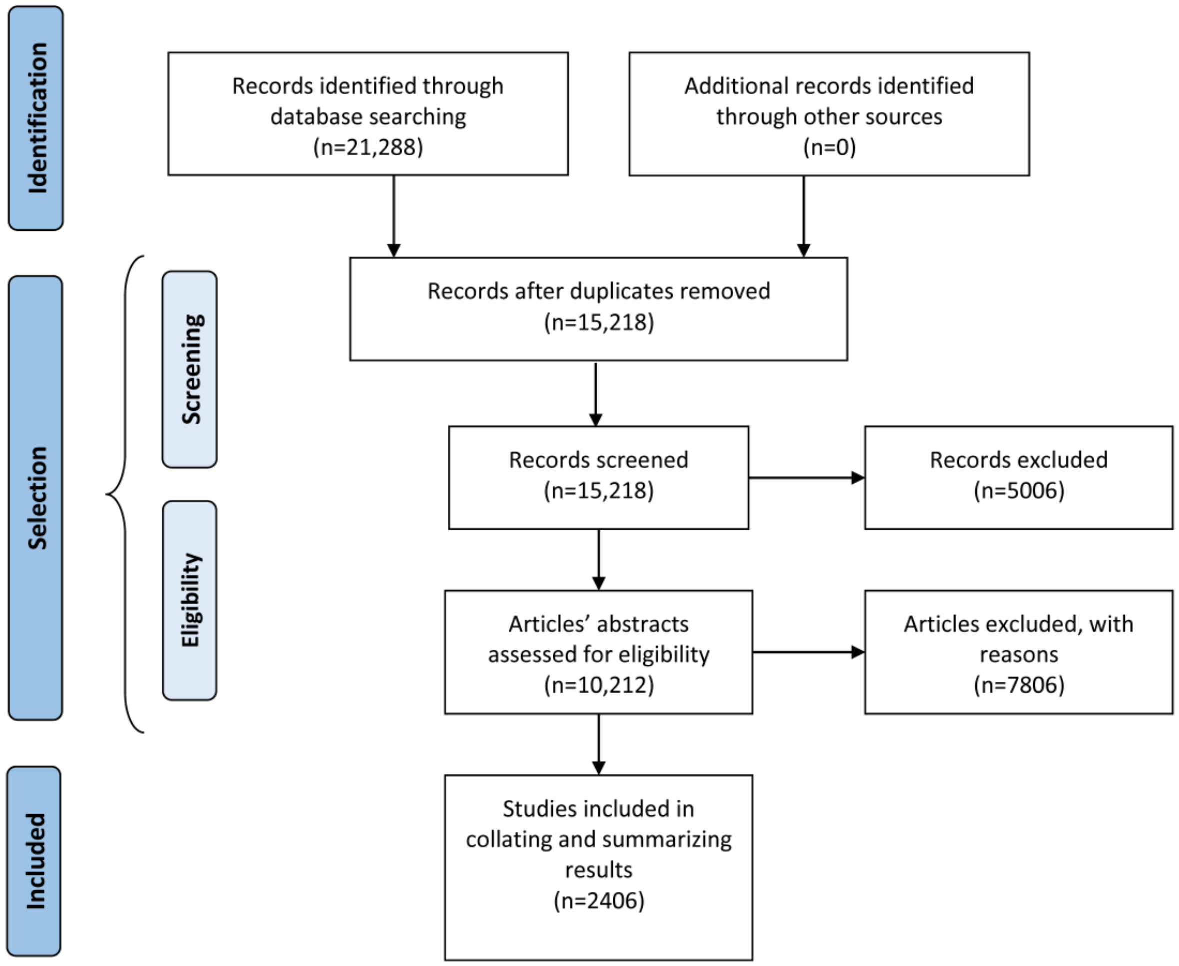
- Identification of a research question;
- Identification of relevant studies;
- Study selection;
- Charting the data; and
- Collating, summarizing, and reporting the results.

The identification of relevant studies and study selection stages in the scoping review methodology corresponds to the PRISMA workflow phases: study collection, scanning, and eligibility evaluation. To ensure transparency, we have enclosed the workflow chart to illustrate the number of identified, scanned, and included articles in this scoping review (Figure 1).

The scope of this study was substantial and the collected research evidence on wearables surpassed the potentials of a manual search. Relying on the advances in NLP algorithms, the NLP tool kit [33] was used to ensure an efficient and exhaustive search of the literature corpus. The NLP tool kit is designed to automate the literature search, scanning, and eligibility assessment in the PRISMA methodological framework for systematic reviews [18], which are aligned with the scoping review phases: identification of relevant studies and study selection.

In the following sections, we clarify the usage of the NLP tool kit for study identification, selection (ie, scanning procedures and eligibility criteria assessment), and charting the data. It is worth noting that no quality assessment of the selected articles has been conducted, as this review has a scoping character. Instead, in the final step, we collated, summarized, and reported the results by aggregating the included studies to address the objectives of this review.

Figure 1. The review workflow reflecting the number of articles identified, screened, processed and removed in each step. The remaining articles were used for aggregating and summarizing the results.



Setting Up the Natural Language Processing Tool Kit

This stage concerns the development of a plan comprising decisions on which digital libraries will be queried, relevant time span, suitable keywords, and properties that should be satisfied. This scoping review has employed the NLP tool kit [33] enabling both automated search, scanning, and processing procedures. The NLP tool kit ensures compliance with the terms of use of the digital libraries, with regard to the number of requests per unit time.

The NLP tool kit input parameters are a collection of keywords that are used to identify potentially relevant articles and a set of properties that should be satisfied by the identified articles. The input is further expanded by proposing synonyms to the search keywords and properties. Synonyms can be provided by the user or proposed by the tool kit and fine-tuned if needed.

Keywords are search terms or phrases that are used to query a digital library (eg, “health” and “ambient and assisted living,” “health” and “enhanced living environment”). Eventual duplicates in the results are removed in a later phase. *Properties* are words or phrases that are being searched in the title, abstract or keywords section of the articles identified with the keywords. Examples of such properties employed in this study are “monitoring,” “recommendation,” and “detection.” *Property groups* are thematically, semantically, or otherwise grouped properties for a more comprehensive presentation of results. For example, the property group for the set of properties given in the example above can be “information delivery.” Table 1 summarizes the relevant input categories used in this work.

Start year indicates the starting year of publishing (inclusive) for the papers to be included in the study. *End year* is the last year of publishing (inclusive) to be considered in the study. This review encompasses studies published from January 2010 to February 2019. *Minimum of the relevant properties* is a number denoting the minimum number of properties that an article has to contain to be considered as relevant. In this study, this value was set to 3, providing a right balance between falsely identifying relevant papers and potentially missing a relevant paper.

When researches perform a scoping review according to the above-mentioned methodology, the actual tasks they perform involve searching digital libraries with different search phrases, often involving complex Boolean conditions. The NLP tool kit counterpart to these phrases are the keywords described above. By screening the title and abstract, a reviewer determines whether the article is indeed relevant for the study. In the NLP tool kit, this process is automated using the properties and their synonyms to define what we are looking for in an article. Articles that contain more properties are considered as more relevant. Undoubtedly, a human reader might better understand the context and better assess the relevance of an article. However, the NLP tool kit is mimicking these tasks, but in an automated and more thorough way, providing incredible efficiency of the scoping review process. For more information about the actual implementation, we refer the reader to the study by Zdravevski et al [33].

Table 1. The natural language processing tool kit input parameters: keywords, properties, and property groups.

Input parameters	Natural language processing tool kit input parameters
Keywords	• health AND any of (ambient assisted living, ambient intelligence, assistive engineering, assistive technologies, enhanced life environment, enhanced living environment, enhanced support environment, hearables, home technologies, nearables, smart environment, smart home, wearables)
Property groups (properties)	• Technology (cloud, communication, communicating, transmission, data processing, data analytics, battery, energy, fog, edge, protocol, sensor, sensing, detector) • Information delivery (monitoring, recommendation, detection, supervise, catching, detecting, spotting) • Concerns (acceptance, adoption, privacy, concealment, intrusiveness, intrusive, technology acceptance, technology adoption, seclusion, meddlesomeness) • Safety and security (protection, reliability, dependability, safety, safe, security)

Identification of Relevant Studies

Upon provision of the defined input categories, the literature search was started using only the specified keywords to query the selected digital libraries. The NLP tool kit indexed the following digital libraries (ie, sources): IEEE Xplore, Springer, and PubMed. It is worth noting that the NLP tool kit has used search engines of the corresponding publishers and retrieved the search results. Depending on the digital library in each search, the number of retrieved articles was constrained. In the PubMed library, all articles matching the given search criteria were retrieved for further analysis. The IEEE’s search engine limits the number of articles in each search to 2000, all of which were retrieved. For Springer, the search for each keyword separately is limited to 1000 articles or 50 pages with results, whichever comes first, sorted by relevance determined by Springer.

Study Selection

After articles had been identified based on the specified keywords and retrieved from the publishers, the study selection (screening and eligibility assessment) procedures followed.

Upon merging the results from multiple independent keyword-based searches, some articles could be present multiple times because they could be identified by different keywords or in multiple libraries. Therefore, the collected articles were screened, and duplicates were removed using their digital object identifier (DOI). In addition, the screening process discarded articles that were not published in the required time span (ie, last 10 years) or for which the title or abstract could not be analyzed because of parsing errors, unavailability, or other reasons.

The selection of studies from the remaining subset of articles relied on the advanced functionalities enabled by NLP tools. The NLP tool kit automates analysis of a title and abstract for each study, significantly reducing the number of articles for manual screening. The automated eligibility analysis involved the following processing: tokenization of sentences [34,35] and English stop words removal, stemming, and lemmatization [35] using the Natural Language Tool kit library [36]. Stemmed and lemmatized properties were searched in the cleaned abstracts and titles, and each article was tagged with the properties it contained.

The processed articles were selected (ie, labeled as relevant) if they contained at least 3 of the predefined properties in its title or abstract (considering the above NLP-enhanced searching capabilities, thus performing a rough screening). To help in the eligibility analysis, the selected articles were sorted by the number of identified property groups, number of identified properties, number of citations (if available), and year of publication, all in descending order. For the relevant articles, the tool kit automatically generated a bibliographic file (as defined by BibTeX reference management software) for simplified citations.

The information flow diagram illustrating the numbers of identified, screened, processed, and removed studies in the automated NLP procedure is presented in the Results section (Figure 1) to ensure transparency and reproducibility.

The listing of the relevant identified articles extracted from IEEE, PubMed, and Springer is available in Multimedia Appendix 1 as an Excel file with the following fields: DOI, link, title, authors, publication date, publication year, number of citations, abstract, keyword, source, publication title, affiliations, number of different affiliations, countries, number of different countries, number of authors, BibTeX cite key, number of found property groups, and number of found properties. These additional files facilitate refined manual search of the articles with specific filtering criteria. The subset of targeted articles can subsequently be retrieved from their publisher and manually analyzed for potential inclusion in the qualitative and quantitative synthesis. It should be noted that not all the references provided within this study are from the identified set of relevant papers. Some additional papers identified in a manual search were used to illustrate and confirm the findings of this scoping review. However, these referenced papers from other libraries have not been used to identify trends in this scoping review.

To replicate the results manually, the keywords in Table 1 have to be used to inquire the selected digital libraries using their search engines. The properties serve for identification of the relevant articles by scanning titles and abstracts of the identified studies. The results can be compared with the resulting list of included studies, provided in Multimedia Appendix 1.

Charting the Data

To answer the research questions, we defined indicators to be extracted from the selected studies. The trends in the past decade were analyzed relying on a broad scope of literature. The processed and retained relevant articles were aggregated by several criteria:

- Source (digital library) and relevance selection criteria;
- Publication year;
- Digital library and publication year;
- Search keyword and digital library;
- Search keyword and year;
- Property group and year;
- Property and year, generating separate charts for each property group; and
- Number of countries, number of distinct affiliations and authors, aiming to simplify the identification of collaboration patterns (eg, written by multiple authors with different affiliations).

These aggregated metrics are available in the form of comma-separated values files and charts. The plotting of the aggregate results was integrated and streamlined using the Matplotlib library [37] and NetworkX [38]. The NLP tool kit enables graphical visualization of the results, where each node represents one of the properties, each edge connects 2 different properties, and its weight is determined by the number of articles containing both properties connected by that edge. Articles that do not contain at least 2 properties, and properties that were not present in at least 2 articles were excluded. For a clearer visualization, only the top 25% property pairs by the number of occurrences are shown in a graph.

A similar graph for the countries of affiliations was generated. The top 50 countries by the number of collaborations were considered for this graph. Countries and an edge between them are shown if the number of bilateral or multilateral collaborations was in the top 10% (above 90th percentile) within those 50 countries.

Collating, Summarizing, and Reporting Results

Using charted data and extracted evidence, we were able to analyze the trends in data and provide qualitative analysis for each thematic segment (as defined by the property groups). The results were reported with regard to the raised research questions. The meaning of these findings was related to the study purpose, and the potential impact on the future research direction was discussed.

Results

Number and Distribution of Identified Articles

Using the NLP tool kit and searching 3 digital libraries: PubMed, IEEE, and Springer, we identified 21,288 studies with potential relevance (Figure 1). Duplicates that emerged in multiple independent searches were removed, reducing the total number to 15,218 studies. The first screening process further eliminated 5006 studies published before 2010 or for which the title or abstract could not be analyzed because of parsing errors, unavailability, or any other reason. The remaining 10,212 studies underwent an automated eligibility assessment using the advanced NLP tool kit functionalities. After processing, the articles were tagged with identified properties, and all articles containing less than 3 properties were removed. Overall, 2406 articles were deemed eligible for further manual inspection and inclusion in identifying the research trends and summarizing the results. The statistics on the number of the collected articles, duplicates, articles with invalid time span or the articles with incomplete data, and relevant articles are presented in Figure 2 for each digital library.

The distribution of the number of collected and relevant articles per year is presented in Figure 3. An increasing trend in the number of collected articles can be noticed from January 2010 to February 2019. The same trend is followed by the number of included articles, which rises from 136 in 2010 to 393 in 2018.

Combining the information on the digital library (source) and publication year of the identified relevant articles, the obtained distribution reveals that IEEE, being a more technology-oriented library, has an increasing trend in the number of relevant articles from 2010, peaking in 2017 (Figure 4). PubMed leads in the number of articles dealing with CH and assisted living and covers more of the searched properties related to user concerns. The number of PubMed articles follows an increasing trend from 2010 and saturates in research evidence from 2016 onward. The Springer library shows an oscillating trend from 2010 to February 2019, with an average of around 50 articles per year.

Figure 2. Statistics on the number of articles in the identification and study selection (screening and eligibility assessment) phase for each digital library. IEEE: Institute of Electrical and Electronics Engineers.

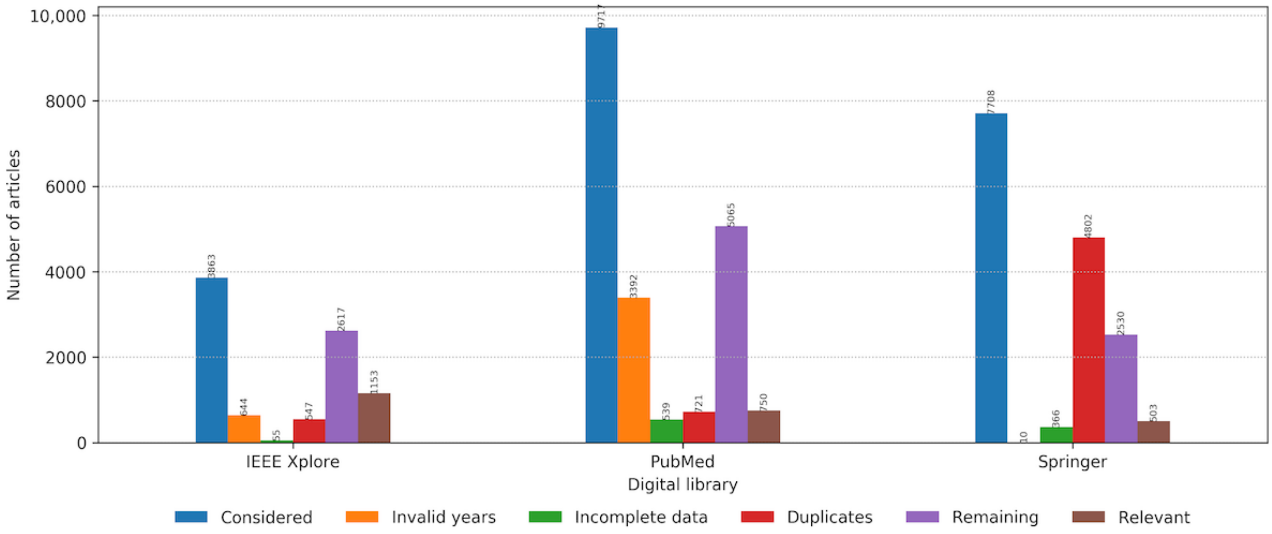


Figure 3. The number of analyzed articles versus the number of selected relevant articles on wearable technologies per year from January 2010 to February 2019.

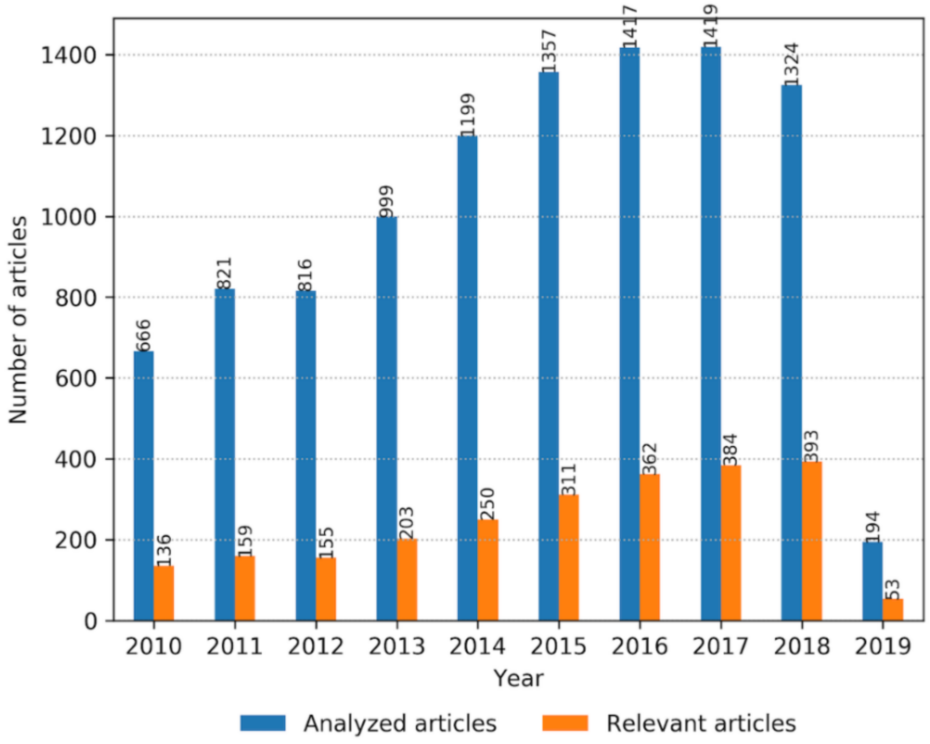
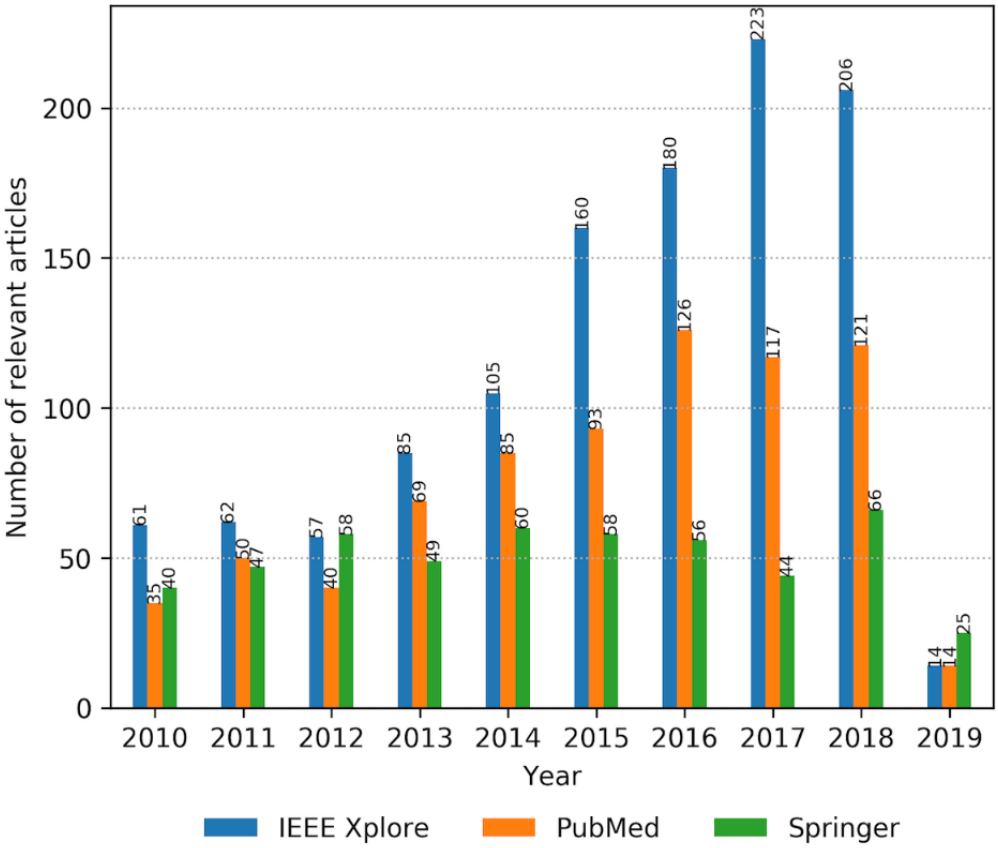


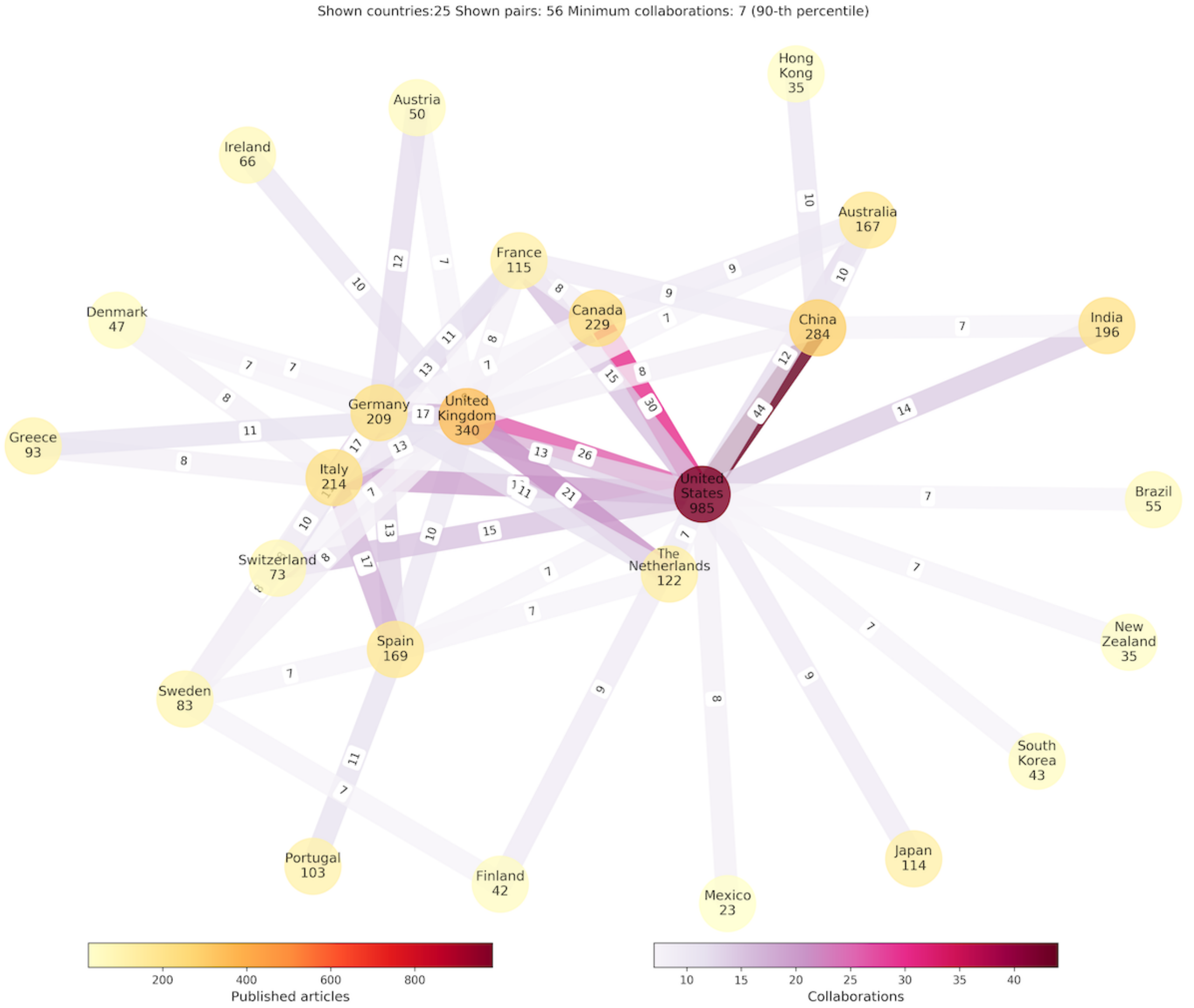
Figure 4. The number of identified relevant articles per year from January 2010 to February 2019, sorted by the originated digital library (Institute of Electrical and Electronics Engineers [IEEE], PubMed, and Springer).



Geographical Distribution and Collaboration Evidence

The authors' affiliations were used to identify wearables' research community clusters and eventual hubs at the research forefront. Multiple country associations were discovered, but for the sake of presentation clarity, the graph in Figure 5 shows 25 countries (nodes) and 56 edges with at least 7 joint articles (90th percentile) specified as edge weights. The number of papers per presented node is color coded, where violet corresponds to the higher and yellowish (paler) color to the lower number of articles. The identified hubs, United States, Canada, United Kingdom, Germany, China, and Italy, feature both national and international scientific production, whereas the strongest edges exist between the United States and Canada and between the United States and China. The collaboration patterns largely correspond to the neighboring geographical areas. The European countries demonstrate active collaboration scheme as well. The United States, United Kingdom, and China have significant national scientific production in the analyzed research domains.

Figure 5. The number of relevant research papers per country and collaboration links with the annotated number of joint articles. For clarity, only 25 countries (nodes) with significant research contribution and minimum 7 joint collaborations (edges) are presented.



Keywords Statistics

The selected keywords used to map the literature corpus on wearables with respect to the set research questions appear in the relevant articles with different distributions. Figure 6 presents the annual number of research papers identified by the search engines of 3 libraries with the defined keywords and additionally filtered manually based on their relevance to the defined properties. Please note that the internals of their search engines are not known, and the libraries might differ in the way they look for these keywords: only in a title, keywords section, abstract, or a whole article. Depending on the digital library, the ratio of the relevant papers containing specific keywords changes (Figure 7). The IEEE digital library has a focus on enabling technology for CH, in terms of novelties in wearable sensing, data processing analytics, computing, and communication protocols. PubMed publications are also oriented toward CH technologies from an assistive and supportive perspective. Springer publications cover slightly different topics, focusing mainly on ambient assisted living (AAL) and ambient intelligence and generally contain more technical articles that address assistive technologies.

Figure 6. Distribution of the number of relevant articles with each of the defined keywords on an annual basis from January 2010 to February 2019.

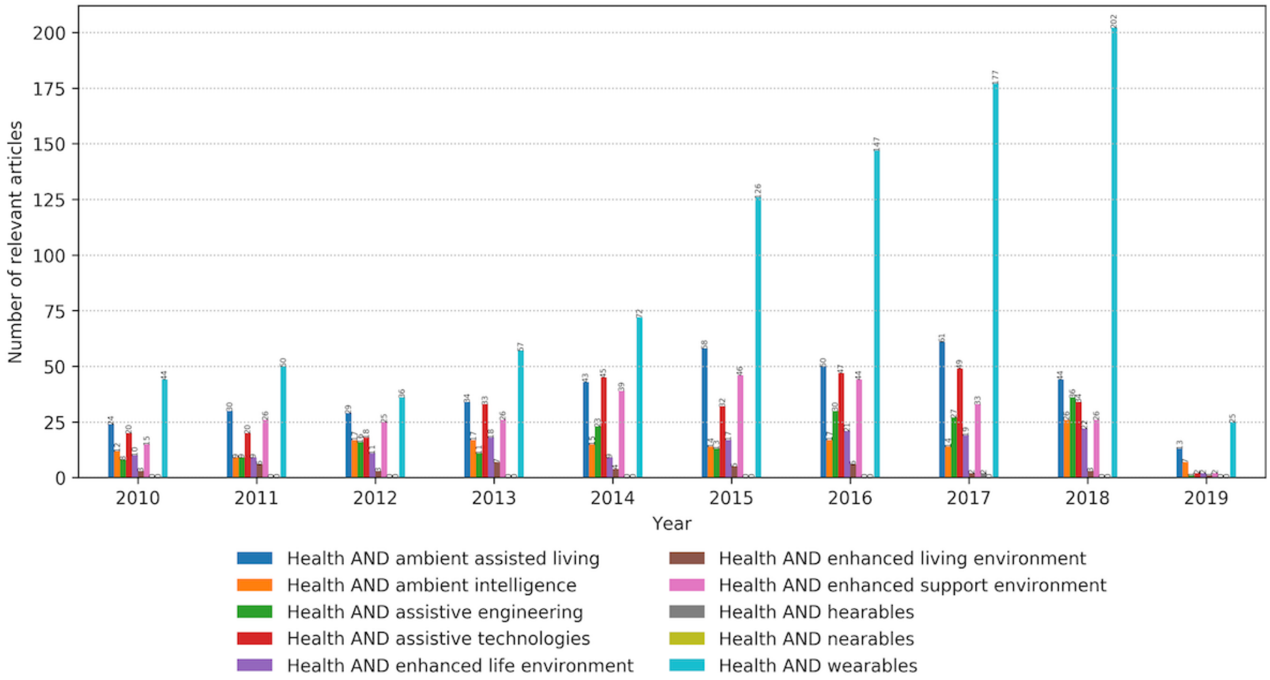
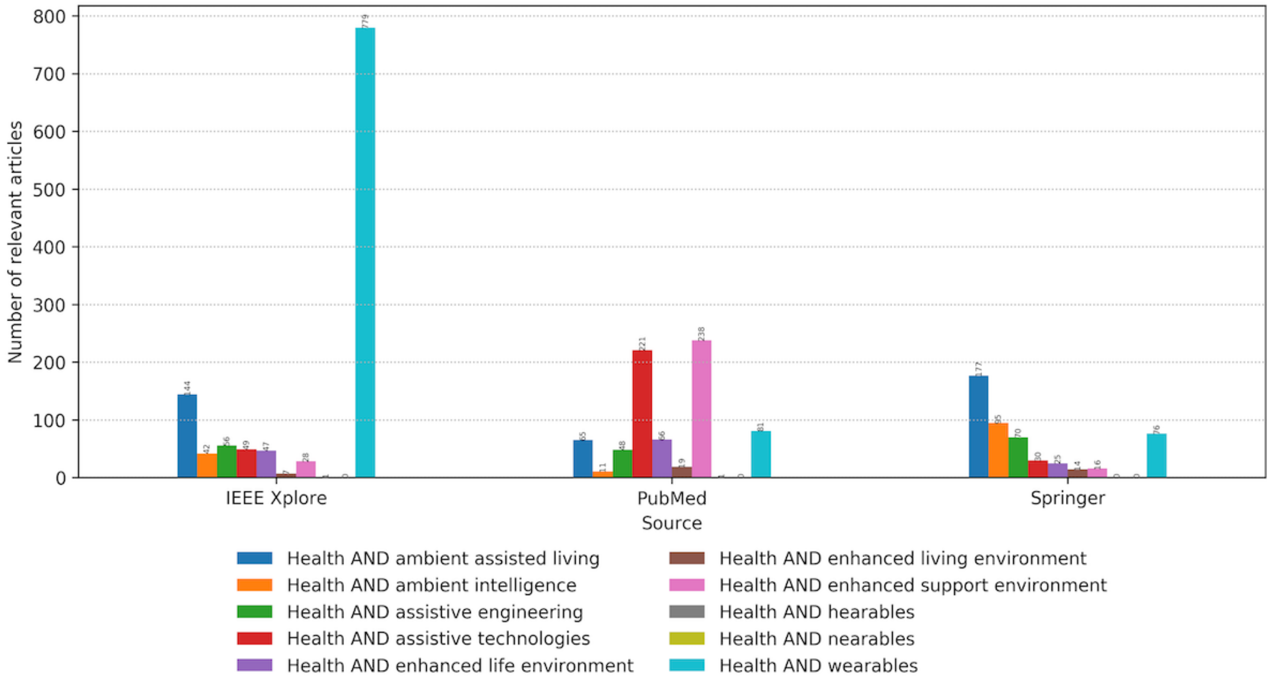


Figure 7. The number of relevant articles containing each one of the keywords per digital library. The data are aggregated within the defined period. IEEE: Institute of Electrical and Electronics Engineers.



Statistics of Properties

As the number of research articles increases within the observed time frame, the number of articles dealing with associated topics summarized in property groups increases accordingly (Figure 8). The increasing trend is accompanied by the stable ratio of papers, with technology-related publications being the leading in number, followed by research related to information delivery, safety and security, and user concerns. When the view is zoomed from property groups to properties, the graph reveals the centrality of *monitoring* as the essential function of a wearable system tightly connected with the key technology: *sensing* (Figure 9). The 2 properties interrelate with communication, detection, reliability, safety, security, transmission, data analytics, and privacy as technologically empowered concepts. Acceptance is the key user-related property in the graph core, with privacy and protection to follow.

Figure 8. The number of relevant articles related to each property group per year within the predefined time frame from January 2010 to February 2019.

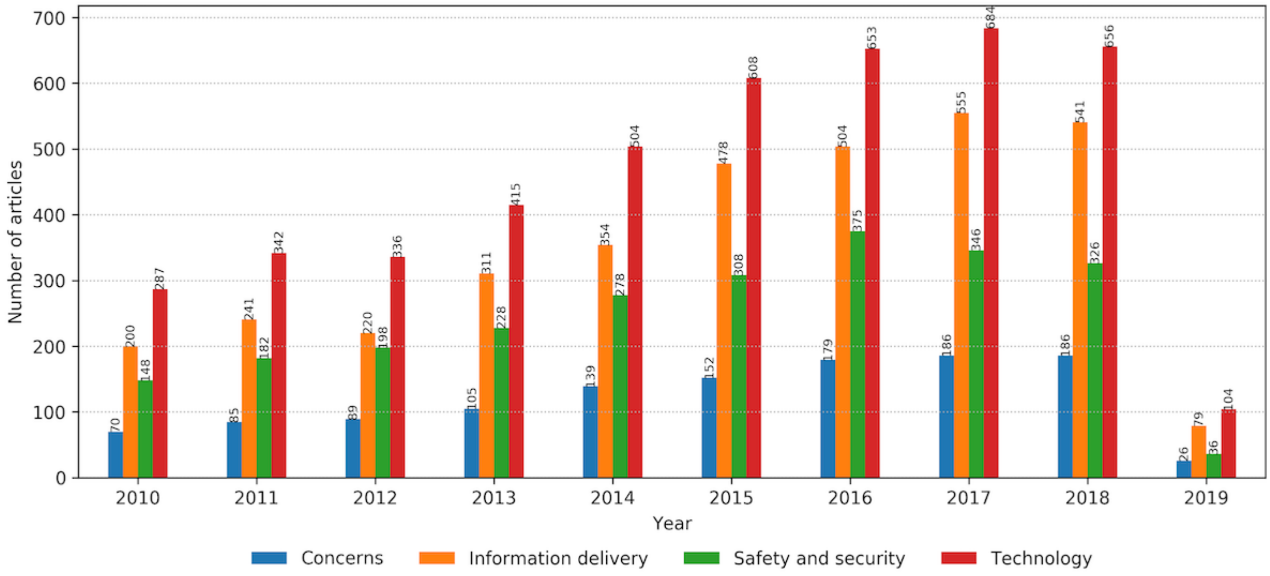
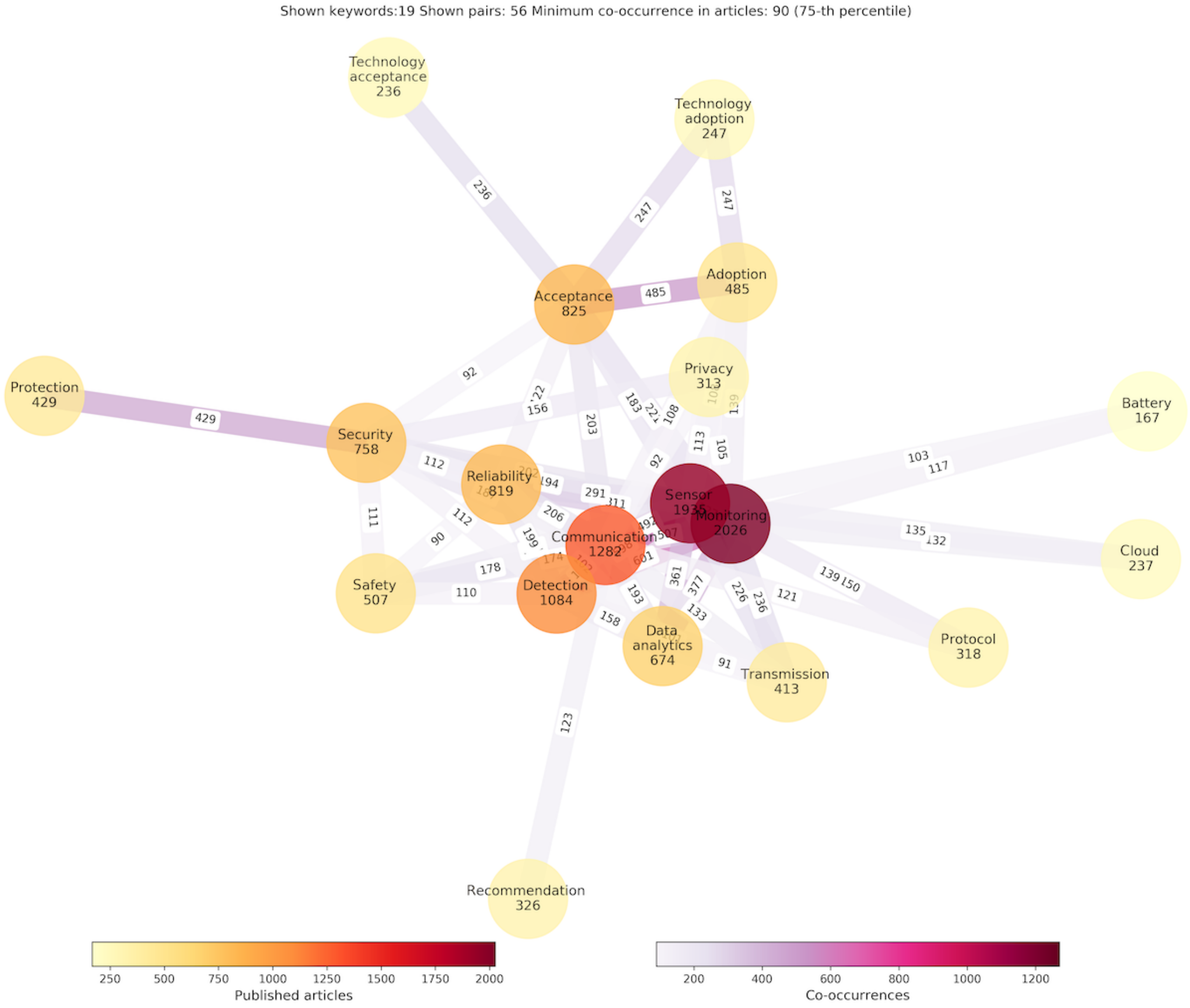


Figure 9. The property graph indicates the number of relevant articles with each property and the co-occurrences of properties in those.



Discussion

Principal Findings

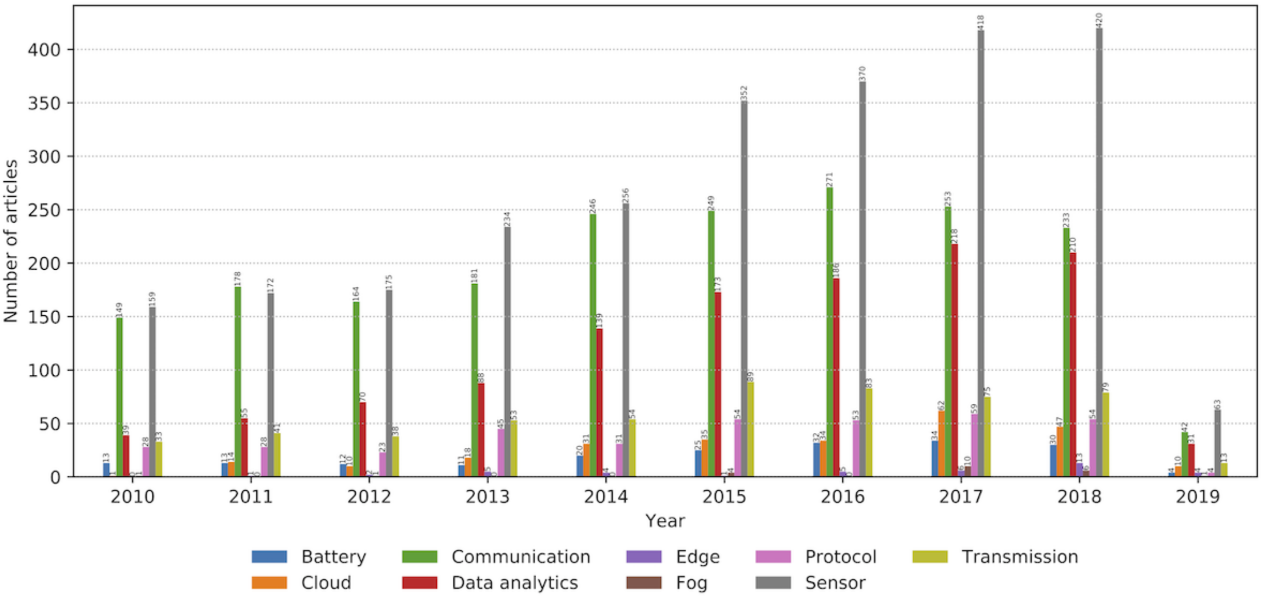
Wearable medical devices play a critical role as an enabling technology and as a key driver that has facilitated the emergence of CH solutions. This paper presents an overview of the most important milestones and trends that have driven research and development initiatives on wearable technology domains in the last decade. Simultaneously, it aimed to identify the most critical barriers or concerns, as far as technology and user aspects are concerned, that hinder the generalized adoption of wearables and still require further research.

The adopted methodology used the NLP tool kit for searching in 3 digital libraries, PubMed, IEEE, and Springer, for papers that address research on wearable technologies for medical applications. In the following, we address the findings related to the research trends in technology, information delivery, user concerns, safety, and security.

Technology as a Key Driver

The literature (Figure 10) reflects the intense research and development in sensor design, communication protocols, and data processing and analytics. The emergence and evolution of concepts of edge computing, cloud, and fog could be easily tracked. As technology is a key enabler of future CH systems, we briefly review significant technological advances in the comprising components of a wearable system.

Figure 10. The number of relevant articles containing each property grouped primarily into the technological domain in the period from January 2010 to February 2019.



Evolution of Sensing Technology

Available sensors and their characteristics largely influence the design of CH systems. The direct sensor's contact with the body implies their stiffness and size, as the most important features concerning comfort and measurement accuracy. The placement of wearable sensors influences their characteristics, user acceptance, and engineering requirements. As sensors evolve from wearable and implantable to ingestible sensors, barriers arise on multiple pathways: regulatory, technical, and translational [39].

The marked progress in wearable sensors is linked to advances in material science and embedded systems. Smart garments or electronic textiles, featuring sensor flexibility, made the first promise toward seamless and pervasive monitoring. The sensor integration into fabrics varies from *garment level*, assuming sensor integration at a later stage, to *fabric level* implying sensor integration by application of coatings to the fabrics [40]. The striving level is a *fiber level* [40] implying integration of conductive threads and fibers in the knitting process to result in a *smart fabric* (a concept first proposed about 20 years ago [41]).

Microcontroller-based systems can as well be included within different textile fabric for health applications [42]. Some products have already been approved and introduced to the market, but most of them are at a prototyping stage. The limitations arise at the electronic and textile integration step, slowing down technology transfer. In addition, there are multiple regulatory concerns, such as safety, reliability, and recycling [43]. Another promising technology for wearable CH solutions is microfluidics. Both sensing and drug delivery can be realized by combining microfabrication and liquid manipulation techniques with conductive elements on stretchable and flexible materials [44,45].

Low-power microelectronics, biocompatible materials, micro- and nano-fabrication, advances in data transmission, and management of sensor drift have driven the development of implantable biosensors [46]. Recent advances report the use of polyamide, flexible material for sensor platforms [47,48]. Research on flexible mechanical and electrical sensing has demonstrated great potential in in vitro diagnostics [49] and advanced therapy delivery [50]. Polymer-based switching matrices used for electronic skin to enable pressure sensing (robots, displays, and prosthetics), evolved into skin-attachable wearable electronic devices [48]. Another use-case involves surgical procedures, where these matrices are used in surgical procedures as part of mapping systems attached to the surface of the organs [50]. Active research directions in polymer sensors are focused on transparency [51], self-powering [52], and self-healing [53] capabilities.

The new generation of implantable sensing solutions for tissue and organ monitoring is enabled by advances in epidermal electronics based on soft lithography and thin-film sensors [46,54]. For example, electrocardiogram, blood glucose, and blood pressure sensors integrated with microstructures provide optical, thermal, and electrical stimulation [55].

Hearables are one of the latest wearable devices aiming to integrate sensing of multiple physiological signals into a single device [56]. The *in-ear* placement of such a device requires a flexible and comfortable fit and provides stable position regardless of the subject’s gross movements. The viscoelastic foam used as a substrate additionally ensures artefacts absorption, as the ear channel is affected by small movements, when speaking, swallowing, or chewing. The solution proposed by Goverdovsky et al [56] offers continuous measurements of cardiac, brain, and respiratory functions.

Implantable pacemakers, pressure sensors, cochlear implants, drug infusion pumps, and stimulators are all examples of implantable devices delivering therapy or providing physiological monitoring [39]. The majority of implantable devices currently operate in an open loop. New research challenges are focused on combining monitoring and therapy delivery for the optimized closed-loop personalized therapy [39]. The neural signal recording is ultimately the most demanding task, as it requires precise, low-power, and low-noise electronics and miniaturized and light weight implantable designs [57]. Neural implants face the hardest challenges in the translational pathway of the research-grade solutions into clinically approved products.

Ingestible sensors for image and data recording in gastrointestinal endoscopy have already proven their benefits in early detection of gastrointestinal cancers [58]. Ingestible, similarly to implantable devices, face challenges that shape the ongoing research: operation frequency selection, amplifiers, antenna design and performance, wireless channel modeling, increasing data rates, and power considerations.

Besides tracking basic physiological parameters (electrocardiogram, blood pressure, blood oxygen saturation, temperature, etc) sensing functions in wearable medical devices have also moved off the body toward contactless or seamless ambient embedded physiological sensing in, for example, keyboards, joysticks, steering wheels, bicycle handles, doors [59], mattresses [60], beds [61], and toilet seats [62]. The combination of such monitoring products with the data-driven services has promoted the development of the AAL concept. The AAL is a new ambient intelligence paradigm where new technologies are associated with the social environment, to transparently improve and assist the daily quality of peoples’ lives. Despite the high number of research and industry organizations already active in the AAL field, significant efforts are still needed to bring these technologies into a real-world usage [15].

Powering Wearables: Constraining Consumption and Energy Harvesting

One of the limitations for a widespread adherence to wearable electronic products concerns the power supply needs [7,9,63]. Active wearable systems need to be comfortable, light, user-friendly, and power efficient. The identified research trends reveal that research on battery technology lags compared with research on other wearable system components (Figure 10). This implies that energy efficacy and efficiency remain an important design concern, both for wearable systems and in the design of networks to serve future landscape of wearables (notably fifth generation [5G] architectures).

Energy harvesting technologies have been explored as an alternative energy source to recharge power batteries or super capacitors. The ongoing research in this domain has investigated technologies to explore motion [64,65], thermal [66,67], optical, electromagnetic [68], solar [69], and chemical forms of energy [70]. However, miniature devices that can harvest proper levels of energy are still in their infancy.

Complementary efforts are being invested in the integration of power-efficient technologies and design techniques in wearable systems. Among those are energy-efficient and low-power wireless communication, voltage scaling, low-leakage and low-voltage complementary metal oxide semiconductors [71], and power-performance management.

Communication Protocols for Wearable Systems

The medical data are low in volume, but with strict requirements in terms of latency, link reliability, and security [7]. Wearable body sensor networks or BANs refer to sensor networks applied for acquisition or monitoring of vital physiological body parameters unobtrusively. These systems can be used in clinical settings or at home by patients or even healthy people who want to improve or monitor their health conditions.

BANs enable wireless communication in and around a human body in 3 different tiers: intra-BAN, inter-BAN, and the beyond-BAN. Intra-BAN communications refer to communications between on-body sensors, within the surrounding body area, enabling wireless data transmission to a personal server. According to the application and design parameters, the intranetwork can be wired or wireless, or even use the human body as a communication medium. Wired networks, as a second type of communication infrastructure for BAN applications, provide high-speed, reliable, and low-power solutions [72].

The international IEEE 802.15.6 standard enables delivering of low power, short range (in the vicinity or inside, within the human body) reliable wireless communications, with data rates from 75.9 kbps to 15.6 Mbps, making use of industrial, scientific, and medical bands, as well as frequency bands approved by national medical and regulatory authorities [73].

The inter-BAN communications include communicating data from personal devices such as smartphones to the access points, either in an infrastructure-based manner or in an ad hoc manner. Wireless BANs can interact with other existing wireless technologies such as ZigBee, wireless local area networks (WLAN), Bluetooth, wireless personal area network, video surveillance systems, and cellular networks [73].

Finally, the beyond-BAN tier connects the access points to the internet and other networks. Beyond-BAN architectures can be implemented in cloud or fog network infrastructures [74] implying *protocols*, *cloud-based systems*, and *fog systems* as research topics in the wearable CH domain. The major challenges in BAN are associated with media (path loss because of the body absorption), physical layer (minimization of power consumption with uncompromised reliability and interference), medium-access control layer (supporting multiple BANs in parallel application), security, and transmission (loss and delay sensitive real-time transmission) [75].

Limited spectra and the need for higher data rates drive the communication community toward the new generation of cellular networks such as 5G [22,63]. The high-speed data and low-latency features of 5G networks will allow wearable devices to communicate faster (in less than 1 millisecond) and perform real-time control. 5G will be a platform for various services and applications, with support to different communication requirements. The transition to millimeter wave (mmW) frequencies will require new communication architectures to be designed for specific mmW propagation. For protection and regulation of exposures to such frequencies, more appropriate metrics are needed, such as temperature elevation of the contact area [7].

The design of wearable antennas, with safety concerns, device-centric architectures, and smart device communication are some of the changes 5G will require. The development of 5G brought the promises supporting the wearables market, such as radio-frequency sensor charging [63], reduction in latency, high data rates and capacity, and network densification, enabling the massive number of deployed wearables per micro- or picocell [22].

The 5G architectures proposed to serve wearables include microbase stations for blanket coverage, whereas local coverage and data throughput should be ensured with small base stations and remote radio headers (RRHs) [7]. RRHs can also support different wireless technologies to ensure backward compatibility (Bluetooth, visible light communication (VLC), etc). The connection to cloud data servers via base stations enables storage, retrieval, and analytics of user-specific data. Realization of communications between wearables and network edge nodes can be done using licensed or unlicensed communication bands. Licensed communication bands provide quality of service at an increased cost at several levels: a service provider cost for more expensive licensed chips and more power consumed on licensed communication protocols. Unlicensed communication (eg, Bluetooth, WLAN, and VLC) is a cheaper, power-preserving option but limited in range [7].

Data Processing and Analytics

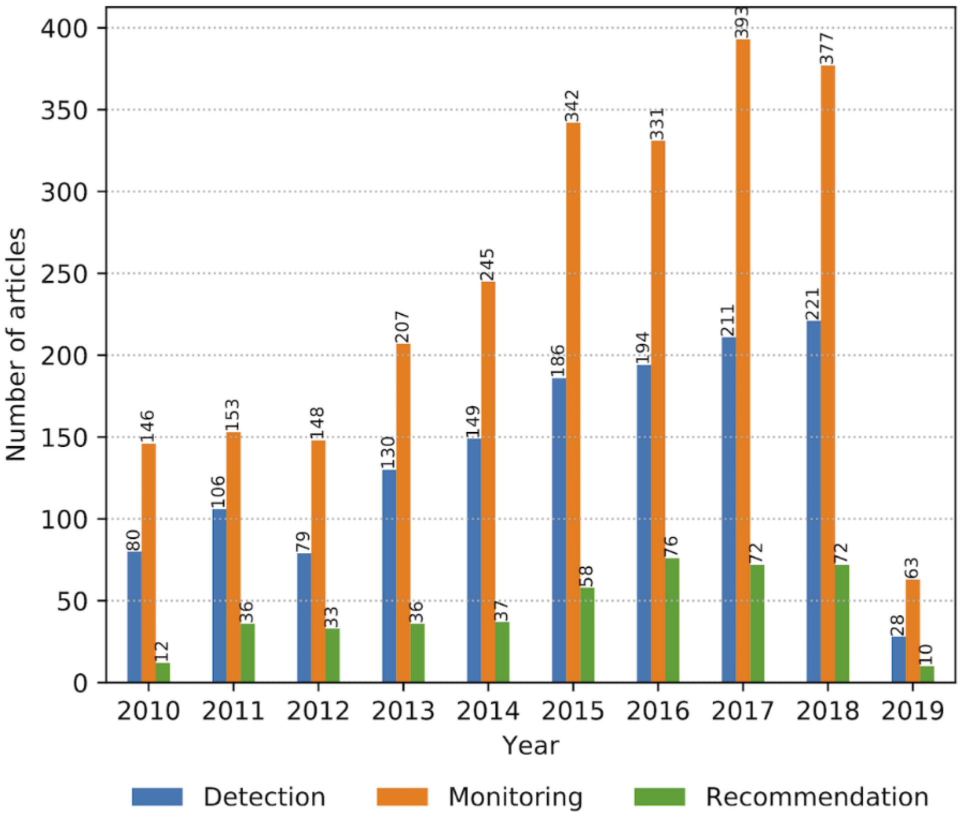
The large volume and heterogeneous data types collected using wearable technology have grown beyond the abilities of commonly used data processing techniques [76]. The necessity for reducing the volumes of captured data at the source, to reduce the power consumption and latency, brought processing closer to the sensor nodes, mapping the data algorithms to ultralow-power microcontrollers [46]. Preprocessing approaches, such as noise filters, peak detection, and feature extraction, allow for significant data reduction at the source [77]. Conversely, advanced data analytics imply sensor data integration, thus relying on the powerful devices located in the cloud. 5G should offer mobile edge computing to reduce latency and traffic demands to the central node. In the wearable scenario, communication between various user devices is fostered by 5G machine-to-machine communications, enabling local processing, low latency, and power saving [7].

High-performance computing permits efficient processing of large data volumes through a map-reduce framework [78]. Advanced data caching and in-memory processing coupled with GPU accelerators and coprocessors support intensive parallel operations. The availability of higher computational power enabled the rebirth of computationally intensive deep neural networks, resulting in superhuman performance and cutting-edge research in multiple domains. These are enabling technologies that will bring to reality the third generation of pervasive sensing platforms [46] that will integrate and extract information from a variety of sources: sensed data, clinical records, genomics, proteomics, and social networks, leading to a system-level approach to human health [79].

Information Delivery and Valuable Feedback

The research in the user-associated information delivery is primarily concerned with recommendations, provision of feedback, and real-time user insight (Figure 11). Current commercial wearable technologies, tracking vital signs and patterns of activity, lack the relevance for many potential consumers, presenting an additional burden [7]. The motivation to buy and use wearable systems has to be justified in a functional CH application context. The clear user benefit comes from a validated system that would transform collected data into manageable and useful information for medical action, safety instructions, or self-performance estimation and improvement.

Figure 11. The number of relevant articles related to recommendations and feedback on health monitoring solutions from January 2010 to February 2019.



To gain wider consumer preference, the information generated by wearables has to be fitted into specific contexts, offering the needed insight and recommendation on actions that should be taken. The second generation of wearable systems, which aims to enable context sensing, needs to integrate many different types of context information, such as sensor information, user profiles and preferences, activity patterns, medical history, and spatial information (location and environment conditions). If not strictly depending on medical condition, the timing, content, and frequency of prompting have to be adjusted to user preferences [80]. As a basic example, the time of day or night implies different content and presentation of the prompting messages because of the different level of user's readiness and wakefulness [81]. The fusion of physiological and context sensing data will rely on sophisticated data analytics for extraction of relevant information and decision making on an action to be prescribed or advised to the user. The feedback to prompting messages generated in day-to-day system's interactions with a user would ensure the adaptation to user preferences in time, relying on reinforcement learning.

The transformation of wearables from measurement devices into resources of reliable real-time information, history mining, and smart and personalized decisions would qualify them for health and performance monitoring solutions.

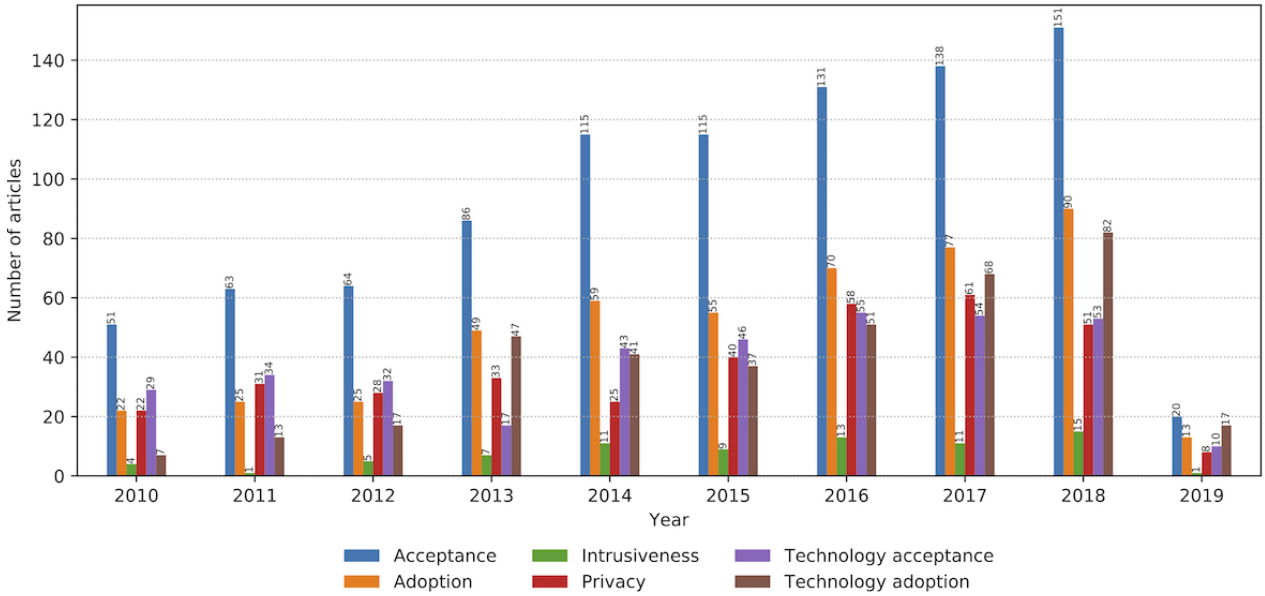
Concerns

Wearables will reshape individuals and society, promoting self-care and health management, moving care outside hospitals, affecting enterprises, and revolutionizing health care [8,82,83]. Their seamless integration into consumers' electronics is well witnessed throughout the Consumer Intelligence Series on wearables from 2014 and 2016 [8,83]. According to these sources, numerous user concerns such as design, accuracy, reliability, security, privacy, and dampened human interaction are becoming less worrying to the users. Research on sensor materials and communication solutions can provide advances in human-centered design and enhance the user experience.

Another big hurdle for deploying wearable systems in the real-world concerns technology acceptance [84,85]. Even though wearables are adopted by the millennials, the older population is still uncomfortable with using and relying on technology. As opposed to the smartphone, the use of wearables in fitness and well-being scenarios does not have clear usage need and benefits. Consumers complain about uncomfortable and unattractive design, short battery life, and frequent connectivity challenges [8]. With the first wearable devices, we have witnessed a *wearables fatigue* attitude, which is noticeable in a significant percentage of wearables being discarded within the first 6 months of use [86].

Our findings are aligned with the outcomes of the user feedback reviews on wearable technologies [8,83], as the identified articles confirm the steady increase in research addressing user-related issues such as technology acceptance, technology adoption, and privacy (Figure 12). The primary design requirements are that a wearable device must be fit for the purpose and seamlessly adapted to the user's lifestyle to be accepted.

Figure 12. User-related properties and concerns addressed in the relevant literature corpus from January 2010 to February 2019.



Preserving privacy and confidentiality is a priority to be considered in design specifications. Communications should be encrypted and secured, and the involved parties should ensure confidentiality. This is particularly important in the case of wireless data communications that are easier to intercept [87]. Personal monitoring devices should unobtrusively authenticate the user identity using biometrics or key physiological signs (owner-aware devices).

Different user concerns, such as quality of experience, security, privacy, technology acceptance, and human-centered design, are relevant research topics in the wearable CH domain and can be used to identify future challenges and research trends. Although some of them (eg, quality of experience and human-centered design) might be decreased as the end user pool gains digital competence and technology matures with time, some of them (security, privacy, and technology acceptance) will probably evolve and mix with other, more societal research topics such as environmental impact, circular economy, and digitalization of society. These can raise a new set of concerns related to the socioeconomic impact of wearable technologies in combination with IoT and 5G technologies used for health care and lifestyle.

It is worth mentioning that another spectrum of concerns and barriers relates to the stakeholders involved in the provision and management of health care. Health professionals need scientific evidence on the reliability of collected data, the performance of analytical models mapping the collected data to disease progression, and eventually positive patient outcome in using wearable-based CH solutions [1]. Reshaping the health care critically depends on research work devoted to the design and evaluation of care pathways, provision of optimized feedback, and eventually providing evidence on long- and short-term cost-effectiveness of CH solutions [1,88].

Safety and Security

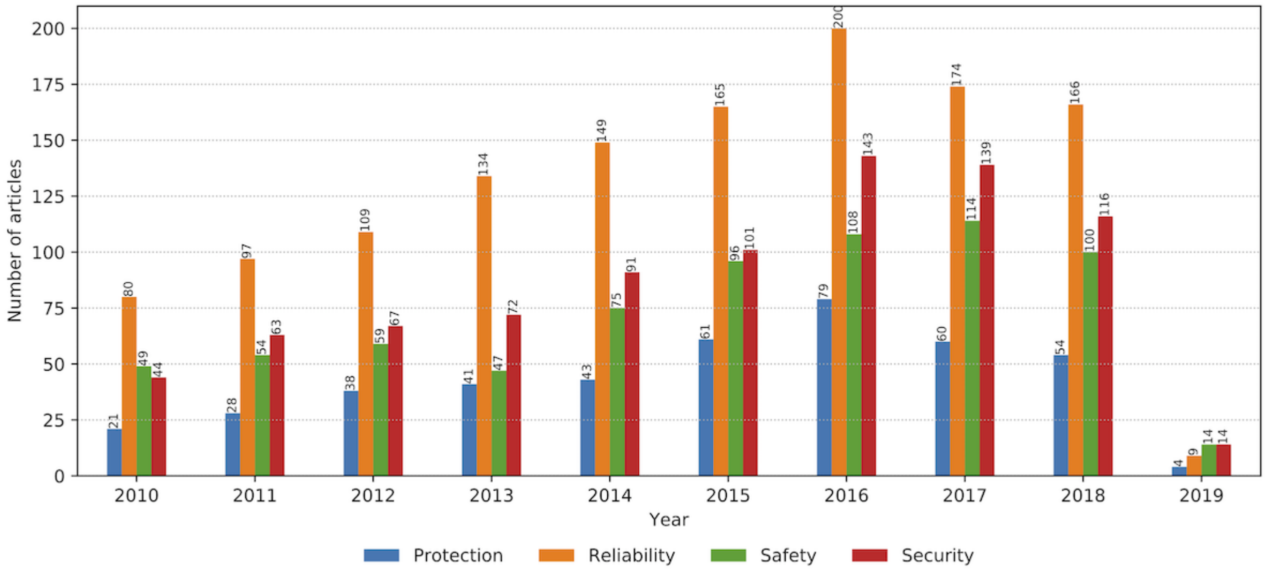
Safety and security are primary considerations for medical devices, tightly coupled with reliability at all system levels. Our findings confirm the increasing importance and research efforts related to these major user concerns (Figure 13). If the wearable device is required to perform safety-critical functions, the tolerance for error is zero. A failure in such a device can cost a life and that requires more effort and time (ultimately cost) to be invested in thoroughly testing and validating the device before it is deemed safe to use.

Along the life cycle of a wearable device, efficient mechanisms are required to detect and diagnose deviations occurring in the captured data. Correct differentiation of errors due to system-related faults from those due to a change in health status is a necessity. Increased level of false alarms (false positive) would prevent user reliance, reduce user alertness, and hamper user adherence to the provided feedback.

Both features, safety and security, are technology conditioned and should be ensured by the system design. Wearable medical devices are required to comply with IEC 62366-1:2015 standards [89] that regulate the application of usability engineering to medical devices to achieve approval. Two new EU regulations on medical devices issued in 2017 that will come into force in May 2020 are Regulation (EU) 2017/745 on medical devices [90] and Regulation (EU) 2017/746 on in vitro diagnostic medical devices [91]. These regulations will have a significant impact on the sector of medical devices that incorporates wearable technology. More stringent procedures for evaluation of medical devices and conformity should improve patient safety.

The CH paradigm involves more connectivity and communications into health care and medical devices. Any device connected to the internet is prone to be targeted for malicious purposes, putting it at a constant threat of damage, theft, and financial cost.

Figure 13. Safety- and security-related properties: protection, reliability, safety, and security. The trends in the period from January 2010 to February 2019 evidence their increasing presence (ie, relevance) as the wearable technology matures.



The number of detected data breaches in health care organizations has increased significantly in the last several years [92]. The reasons are not primarily technical but in part caused by the negligence and lack of knowledge of employees in treating this sensitive data and implementing the information security practices [92]. According to the 2017 Fourth Annual Data Breach Industry Forecast, health care organizations will be the most targeted sector with new, sophisticated attacks emerging. New security frameworks for mHealth are being proposed to ensure security and reliability of medical devices and personal health data [93-96].

After the General Data Protection Regulation was put in place on May 25, 2018, the requirements for data protection and privacy assurance have been raised and unified across Europe. The health care and monitoring systems have to adhere to the privacy by design principle, which requires the incorporation of privacy protection in systems design and not as an afterthought add-on solution.

Limitations of the Study

This study considered only 3 digital libraries, and some relevant articles from nonindexed publishers were not considered. However, keeping in mind the size of the considered digital libraries, we believe that the obtained results are indicative for the purpose of the study.

All digital libraries that were used in this work have different internal search engines with different rules for the maximum number of papers that can be retrieved and different formatting of search results. The papers obtained for this study are the results of the same search query sent to those different search engines. However, keeping in mind the number of papers that were analyzed within this scoping review, we believe that specificities of the publishers' search engines have limited impact and have not influenced the findings of this work.

In the future, the NLP tool kit needs to be extended to process more digital libraries. In addition, there is an apparent need of a Web app that will make it available to a wider audience. Until then, readers are encouraged to contact the authors if they are interested in using the tool kit.

Conclusions

Wearable medical solutions, integrated into the wider concept of IoT, provide for pervasive data acquisition from a body and beyond, and rely on powerful data analytics, smart networking, and machine-to-machine communications to facilitate patient-centric, personalized, and holistic care. Although technological innovations and availability support the emergence of CH solutions, the widespread adoption of wearables is still hindered by numerous concerns related to reliability, security, and cost-effectiveness.

This scoping review maps the scientific literature related to wearable technology in health care starting from January 2010 to February 2019, identifying the research trends related to enabling technology, and the trends in addressing the concerns from both user and technology perspectives. The NLP tool kit supported search procedures applied over 3 large digital libraries, IEEE, PubMed, and Springer, which provided for a representative subset of 2406 articles on wearable technologies for medical applications.

On the basis of the investigated sample, the main findings reflect key drivers in the field, some research gaps and relevant topics that would benefit from more systematic qualitative knowledge synthesis:

- User concerns were the least addressed topic, whereas the enabling technology research was the main focus in the literature within the observed time period;
- Major breakthroughs were made in sensor technology, data analytics, communications, and computing architectures (edge and cloud);
- Research on battery technology and efficient solutions for energy harvesting has lagged, implying energy efficiency

as one of the major constraints in designing wearable solutions for pervasive monitoring;

- Research on communication technologies focuses on 5G featuring low-latency, massive connectivity, and high capacity to mitigate the current challenges with respect to real-time feedback, energy, and computing constraints;
- The research related to the user-associated information delivery was mainly focused on monitoring and measurement information and much less on the provision of feedback recommendation and prescriptive insight; and
- The most addressed concerns from the user perspective were technology acceptance and issues related to safety and security, implying privacy and reliability as the most central topics.

This study confirms that applications of the wearable technology in the CH domain are becoming mature and established as a scientific domain. However, further research and development are required to improve their reliability, comfortability, and dependability levels. The research focus shifts from sensors and data analytics toward the sustainable delivery of valuable recommendations, reliable, energy-efficient, and low-latency communications and computation offloading. Sensor data integration goes beyond body-level integration to include context sensing, location and environment metrics, medical history, pattern of activities, and user preferences. This is essential for making wearables a robust patients' representation interface and reliable node of the IoT infrastructure that makes CH a reality.

There is a further need to explore and provide the literature evidence supporting the positive experiences, improved patient outcomes, and cost-effectiveness of CH solutions. Practical adoption in the field still demands design and validation of new care pathways, optimization of interventional strategies, and a sound business model.

Acknowledgments

This article/publication is based upon work from COST Action ENJECT TD 1405, supported by COST (European Cooperation in Science and Technology; www.cost.eu).

Authors' Contributions

TLT and VT conceived of the idea of scoping review, contributed to the scoping review, and drafted and edited the manuscript. EZ contributed to coding for the platform for scoping reviews and visualization of results. JMS contributed to the scoping review and editing of the manuscript. IC contributed to the scoping review methodology and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The list of the identified relevant research articles from the three selected digital libraries.
[XLSX File (Microsoft Excel File), 1MB - jmir_v21i9e14017_app1.xlsx]

References

1. Caulfield BM, Donnelly SC. What is connected health and why will it change your practice? QJM 2013 Aug;106(8):703-707. [doi: 10.1093/qjmed/hct114] [Medline: 23676416]
2. Kirk S. The wearables revolution: is standardization a help or a hindrance?: mainstream technology or just a passing phase? IEEE Consum Electr M 2014 Oct;3(4):45-50 [FREE Full text] [doi: 10.1109/mce.2014.2345996]
3. Tresp V, Overhage JM, Bundschus M, Rabizadeh S, Fasching PA, Yu S. Going digital: a survey on digitalization and large-scale data analytics in healthcare. Proc IEEE 2016 Nov;104(11):2180-2206 [FREE Full text] [doi: 10.1109/JPROC.2016.2615052]
4. Framingham M. IDC. 2019. IDC Reports Strong Growth in the Worldwide Wearables Market, Led by Holiday Shipments of Smartwatches, Wrist Bands, and Ear-Worn Devices URL: https://www.idc.com/getdoc.jsp?containerId=prUS44901819 [accessed 2019-03-13] [WebCite Cache ID 76qs7ZHS3]
5. Hunke N, Yusuf Z, Rüßmann M, Schmieg F, Bhatia A, Kalra N. The Boston Consulting Group. 2017. Winning in IoT: It's All About the Business Processes URL: https://www.bcg.com/publications/2017/hardware-software-energy-environment-winning-in-iot-all-about-winning-processes.aspx [accessed 2019-03-13] [WebCite Cache ID 76qsTyBwn]
6. Guler SD, Gannon M, Sicchio K. Crafting Wearables: Blending Technology with Fashion. Berkely, CA, USA: Apress; 2016.
7. Sun H, Zhang Z, Hu R, Qian Y. Wearable communications in 5G: challenges and enabling technologies. IEEE Veh Technol Mag 2018 Sep;13(3):100-109 [FREE Full text] [doi: 10.1109/MVT.2018.2810317]
8. Bothun D, Lieberman M. PwC Sverige. 2016. The Wearable Life 2.0: Connected Living in a Wearable World URL: https://www.pwc.se/sv/pdf-reports/the-wearable-life-2-0.pdf [accessed 2019-03-13] [WebCite Cache ID 76qsrO1TS]

9. 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]
10. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014 Feb;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]
11. Uddin MZ, Khaksar W, Torresen J. Ambient sensors for elderly care and independent living: a survey. Sensors (Basel) 2018 Jun 25;18(7):E2027 [FREE Full text] [doi: 10.3390/s18072027] [Medline: 29941804]
12. Erden F, Velipasalar S, Alkar AZ, Cetin AE. Sensors in assisted living: a survey of signal and image processing methods. IEEE Signal Process Mag 2016 Mar;33(2):36-44 [FREE Full text] [doi: 10.1109/MSP.2015.2489978]
13. 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]
14. Marcelino I, Laza R, Domingues P, Gómez-Meire S, Fdez-Riverola F, Pereira A. Active and assisted living ecosystem for the elderly. Sensors (Basel) 2018 Apr 17;18(4):E1246 [FREE Full text] [doi: 10.3390/s18041246] [Medline: 29673234]
15. Memon M, Wagner S, Pedersen C, Beevi F, Hansen F. Ambient assisted living healthcare frameworks, platforms, standards, and quality attributes. Sensors (Basel) 2014 Mar 4;14(3):4312-4341 [FREE Full text] [doi: 10.3390/s140304312] [Medline: 24599192]
16. Rodrigues J, de Rezende DD, Junqueira HA, Sabino MH, Prince RM, Al-Muhtadi J, et al. Enabling technologies for the internet of health things. IEEE Access 2018;6:13129-13141 [FREE Full text] [doi: 10.1109/ACCESS.2017.2789329]
17. Moher D, Stewart L, Shekelle P. All in the family: systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. Syst Rev 2015 Dec 22;4:183 [FREE Full text] [doi: 10.1186/s13643-015-0163-7] [Medline: 26693720]
18. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
19. Wilde LJ, Ward G, Sewell L, Müller AM, Wark PA. Apps and wearables for monitoring physical activity and sedentary behaviour: a qualitative systematic review protocol on barriers and facilitators. Digit Health 2018;4:2055207618776454 [FREE Full text] [doi: 10.1177/2055207618776454] [Medline: 29942637]
20. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016 Jun 9;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
21. Liu Y, Pharr MM, Salvatore GA. Lab-on-skin: a review of flexible and stretchable electronics for wearable health monitoring. ACS Nano 2017 Oct 24;11(10):9614-9635. [doi: 10.1021/acsnano.7b04898] [Medline: 28901746]
22. Aun NF, Soh PJ, Al-Hadi AA, Jamlos MF, Vandenbosch GA, Schreurs D. Revolutionizing wearables for 5G: 5G technologies: recent developments and future perspectives for wearable devices and antennas. IEEE Microw Mag 2017 May;18(3):108-124. [doi: 10.1109/MMM.2017.2664019]
23. Garcia-Perez C, Diaz-Zayas A, Rios A, Merino P, Katsalis K, Chang CY, et al. Improving the efficiency and reliability of wearable based mobile ehealth applications. Pervasive Mob Comput 2017 Sep;40:674-691 [FREE Full text] [doi: 10.1016/j.pmcj.2017.06.021]
24. Düking P, Fuss FK, Holmberg HC, Sperlich B. Recommendations for assessment of the reliability, sensitivity, and validity of data provided by wearable sensors designed for monitoring physical activity. JMIR Mhealth Uhealth 2018 Apr 30;6(4):e102 [FREE Full text] [doi: 10.2196/mhealth.9341] [Medline: 29712629]
25. Adapa A, Nah FF, Hall RH, Siau K, Smith SN. Factors influencing the adoption of smart wearable devices. Int J Hum-Comput Int 2017 Sep 14;34(5):399-409 [FREE Full text] [doi: 10.1080/10447318.2017.1357902]
26. Lunney A, Cunningham NR, Eastin MS. Wearable fitness technology: a structural investigation into acceptance and perceived fitness outcomes. Comput Hum Behav 2016 Dec;65:114-120 [FREE Full text] [doi: 10.1016/j.chb.2016.08.007]
27. Starner T. How wearables worked their way into the mainstream. IEEE Pervasive Comput 2014 Oct;13(4):10-15. [doi: 10.1109/MPRV.2014.66]
28. Sartor F, Papini G, Cox LG, Cleland J. Methodological shortcomings of wrist-worn heart rate monitors validations. J Med Internet Res 2018 Jul 2;20(7):e10108 [FREE Full text] [doi: 10.2196/10108] [Medline: 29967000]
29. Knapp M, Barlow J, Comas-Herrera A, Damant J, Freddolino PP, Hamblin K, et al. Policy Innovation Research Unit. 2015. The Case for Investment in Technology To Manage the Global Costs of Dementia URL: https://piru.lshtm.ac.uk/assets/files/Dementia_IT_PIRU_publ_18.pdf [accessed 2019-06-05] [WebCite Cache ID 78u9BHwCV]
30. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
31. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
32. Colquhoun HL, Levac D, O'Brien KK, Straus S, Tricco AC, Perrier L, et al. Scoping reviews: time for clarity in definition, methods, and reporting. J Clin Epidemiol 2014 Dec;67(12):1291-1294. [doi: 10.1016/j.jclinepi.2014.03.013] [Medline: 25034198]

33. Zdravevski E, Lameski P, Chorbev I, Goleva R, Pombo N, Garcia NM. Automation in systematic, scoping and rapid reviews by an NLP toolkit: a case study in enhanced living environments. In: Ganchev I, Garcia NM, Dobre C, Mavromoustakis CX, Goleva R, editors. Enhanced Living Environments: Lecture Notes in Computer Science. Volume 11369. Switzerland: Springer International Publishing; 2019:1-18.
34. Webster JJ, Kit C. Tokenization as the Initial Phase in NLP. In: Proceedings of the 14th Conference on Computational linguistics. 1992 Presented at: COLING'92; August 23-28, 1992; Nantes, France p. 1106-1110. [doi: 10.3115/992424.992434]
35. Manning C, Surdeanu M, Bauer J, Finkel J, Bethard S, Mc-Closky D. The Stanford CoreNLP Natural Language Processing Toolkit. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations. 2014 Presented at: ACL'14; June 22-27, 2014; Baltimore, Maryland p. 55-60 URL: https://doi.org/10.3115/v1/P14-5010 [doi: 10.3115/v1/P14-5010]
36. Bird S, Klein E, Loper E. Natural Language Processing With Python: Analyzing Text With The Natural Language Toolkit. United States of America: O'Reilly Media; 2009.
37. Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng 2007 May;9(3):90-95 [FREE Full text] [doi: 10.1109/MCSE.2007.55]
38. Hagberg A, Swart P, Chult DS. Exploring Network Structure, Dynamics, and Function Using NetworkX. In: Proceedings of the 7th Python in Science Conference. 2008 Presented at: SciPy'08; August 19-24, 2008; Pasadena, CA, USA p. 11-16 URL: https://conference.scipy.org/proceedings/scipy2008/paper_2/full_text.pdf
39. Meng E, Sheybani R. Insight: implantable medical devices. Lab Chip 2014 Sep 7;14(17):3233-3240. [doi: 10.1039/c4lc00127c] [Medline: 24903337]
40. Castano LM, Flatau AB. Smart fabric sensors and e-textile technologies: a review. Smart Mater Struct 2014 Apr 1;23(5):053001. [doi: 10.1088/0964-1726/23/5/053001]
41. Otsuka K, Wayman CM. Shape Memory Materials. Camebridge, UK: Cambridge University Press; 1999.
42. Syduzzaman M, Patwary SU, Farhana K, Ahmed S. Smart textiles and nano-technology: a general overview. J Textile Sci Eng 2015;5(1):1000181 [FREE Full text] [doi: 10.4172/2165-8064.1000181]
43. European Commission. 2018. Feedback From Stakeholders on the Smart Wearables Reflection and Orientation Paper URL: https://ec.europa.eu/digital-single-market/en/news/feedback-stakeholders-smart-wearables-reflection-and-orientation-paper [accessed 2019-03-14] [WebCite Cache ID 76rhGB7z4]
44. Yeo JC, Kenry K, Lim CT. Emergence of microfluidic wearable technologies. Lab Chip 2016;16(21):4082-4090 [FREE Full text] [doi: 10.1039/C6LC00926C]
45. Liana DD, Raguse B, Gooding JJ, Chow E. Recent advances in paper-based sensors. Sensors (Basel) 2012;12(9):11505-11526 [FREE Full text] [doi: 10.3390/s120911505] [Medline: 23112667]
46. Andreu-Perez J, Leff DR, Ip HM, Yang G. From wearable sensors to smart implants--toward pervasive and personalized healthcare. IEEE Trans Biomed Eng 2015 Dec;62(12):2750-2762. [doi: 10.1109/TBME.2015.2422751] [Medline: 25879838]
47. Pang C, Lee C, Suh KY. Recent advances in flexible sensors for wearable and implantable devices. J Appl Polym Sci 2013 Jun 26;130(3):1429-1441 [FREE Full text] [doi: 10.1002/app.39461]
48. Li X, Ballerini DR, Shen W. A perspective on paper-based microfluidics: current status and future trends. Biomicrofluidics 2012 Mar;6(1):11301-11313 [FREE Full text] [doi: 10.1063/1.3687398] [Medline: 22662067]
49. Bae WG, Kim D, Kwak MK, Ha L, Kang SM, Suh KY. Enhanced skin adhesive patch with modulus-tunable composite micropillars. Adv Healthc Mater 2013 Jan;2(1):109-113. [doi: 10.1002/adhm.201200098] [Medline: 23184425]
50. Kim D, Shin G, Kang YJ, Kim W, Ha JS. Fabrication of a stretchable solid-state micro-supercapacitor array. ACS Nano 2013 Sep 24;7(9):7975-7982. [doi: 10.1021/nn403068d] [Medline: 23952841]
51. Lipomi DJ, Lee JA, Vosgueritchian M, Tee BC, Bolander JA, Bao Z. Electronic properties of transparent conductive films of PEDOT:PSS on stretchable substrates. Chem Mater 2012 Jan 10;24(2):373-382 [FREE Full text] [doi: 10.1021/cm203216m]
52. Fan FR, Lin L, Zhu G, Wu W, Zhang R, Wang ZL. Transparent triboelectric nanogenerators and self-powered pressure sensors based on micropatterned plastic films. Nano Lett 2012 Jun 13;12(6):3109-3114. [doi: 10.1021/nl300988z] [Medline: 22577731]
53. Tee BC, Wang C, Allen R, Bao Z. An electrically and mechanically self-healing composite with pressure- and flexion-sensitive properties for electronic skin applications. Nat Nanotechnol 2012 Dec;7(12):825-832. [doi: 10.1038/nnano.2012.192] [Medline: 23142944]
54. Rogers JA, Nuzzo RG. Recent progress in soft lithography. Mater Today 2005 Feb;8(2):50-56 [FREE Full text] [doi: 10.1016/S1369-7021(05)00702-9]
55. Xu L, Gutbrod SR, Bonifas AP, Su Y, Sulkin MS, Lu N, et al. 3D multifunctional integumentary membranes for spatiotemporal cardiac measurements and stimulation across the entire epicardium. Nat Commun 2014 Feb 25;5:3329 [FREE Full text] [doi: 10.1038/ncomms4329] [Medline: 24569383]
56. Goverdovsky V, von Rosenberg W, Nakamura T, Looney D, Sharp DJ, Papavassiliou C, et al. Hearables: multimodal physiological in-ear sensing. Sci Rep 2017 Jul 31;7(1):6948 [FREE Full text] [doi: 10.1038/s41598-017-06925-2] [Medline: 28761162]

57. Ng KA, Greenwald E, Xu YP, Thakor N. Implantable neurotechnologies: a review of integrated circuit neural amplifiers. Med Biol Eng Comput 2016 Jan;54(1):45-62 [FREE Full text] [doi: 10.1007/s11517-015-1431-3] [Medline: 26798055]
58. Kiourti A, Psathas KA, Nikita KS. Implantable and ingestible medical devices with wireless telemetry functionalities: a review of current status and challenges. Bioelectromagnetics 2014 Jan;35(1):1-15. [doi: 10.1002/bem.21813] [Medline: 24115132]
59. da Silva HP, Fred A, Martins R. Biosignals for everyone. IEEE Pervasive Comput 2014 Oct;13(4):64-71 [FREE Full text] [doi: 10.1109/MPRV.2014.61]
60. Yonezawa Y, Miyamoto Y, Maki H, Ogawa H, Ninomiya I, Sada K, et al. A new intelligent bed care system for hospital and home patients. Biomed Instrum Technol 2005;39(4):313-319. [Medline: 16111412]
61. Gu WB, Poon CC, Leung HK, Sy MY, Wong MY, Zhang YT. A Novel Method for the Contactless and Continuous Measurement of Arterial Blood Pressure on a Sleeping Bed. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2009 Presented at: EMBS'09; September 3-6, 2009; Minneapolis, MN, USA p. 6084-6086 URL: https://doi.org/10.1109/IEMBS.2009.5335393 [doi: 10.1109/IEMBS.2009.5335393]
62. Kim KK, Lim YK, Park KS. The Electrically Non-Contacting ECG Measurement on the Toilet Seat Using the Capacitively-Coupled Insulated Electrodes. In: Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2004 Presented at: EMBS'04; September 1-5, 2004; San Francisco, CA, USA p. 2375-2378 URL: https://doi.org/10.1109/IEMBS.2004.1403688 [doi: 10.1109/IEMBS.2004.1403688]
63. Galinina O, Tabassum H, Mikhaylov K, Andreev S, Hossain E, Koucheryavy Y. On feasibility of 5G-grade dedicated RF charging technology for wireless-powered wearables. IEEE Wireless Commun 2016 Apr;23(2):28-37. [doi: 10.1109/MWC.2016.7462482]
64. Pillatsch P, Yeatman E, Holmes AS. A piezoelectric frequency up-converting energy harvester with rotating proof mass for human body applications. Sensor Actuat A-Phys 2014 Feb;206:178-185. [doi: 10.1016/j.sna.2013.10.003]
65. Magno M, Spadaro L, Singh J, Benini L. Kinetic Energy Harvesting: Toward Autonomous Wearable Sensing for Internet of Things. In: Proceedings of the International Symposium on Power Electronics, Electrical Drives, Automation and Motion. 2016 Presented at: SPEEDAM'16; June 22-24, 2016; Anacapri, Italy URL: https://doi.org/10.1109/SPEEDAM.2016.7525995 [doi: 10.1109/SPEEDAM.2016.7525995]
66. Tian R, Wan C, Hayashi N, Aoai T, Koumoto K. Wearable and flexible thermoelectrics for energy harvesting. MRS Bull 2018 Mar 9;43(3):193-198 [FREE Full text] [doi: 10.1557/mrs.2018.8]
67. Torfs T, Leonov V, Yazicioglu RF, Merken P, van Hoof C, Vullers RJ, et al. Wearable Autonomous Wireless Electro-Encephalography System Fully Powered by Human Body Heat. In: Proceedings of IEEE SENSORS. 2008 Presented at: SENSORS'08; October 26-29, 2008; Lecce, Italy p. 1269-1272 URL: https://doi.org/10.1109/ICSENS.2008.4716675 [doi: 10.1109/ICSENS.2008.4716675]
68. Soyata T, Copeland L, Heinzelman W. RF energy harvesting for embedded systems: a survey of tradeoffs and methodology. IEEE Circuits Syst Mag 2016;16(1):22-57 [FREE Full text] [doi: 10.1109/MCAS.2015.2510198]
69. Chai Z, Zhang N, Sun P, Huang Y, Zhao C, Fan HJ, et al. Tailorable and wearable textile devices for solar energy harvesting and simultaneous storage. ACS Nano 2016 Oct 25;10(10):9201-9207. [doi: 10.1021/acsnano.6b05293] [Medline: 27701868]
70. Jia W, Wang X, Imani S, Bandodkar AJ, Ramírez J, Mercier PP, et al. Wearable textile biofuel cells for powering electronics. J Mater Chem A 2014;2(43):18184-18189 [FREE Full text] [doi: 10.1039/C4TA04796F]
71. Hiramoto T, Takeuchi K, Mizutani T, Ueda A, Saraya T, Kobayashi M, et al. Ultra-Low Power and Ultra-Low Voltage Devices and Circuits for IoT Applications. In: Proceedings of the Conference on Silicon Nanoelectronics Workshop. 2016 Presented at: SNW'16; June 12-13, 2016; Honolulu, HI, USA URL: https://doi.org/10.1109/SNW.2016.7578025 [doi: 10.1109/SNW.2016.7578025]
72. Lee HS, Park CB, Noh KJ, Sunwoo J, Choi H, Cho IY. Wearable personal network based on fabric serial bus using electrically conductive yarn. ETRI J 2010 Oct 6;32(5):713-721 [FREE Full text] [doi: 10.4218/etrij.10.1510.0084]
73. IEEE Xplore Digital Library. 2012. 802.15.6-2012 - IEEE Standard for Local and Metropolitan Area Networks - Part 15.6: Wireless Body Area Networks URL: https://ieeexplore.ieee.org/document/6161600 [accessed 2019-03-13]
74. Suciu G, Suciu V, Martian A, Craciunescu R, Vulpe A, Marcu I, et al. Big data, internet of things and cloud convergence--an architecture for secure e-health applications. J Med Syst 2015 Nov;39(11):141. [doi: 10.1007/s10916-015-0327-y] [Medline: 26345453]
75. Chen M, Gonzalez-Valenzuela S, Vasilakos AV, Cao H, Leung VC. Body area networks: a survey. Mobile Netw Appl 2010 Aug 18;16(2):171-193 [FREE Full text] [doi: 10.1007/s11036-010-0260-8]
76. Poon CC, Lo BP, Yuce MR, Alomainy A, Hao Y. Body sensor networks: in the era of big data and beyond. IEEE Rev Biomed Eng 2015;8:4-16. [doi: 10.1109/RBME.2015.2427254] [Medline: 25935046]
77. Hulzink J, Konijnenburg M, Ashouei M, Breeschoten A, Berset T, Huisken J, et al. An ultra low energy biomedical signal processing system operating at near-threshold. IEEE Trans Biomed Circuits Syst 2011 Dec;5(6):546-554. [doi: 10.1109/TBCAS.2011.2176726] [Medline: 23852552]
78. Mohammed EA, Far BH, Naugler C. Applications of the MapReduce programming framework to clinical big data analysis: current landscape and future trends. BioData Min 2014;7:22 [FREE Full text] [doi: 10.1186/1756-0381-7-22] [Medline: 25383096]

79. Gravina R, Alinia P, Ghasemzadeh H, Fortino G. Multi-sensor fusion in body sensor networks: state-of-the-art and research challenges. Inform Fusion 2017 May;35:68-80 [FREE Full text] [doi: 10.1016/j.inffus.2016.09.005]
80. Kamar E, Horvitz E. Jogger: Models for Context-Sensitive Reminding. In: Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems. 2011 Presented at: AAMAS'11; May 2–6, 2011; Taipei, Taiwan p. 1089-1090.
81. Mulvenna M, Carswell W, Mccullagh P, Augusto JC, Zheng H, Jeffers P, et al. Visualization of data for ambient assisted living services. IEEE Commun Mag 2011 Jan;49(1):110-117 [FREE Full text] [doi: 10.1109/MCOM.2011.5681023]
82. Schüll ND. Data for life: wearable technology and the design of self-care. Biosocieties 2016 Oct 13;11(3):317-333 [FREE Full text] [doi: 10.1057/biosoc.2015.47]
83. PwC Sverige. 2014. The Wearable Future URL: https://www.pwc.se/sv/pdf-reports/consumer-intelligence-series-the-wearable-future.pdf [accessed 2019-03-13] [WebCite Cache ID 76qxdtiqw]
84. Cosco TD, Firth J, Vahia I, Sixsmith A, Torous J. Mobilizing mhealth data collection in older adults: challenges and opportunities. JMIR Aging 2019 Mar 19;2(1):e10019 [FREE Full text] [doi: 10.2196/10019]
85. Hill R, Betts LR, Gardner SE. Older adults' experiences and perceptions of digital technology: (dis)empowerment, wellbeing, and inclusion. Comput Human Behav 2015 Jul;48:415-423 [FREE Full text] [doi: 10.1016/j.chb.2015.01.062]
86. Triteq. 2015. Wearable Technologies by Triteq URL: https://triteq.com/news/wearable-technologies-by-triteq [accessed 2019-03-13] [WebCite Cache ID 76qx2iIR7]
87. Al Ameen M, Liu J, Kwak K. Security and privacy issues in wireless sensor networks for healthcare applications. J Med Syst 2012 Feb;36(1):93-101 [FREE Full text] [doi: 10.1007/s10916-010-9449-4] [Medline: 20703745]
88. Noah B, Keller MS, Mosadeghi S, Stein L, Johl S, Delshad S, et al. Impact of remote patient monitoring on clinical outcomes: an updated meta-analysis of randomized controlled trials. NPJ Digit Med 2018 Jan 15;1(1):20172 [FREE Full text] [doi: 10.1038/s41746-017-0002-4]
89. International Organization for Standardization. 2015. Medical Devices — Part 1: Application of Usability Engineering to Medical Devices URL: https://www.iso.org/obp/ui/#iso:std:iec:62366:-1:ed-1:v1:en,fr [accessed 2019-03-13] [WebCite Cache ID 76roDXeeu]
90. EUR-Lex. 2017. Regulation (EU) 2017/745 of the European Parliament and of the Council on Medical Devices URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745 [accessed 2019-03-14] [WebCite Cache ID 76rooand3]
91. EUR-Lex. 2017. Regulation (EU) 2017/746 of the European Parliament and of the Council in in vitro Diagnostic Medical Devices URL: https://eur-lex.europa.eu/eli/reg/2017/746/oj [accessed 2019-03-13] [WebCite Cache ID 76qy05are]
92. Kierkegaard P. Medical data breaches: notification delayed is notification denied. Comput Law Secur Rev 2012 Apr;28(2):163-183 [FREE Full text] [doi: 10.1016/j.clsr.2012.01.003]
93. Matousek K. Security and Reliability Considerations for Distributed Healthcare Systems. In: Proceedings of the 42nd Annual IEEE International Carnahan Conference on Security Technology. 2008 Presented at: ICCST'08; October 13-16, 2008; Prague, Czech Republic URL: https://doi.org/10.1109/CCST.2008.4751326 [doi: 10.1109/CCST.2008.4751326]
94. Eldosouky A, Saad W. On the Cybersecurity of m-Health IoT Systems With LED Bitslice Implementation. In: Proceedings of the International Conference on Consumer Electronics. 2018 Presented at: ICCE'18; January 12-14, 2018; Las Vegas, NV, USA URL: https://doi.org/10.1109/ICCE.2018.8326298 [doi: 10.1109/ICCE.2018.8326298]
95. Al Hamid HA, Rahman SM, Hossain MS, Almogren A, Alamri A. A security model for preserving the privacy of medical big data in a healthcare cloud using a fog computing facility with pairing-based cryptography. IEEE Access 2017;5:22313-22328 [FREE Full text] [doi: 10.1109/ACCESS.2017.2757844]
96. Classen J, Wegemer D, Patras P, Spink T, Hollick M. Anatomy of a vulnerable fitness tracking system. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018 Mar 26;2(1):1-24 [FREE Full text] [doi: 10.1145/3191737]
97. European Cooperation in Science and Technology. URL: https://www.cost.eu/

Abbreviations

5G: fifth generation
AAL: ambient assisted living
BAN: body area network
CH: connected health
DOI: digital object identifier
IEEE: Institute of Electrical and Electronics Engineers
IoT: Internet of Things
mHealth: mobile health
mmW: millimeter wave
NLP: natural language processing
PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analysis
RRH: remote radio header

VLC: visible light communication
WLAN: wireless local area network

Edited by B Caulfield; submitted 14.03.19; peer-reviewed by V Curcin, PA Silva, E Guisado-Fernandez; comments to author 18.05.19; revised version received 09.06.19; accepted 19.06.19; published 05.09.19.

Please cite as:
Loncar-Turukalo T, Zdravevski E, Machado da Silva J, Chouvarda I, Trajkovik V
Literature on Wearable Technology for Connected Health: Scoping Review of Research Trends, Advances, and Barriers
J Med Internet Res 2019;21(9):e14017
URL: https://www.jmir.org/2019/9/e14017/
doi:10.2196/14017
PMID:31489843

Original Paper

An Assessment of Physical Activity Data Collected via a Smartphone App and a Smart Band in Breast Cancer Survivors: Observational Study

Il Yong Chung[1*], MD; Miyeon Jung[2*], BSc; Sae Byul Lee[1], MD, PhD; Jong Won Lee[1], MD, PhD; Yu Rang Park[3], PhD; Daegon Cho[2], PhD; Haekwon Chung[4], BSc; Soyoung Youn[5], MD; Yul Ha Min[6], PhD; Hye Jin Park[1], BSN; Minsun Lee[1], BSN; Seockhoon Chung[5], MD, PhD; Byung Ho Son[1], MD, PhD; Sei-Hyun Ahn[1], MD, PhD

[1]Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[2]College of Business, Korea Advanced Institute of Science and Technology, Seoul, Republic of Korea
[3]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea
[4]Swallaby Co, Ltd, Seoul, Republic of Korea
[5]Department of Psychiatry, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[6]College of Nursing, Gachon University, Incheon, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Sae Byul Lee, MD, PhD
Department of Surgery
Asan Medical Center
University of Ulsan College of Medicine
88, Olympic-ro 43-gil, Songpa-gu, 05505
Seoul,
Republic of Korea
Phone: 82 10 7209 4620
Email: newstar153@hanmail.net

Abstract

Background: Although distress screening is crucial for cancer survivors, it is not easy for clinicians to recognize distress. Physical activity (PA) data collected by mobile devices such as smart bands and smartphone apps have the potential to be used to screen distress in breast cancer survivors.

Objective: The aim of this study was to assess data collection rates of smartphone apps and smart bands in terms of PA data, investigate the correlation between PA data from mobile devices and distress-related questionnaires from smartphone apps, and demonstrate factors associated with data collection with smart bands and smartphone apps in breast cancer survivors.

Methods: In this prospective observational study, patients who underwent surgery for breast cancer at Asan Medical Center, Seoul, Republic of Korea, between June 2017 and March 2018 were enrolled and asked to use both a smartphone app and smart band for 6 months. The overall compliance rates of the daily PA data collection via the smartphone walking apps and wearable smart bands were analyzed in a within-subject manner. The longitudinal daily collection rates were calculated to examine the dropout pattern. We also performed multivariate linear regression analysis to examine factors associated with compliance with daily collection. Finally, we tested the correlation between the count of daily average steps and distress level using Pearson correlation analysis.

Results: A total of 160 female patients who underwent breast cancer surgeries were enrolled. The overall compliance rates for using a smartphone app and smart bands were 88.0% (24,224/27,513) and 52.5% (14,431/27,513), respectively. The longitudinal compliance rate for smartphone apps was 77.8% at day 180, while the longitudinal compliance rate for smart bands rapidly decreased over time, reaching 17.5% at day 180. Subjects who were young, with other comorbidities, or receiving antihormonal therapy or targeted therapy showed significantly higher compliance rates to the smartphone app. However, no factor was associated with the compliance rate to the smart band. In terms of the correlation between the count of daily steps and distress level, step counts collected via smart band showed a significant correlation with distress level.

Conclusions: Smartphone apps or smart bands are feasible tools to collect data on the physical activity of breast cancer survivors. PA data from mobile devices are correlated with participants' distress data, which suggests the potential role of mobile devices in the management of distress in breast cancer survivors.

Trial Registration: ClinicalTrials.gov NCT03072966; https://clinicaltrials.gov/ct2/show/NCT03072966

(J Med Internet Res 2019;21(9):e13463) doi:10.2196/13463

KEYWORDS

telemedicine; breast neoplasms; mobile apps; quality of life; stress, psychological; patient compliance; smartphone; mobile phone; wearable electronic devices; survivorship

Introduction

Distress screening is important in cancer survivorship care. The prevalence of depression and anxiety were reported to be high among cancer survivors [1,2]. The National Comprehensive Cancer Network and American Society of Clinical Oncology guidelines recommend all cancer patients be screened for depressive symptoms [3,4].

Although distress screening is crucial for cancer survivors, it is not easy for clinicians to recognize distress. Some studies have demonstrated that the prevalence and severity of depression are underestimated [4,5]. Moreover, conventional screening tools for distress are patient-reported outcomes (PROs) based on paper questionnaires, which are subject to recall bias and do not reflect real-time episodes of distress.

Electronic PROs (ePROs) using a smartphone app can be used as a screening tool for depression in practice [6]. While the completion rate of online questionnaires in cancer patients was about 15.0% from home, the overall compliance rate increased when patients used smartphone apps [7-9]. Although ePROs using apps is a feasible tool for distress screening, there is still room for improvement in adherence rates.

Researchers have recently suggested that changes in physical activities (PA) per the data collected by wearable devices can predict mood changes [10-11]. PA data collected by mobile devices such as smart bands and smartphone apps to screen distress in cancer survivors could offer tremendous advantages. However, the previous studies were not conducted in cancer survivors, the sample size was relatively small, and distress or depression was not evaluated by conventional screening tools.

The prerequisite and essential condition for using mobile devices for data collection is compliance to mobile devices. However, long-term compliance to mobile devices, especially in clinical setting, has rarely been studied. According to a systematic review [12], among compliance studies with mobile device–based ecological momentary assessment, about half of the studies (24/42, 62%) recruited youth from nonclinical settings, the length of studies ranged only from 2 to 42 days, and the sample size in clinical settings was small (range 5 to 84 participants). Until now, there has been a lack of evidence about long-term compliance and data collection using mobile devices.

We conducted a single-center, prospective, observational study with 160 breast cancer patients who underwent breast cancer surgery and 6-month follow-up. The aim of this study was to assess data collection rates of smartphone apps and smart bands in terms of PA data to investigate the correlation between PA data from mobile devices and distress-related questionnaire data from smartphone apps and demonstrate factors associated with data collection with smart bands and smartphone apps in breast cancer survivors.

Methods

Study Setting and Subjects

This prospective observational study recruited patients who underwent surgeries for breast cancer at Asan Medical Center, Seoul, Republic of Korea. Patients were eligible for study participation if they were women between the ages of 20 and 65 years and had their own Android smartphones compatible with the WalkOn app (Swallaby Co, Ltd), a free activity tracking app modified for this study. Patients who had distant metastasis, recurrent breast cancer, or severe medical conditions such as cardiovascular disease and those who had no capability of using a smartphone were excluded. Patients on chemotherapy were also excluded.

Written informed consent was obtained from all subjects at enrollment. The study protocol was approved by the institutional review board at Asan Medical Center (2016-0819). This study was registered on the ClinicalTrial.gov website (NCT03072966).

Recruitment and Follow-Up

During the hospital stay after breast cancer surgery, subjects were contacted by a clinical research assistant. After consenting to participate, subjects completed paper-based questionnaires (Distress Thermometer [DT] and Patient Health Questionnaire–9 [PHQ-9]) at baseline. The clinical research assistant helped to download the Android-based app (WalkOn) to the participants' smartphones.

At the first visit to the clinic after discharge, subjects who are not undergoing adjuvant chemotherapy received wearable devices (Fitbit Charge HR, Fitbit Inc), and the Fitbit apps were installed on their smartphones. Subjects scheduled to receive adjuvant chemotherapy were provided the wearable devices after chemotherapy. The clinical research assistant called participants after 3 weeks to follow up and assist with collection of data from the wearable devices. At the 3- and 6-month follow-ups, participants' smartphone apps and smart bands were checked by the clinical research assistant.

Smartphone App and App-Based Questionnaire

The health-related smartphone app, WalkOn, has been developed by the mobile phone health care app company, Swallaby Co, Ltd. The app provides users with platforms for tracking their daily steps and creates mobile communities where users can communicate with each other and view other members' daily steps to motivate them and promote health-related activities.

App-based self-reporting questionnaires have been included in this app, which exclusively allows study participants to complete daily, weekly, and biweekly questionnaires. The daily questionnaires developed and previously reported by the authors consisted of self-reporting modules for daily anxiety, sleep, and emotion [6,9]. Weekly DT and biweekly PHQ-9 questionnaires were also collected through the app, and push alarms from Sunday to Tuesday every week were sent to subjects' smartphones as notification.

Activity Tracking From Smartphone Apps and Smart Bands

We used the app to collect data on participants' daily steps. Once the app was activated at enrollment, participants were instructed to open the app and pull-to-refresh the front page of the app at least once weekly to send the weekly bundle of daily walking data to a central database system. The central database system collected the daily walking data for each subject with anonymized user ID, item ID, date and time of input, and input value.

As a second channel for data collection, a smart band (Fitbit Charge HR) was also used to monitor daily PA and sleep patterns in real time during the 6-month study period. Briefly, the Fitbit Charge HR is designed to measure various PAs such as steps taken, distance traveled, and calories burned. It also shows users how many minutes they have been active during the day. Similar to the walking app, after Fitbit activation, participants were instructed to access the Fitbit app and pull-to-refresh at least once per week to send the weekly set of daily walking data to the server. The Fitbit app also sent a push notification to the users about changes in daily activity patterns. While the mobile walking app only collects the number of steps per day, the Fitbit collects various indicators about PA each minute. The central database system collected all indicators at both minute- and day-levels using anonymized user ID, item ID, date and time of input, and input values.

Statistical Analysis

The feasibility of the app-based and smart band–based PA collecting systems were analyzed by calculating individual-level data collection rates, defined as the total number of days in which data collection was completed divided by the number of follow-up days for each patient as well as a longitudinal day-level data collection rate, defined as the total number of patients with PA data divided by the number of patients who did not drop out on a single specific day from the start of data collection. The cumulative longitudinal day-level data collection rate was defined as the mean of all longitudinal day-level data collection rates from the starting day until the specified day. For example, the longitudinal day-level collection rate at day 20 indicated the number of patient data collection points on the 20th day divided by 160, while the cumulative longitudinal day-level data collection rate at 20th day was the mean value of all longitudinal day-level data collection rates from day 0 to day 20. The definitions of individual-level, longitudinal day-level, and cumulative longitudinal day-level data collection rates were similar with the measurements of compliance which were described previously [9].

Student *t* tests and chi-square tests were used to investigate the differences in patient characteristics between high and low data collection rate groups. The variables of interest included age, marital status, education, occupation, comorbidity, past episode of depression, surgery, chemotherapy, radiation therapy, antihormonal therapy, targeted therapy, and stage. The mean value of individual-level data collection rates was used as the cutoff to divide the high and low collection rate groups. The mean individual-level rates were 91.1% for the smartphone app and 55.0% for the smart band.

The Pearson correlation test was used to calculate the correlation between app-based and band-based collection rates. To more comprehensively investigate the factors affecting individual-level collection rates, we conducted multivariate linear regression analysis for both app- and band-based data collection rates with all variables. Since two patients did not report baseline distress levels in presurvey, we exclude them from step 2 and 3 analyses.

Finally, we examined the correlation between app-based and band-based step counts and distress-related psychological questionnaires using Pearson correlation tests.

Results

Descriptive Statistics

From June 2017 to January 2018, we assessed for study eligibility 1247 consecutive patients who underwent breast cancer surgery (Figure 1). After screening, 591 patients were excluded; 176 patients were not able to be contacted during the hospital stay and 320 patients refused to join the study. A total of 160 patients were enrolled in this study. During follow-up, 33 participants withdrew consent due to inconvenience (n=26), transfer to other hospitals (n=2), app compatibility problems after update (n=4), and breast cancer recurrence (n=1).

Table 1 summarizes descriptive statistics of the demographic and clinical characteristics of the subjects as absolute and relative frequencies. Subjects were aged mean 45.3 [SD 6.8] years (range 28 to 64 years), and about three-quarters (121/160, 75.6%) were younger than 50 years. Sixty-four subjects completed adjuvant or neoadjuvant chemotherapy before beginning daily data collection. Among the 160 patients, 75.6% (121/160) were less than 50 years of age, 66.9% (107/160) had an educational attainment of college level or higher, and 47.5% (76/160) were currently employed. The mean baseline EQ5D-5L score was 0.91 (SD 0.1). Regarding breast cancer stages, 13.1% (21/160) of patients had stage 0, 48.1% (77/160) had stage I, 24.4% (39/160) had stage II, and 14.4% (23/160) had stage III disease.

Figure 1. Subject enrollment.

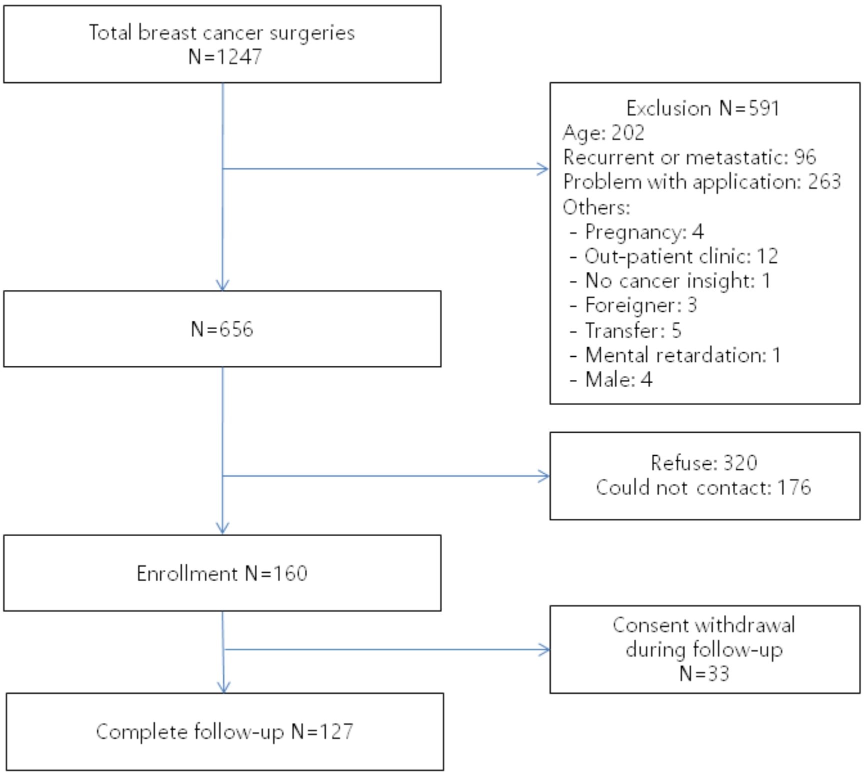


Table 1. Subject demographics.

Characteristics	Total (N=160)	App compliance			Band compliance		
		Higher (n=133)	Lower (n=27)	*P* value	Higher (n=80)	Lower (n=80)	*P* value
Age (years), mean (SD)	**45.3 (6.8)**	**44.9 (6.9)**	**47.1 (6.4)**	**.13**	**44.9 (6.5)**	**45.7 (7.1)**	**.43**
<50, n (%)	121 (75.6)	103 (77.4)	18 (66.7)	.35	65 (81.3)	56 (70.0)	.14
≥50, n (%)	39 (24.4)	30 (22.6)	9 (33.3)	—[a]	15 (18.7)	24 (30.0)	—
Marital status, n (%)				**.63**			**.73**
Married	137 (85.6)	114 (85.7)	23 (85.2)		70 (87.5)	67 (83.8)	
Single	20 (12.5)	17 (12.8)	3 (11.1)		9 (11.3)	11 (13.7)	
Other	3 (1.9)	2 (1.5)	1 (3.7)		1 (1.2)	2 (2.5)	
Education, n (%)				**.49**			**.74**
≤High school	53 (33.1)	42 (31.6)	11 (40.7)		25 (31.3)	28 (35.0)	
>High school	107 (66.9)	91 (68.4)	16 (59.3)		55 (68.7)	52 (65.0)	
Occupation, n (%)				**.33**			**.43**
Yes	76 (47.5)	66 (49.6)	10 (37.0)		41 (51.3)	35 (43.8)	
No	84 (52.5)	67 (50.4)	17 (63.0)		39 (48.8)	45 (56.2)	
Comorbidity, n (%)				**.03**			**.24**
Yes	108 (67.5)	95 (71.4)	13 (48.1)		50 (62.5)	58 (72.5)	
No	52 (32.5)	38 (28.6)	14 (51.9)		30 (37.5)	22 (27.5)	
Past episode of depression, n (%)				**.498**			**.84**
Yes	2 (1.3)	2 (1.5)	0 (0)		1 (1.3)	1 (1.3)	
No	152 (95.0)	127 (95.5)	25 (92.6)		77 (96.3)	75 (93.7)	
No response	6 (3.8)	4 (3.0)	2 (7.4)		2 (2.4)	4 (5.0)	
Surgery, n (%)				**.60**			**.08**
Mastectomy	11 (6.9)	9 (6.8)	2 (59.3)		7 (8.8)	4 (5.0)	
Breast-conserving surgery	107 (66.9)	91 (68.4)	16 (7.4)		58 (72.5)	49 (61.3)	
Mastectomy with reconstruction	42 (26.3)	33 (24.8)	9 (33.3)		15 (18.8)	27 (33.7)	
Previous chemotherapy, n (%)				**.28**			**.33**
Yes	65 (40.6)	51 (38.3)	14 (51.9)		36 (45.0)	29 (36.3)	
No	95 (59.4)	82 (61.7)	13 (48.1)		44 (55.0)	51 (63.7)	
Antihormonal therapy, n (%)				**.04**			**>.99**
Yes	135 (84.4)	116 (87.2)	19 (70.4)		68 (85.0)	67 (83.7)	
No	25 (15.6)	17 (12.8)	8 (29.6)		12 (15.0)	13 (16.3)	
Radiation therapy, n (%)				**.47**			**.34**
Yes	124 (77.5)	105 (78.9)	19 (70.4)		65 (81.3)	59 (73.8)	
No	36 (22.5)	28 (21.1)	8 (29.6)		15 (18.7)	21 (26.3)	
Targeted therapy, n (%)				**.01**			**.76**
Yes	148 (92.5)	127 (95.5)	21 (77.8)		73 (91.3)	75 (93.8)	
No	12 (7.5)	6 (4.5)	6 (22.2)		7 (8.7)	5 (6.2)	
Stage, n (%)				**.31**			**.42**
0	21 (13.1)	16 (12.0)	5 (18.5)		8 (10.0)	13 (16.3)	
I	77 (48.1)	68 (51.1)	9 (33.3)		39 (48.8)	38 (47.5)	
II	39 (24.4)	30 (22.6)	9 (33.3)		23 (28.7)	16 (20.0)	
III	23 (14.4)	19 (14.3)	4 (14.8)		10 (12.5)	13 (16.3)	

Characteristics	Total (N=160)	App compliance			Band compliance		
		Higher (n=133)	Lower (n=27)	*P* value	Higher (n=80)	Lower (n=80)	*P* value
Distress thermometer score, n (%)				**.29**			**.32**
≥5	59 (36.9)	46 (34.6)	13 (48.1)		26 (32.5)	33 (41.3)	
≤5	99 (61.9)	85 (63.9)	14 (51.9)		53 (66.3)	46 (57.5)	
No response	2 (1.2)	2 (1.5)	0 (0)		1 (1.2)	1 (1.2)	
HRQOL[b] with EQ5D-5L[c]				**.14**			**.44**
EQ5D-5L index, mean (SD)	0.91 (0.1)	0.92 (0.1)	0.87 (0.1)		0.92 (0.1)	0.90 (0.1)	
No response, n (%)	2 (1.2)	2 (1.5)	0 (0)		1 (1.2)	1 (1.2)	
PHQ-9[d] total score, n (%)				**.93**			**.71**
≥11	37 (23.2)	30 (22.6)	7 (25.9)		20 (25.0)	17 (21.3)	
≤11	121 (75.6)	101 (75.9)	20 (74.1)		59 (73.8)	62 (77.5)	
No response	2 (1.2)	2 (1.5)	0 (0)		1 (1.2)	1 (1.2)	
Median follow-up in days, median (SD)	164 (39.3)	164 (39.3)	164 (40.4)	.95	160 (46.6)	168 (30.2)	.18

[a] Not applicable.

[b] HRQOL: Health-Related Quality of Life.

[c] EQ5D-5L: 5-dimension 5-level health questionnaire.

[d] PHQ-9: Patient Health Questionnaire–9.

Data Collection Rates

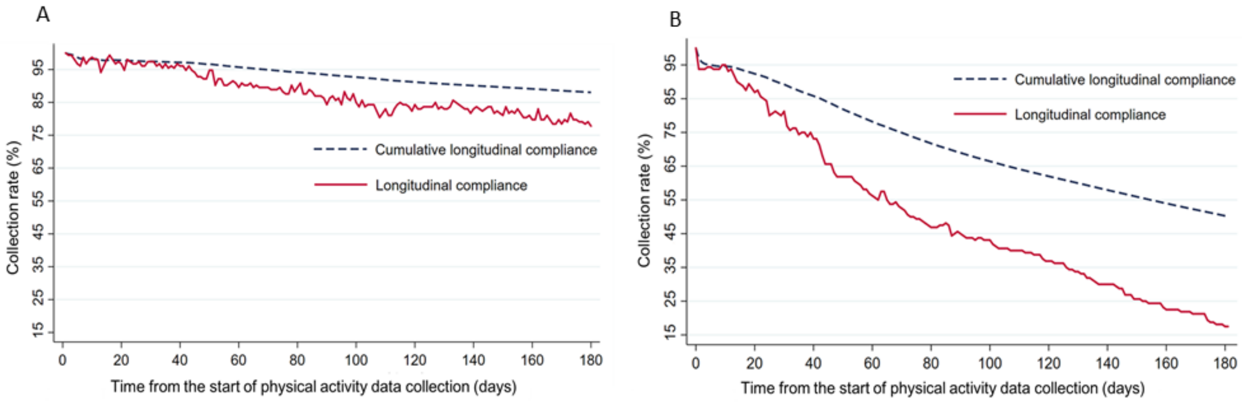
Among 160 participants, 127 completed the 180-day study period, and 33 patients had different follow-up periods depending on the withdrawal date (minimum: 11 days, maximum: 176 days, median: 101 days). The total sum of follow-up days of all participants is 27,513, slightly less than 28,800 (160*180) because of patient dropout.

In total, 24,224 and 14,431 data points were collected via smartphone apps and smart bands, respectively. The overall data collection rates for using the smartphone app and smart band were 88.05% (24,224/27,513) and 52.45% (14,431/27,513), respectively.

The data collection rate via smartphone app was >80% for 141 subjects (88.1%) and via smart band was >80% for 53 participants (33.1%). The longitudinal day-level data collection rates from day 1 to day 180 were calculated at daily intervals and plotted (Figure 2). For the smartphone app, the longitudinal data collection rate remained above 75%, reaching 77.8% at day 180, while the cumulative longitudinal data collection curve showed a steady decrease to 88% (representing overall data collection) at day 180. Although cumulative longitudinal data collection with smart bands was 46% at day 90, longitudinal compliance rate for smart bands more rapidly decreased, reaching 17.5% at day 180.

Daily data collection of mobile questionnaires is relatively smaller than the daily data collection of the semiautomated walking steps collection systems. The daily questionnaires consisted of self-reporting modules for daily anxiety, sleep, and emotion questionnaire [6,9], and in total, 20,733 data points were collected via the smartphone app. The overall data collection rate for using the smartphone app self-reporting modules was 75.36% (20,733/27,513). Therefore, we can use the self-reported value of daily anxiety, sleep, and emotion as focal outcome variables to develop distress screening system in our future study. The weekly DT and biweekly PHQ-9 were also collected through the same app, and the overall weekly data collection rates for them were 42.42% (3597/8480) and 41.86% (1775/4240), respectively. As we expected, the self-reported value was less diligently collected than the semiautomated walking step counts.

Figure 2. Longitudinal day-level data collection rates: (A) smartphone app; (B) smart band.



Factors Related to Data Collection Rates

In univariate analyses, comorbidities, antihormonal therapy, and targeted therapy were significantly associated with greater data collection rates with the walking app (Table 1). Table 2 shows the results of hierarchical regression analysis. In the first step including demographic variables, no factors were associated with data collection rates for smartphone apps and smart bands. In step 2, we added clinical factors to the model. Since clinical factors such as the types of therapy patients received are important to explain the data collection rate, the amount of variance explained by data collection rates with the app increased to 11.7%, with age, comorbidity, antihormonal therapy, and targeted therapy as significant predictors. Age became a significant factor in step 2 as other factors were included in the model. In step 3, the final model explained 12% of the variance in the data collection rates with the smartphone app. All variables with significance in step 2 remained significant in the final model. Young women were more compliant in the collection of PA data via the smartphone app. Subjects with any kind of disease in addition to breast cancer had a higher rate of data collection with the app (*P*=.001; Table 2). Patients on antihormonal (*P*=.009) or targeted (*P*=.011) therapies had significantly higher rates of data collection with the app (Table 2). In contrast, no demographics or clinical factors were associated with compliance with wearing the smart band (Table 3).

We also tested the correlation between the data collection rates for the use of the app and smart band. The 95% confidence interval ranged from –0.22 to 0.09, and the *P* value was 0.41, indicating there was no correlation between these data collection rates.

Table 2. Factors associated with data collection with the smartphone app in multivariate analysis.

Variable	Step 1: demographics (N=160)		Step 2: clinical factors (n=158)		Step 3: depression-related factors (n=158)	
	Beta	*P* value	Beta	*P* value	Beta	*P* value
Constant	1.09	<.001	0.71	.006	0.90	.003
Age	–0.01	.22	–0.01	.04	–0.01	.02
Marital status: single	–0.09	.15	–0.01	.10	–0.09	.14
Marital status: other	0.03	.81	0.11	.41	0.06	.70
≤High school	–0.01	.97	–0.02	.68	–0.02	.62
No occupation	–0.01	.95	0.01	.93	–0.01	.94
Comorbidity: yes	—[a]	—	0.13	.002	0.13	.001
Mastectomy (baseline: breast-conserving surgery)	—	—	–0.11	.25	–0.12	.20
Mastectomy with reconstruction (baseline: breast-conserving surgery)	—	—	–0.11	.10	–0.08	.26
Previous chemotherapy	—	—	0.04	.45	0.04	.49
Antihormonal therapy	—	—	0.14	.01	0.14	.009
Radiation therapy	—	—	–0.08	.25	–0.07	.33
Targeted therapy	—	—	0.17	.02	0.20	.01
Stage I (baseline: stage 0)	—	—	0.03	.62	0.01	.90
Stage II	—	—	–0.04	.58	–0.06	.39
Stage III	—	—	0.10	.25	0.09	.30
EQ5D-5L[b] index	—	—	0.22	.27	0.31	.14
Distress Thermometer ≥5	—	—	—	—	–0.07	.13
PHQ-9[c] ≥11	—	—	—	—	0.04	.49
Previous depression: no	—	—	—	—	–0.24	.17
Previous depression: no response	—	—	—	—	–0.18	.39
Adjusted R^2	–0.01	.68	0.12	.005	0.12	.008

[a] Not applicable.

[b] EQ5D-5L: 5-dimension 5-level health questionnaire.

[c] PHQ-9: Patient Health Questionnaire–9.

Table 3. Factors associated with data collection with smart bands in multivariate analysis.

Variable	Step 1: demographics (N=160)		Step 2: clinical factors (n=158)		Step 3: depression-related factors (n=158)	
	Beta	*P* value	Beta	*P* value	Beta	*P* value
Constant	0.65	.002	0.94	.02	0.81	.09
Age	–0.01	.81	–0.00	.62	–0.01	.58
Marital status: single	–0.06	.53	–0.07	.47	–0.05	.63
Marital status: other	–0.08	.71	–0.15	.498	–0.18	.49
≤High school	–0.01	.88	–0.01	.98	–0.01	.91
No occupation	–0.07	.26	–0.07	.26	–0.07	.28
Comorbidity: yes	—[a]	—	–0.03	.66	–0.03	.67
Mastectomy (baseline: breast-conserving surgery)	—	—	0.20	.19	0.21	.17
Mastectomy with reconstruction (baseline: breast-conserving surgery)	—	—	–0.15	.17	–0.12	.27
Previous chemotherapy	—	—	0.06	.46	0.07	.46
Antihormonal therapy	—	—	–0.03	.70	–0.02	.83
Radiation therapy	—	—	0.05	.67	0.05	.68
Targeted therapy	—	—	–0.07	.56	–0.05	.68
Stage I (baseline: stage 0)	—	—	0.14	.14	0.12	.19
Stage II	—	—	0.15	.19	0.1	.22
Stage III	—	—	0.01	>.99	–0.01	.92
EQ5D-5L[b] index	—	—	–0.29	.36	–0.21	.53
Distress Thermometer ≥5	—	—	—	—	–0.07	.29
PHQ-9[c] ≥11	—	—	—	—	0.11	.17
Previous depression: no	—	—	—	—	0.06	.84
Previous depression: no response	—	—	—	—	0.01	.97
Adjusted R^2	–0.02	.87	–0.01	.51	–0.01	.60

[a] Not applicable.

[b] EQ5D-5L: 5-dimension 5-level health questionnaire.

[c] PHQ-9: Patient Health Questionnaire–9.

Correlation Between Physical Activity and Patient Distress

In Pearson correlation tests, for smart bands, anxiety, emotion, DT, and the sum of PHQ-9 were negatively correlated with the daily average steps (Table 4). Since these psychological factors were asked to be scored high when the patients felt more depressive, the negative correlation means more daily steps relates with lower distress. The sleep disturbance answer was positively correlated with the daily average steps. Since the sleep disturbance score was asked to be scored low when patients felt unable to sleep, the positive correlation means higher daily steps relates with higher quality of sleep. For the smartphone app, all psychological domains show the same direction of correlation with the daily average steps as for the smart band, although the coefficients were not statistically significant for all domains.

Table 4. Correlation between daily steps and examined psychological domains.

Psychological domains	Smartphone app		Smart band	
	Coefficient	*P* value	Coefficient	*P* value
Anxiety (0-10)	–0.07	.38	–0.23	.008
Emotion (0-10)	–0.03	.71	–0.24	.003
Sleep (0-10)	0.08	.36	0.20	.02
Distress Thermometer (0-10)	–0.13	.11	–0.31	<.001
Patient Health Questionnaire–9 sum (0-27)	–0.12	.15	–0.30	<.001

Discussion

Principal Findings

Our results indicate that both smartphone app– and smart band–based technologies are feasible tools to collect PA data from breast cancer survivors. The overall data collection rates using both smartphone walking apps and smart bands were higher than the self-reporting rates with smartphone apps in our previous study, despite the longer follow-up period in this study [9]. Although data collection rates with smart bands rapidly decreased over time, PA data from smart bands were significantly correlated with participants' distress data. Patients who were young, with comorbidities, or receiving antihormonal or targeted therapies were more likely to be adherent to smartphone apps.

To our knowledge, this is the one of the largest prospective studies to assess PA data collection using mobile devices with a longer term of follow-up. In addition, the unique feature of this study is that we enrolled breast cancer patients who are usually recommended to be screened for depressive symptoms, and their depressive symptoms were significantly correlated with PA data from their mobile devices in our study. These results suggest that mobile devices have a significant potential as tools for distress screening in these patients with unmet needs.

Two types of PA data can be collected via smartphone apps. First, self-reported data manually capture health information. For the self-reporting of PA, PA app users self-monitored and recorded exercise more frequently over a 6-month study (2.6 [SD 0.5] days per week) than did non-app users (1.2 [SD 0.5] days per week PA self-monitoring, *P*=.001) [13]. This finding suggested that app-based PA self-reporting is better than non-app–based methods such as paper or Web questionnaires; however, the overall compliance was only 37%. Second, PA data may be obtained from mobile phone sensing via built-in or external sensors. There are a few feasibility studies on the daily collection of PA data through mobile apps or smart bands. One study reported that the number of days with nontypical wear patterns of wearable devices (Fitbit One) ranged from 5% to 9% of all observation points for 5 weeks [14]. Days were marked as nontypical to indicate that the tracker may not have been consistently worn throughout the day and/or data were not recorded, possibly due to a depleted battery. Although the number of missing days was small, the follow-up period was too short. Longer follow-up periods can result in lower compliance rates. Therefore, feasibility studies of daily collection rate of PA through smart bands should be conducted with longer follow-up periods. Our study investigated this issue with more subjects during a longer follow-up period. The results showed that the data collection rates with the smartphone app and smart bands remained higher for 6 months compared with those in other studies.

In terms of cancer survivors, previous studies focused on ePRO data, questionnaires collected via a downloadable app or Web-based portal [15,16]. Authors from a German cancer center reported an adherence rate to a smartphone app–based questionnaire of about 70% in 40 cancer patients over about 100 days [17]. Our previous studies demonstrated an overall compliance rate to an app-based questionnaire of 45% in 30 participants over a 90-day study period [9]. The overall data collection rates with mobile devices in our study were higher than the self-reporting rates in smartphone apps in previous studies.

We found that female breast cancer patients who were young, with comorbidities, or receiving antihormonal or targeted therapies tended to be more adherent to the use of smartphone apps. After enrollment, some patients were administered antihormonal or targeted therapies according to clinical practice guidelines. In terms of antihormonal therapy, patients generally have to take medications daily for 5 to 10 years [18]. Patients administered targeted therapy are usually required to visit the hospital every 3 weeks for approximately 1 year [19]. Patients with other comorbidities may also have to take other medicines or visit the clinic regularly. Thus, one possible explanation is that patients with a regular lifestyle such as taking medicines every day are more likely to check their time schedule with their smartphone, which may lead to increased adherence to smartphone apps. Breast cancer patients are more actively involved in PA collection through smartphone apps because wearing smart bands has caused discomfort to the elderly women. In interviews with patients, we were told that wearing smart bands was inconvenient, and thus, the patients were more active in PA collection through apps than bands.

With early diagnosis and improved treatment, the number of cancer survivors has increased worldwide [20]. Along with the increased cancer survival rates, more attention should be paid to the lives of cancer survivors. Recently, digital footprints, data generated passively through mobile technologies, have been introduced as tools for psychiatric research [10]. Several studies have suggested the potential for the use of data from mobile devices for new measures of mental health [21]. We believe this study can provide future direction to develop distress screening tool using mobile device–based PA data in breast cancer survivors.

Limitations

Several limitations of this study should be noted. First, long-term adherence to wearable devices has not yet been demonstrated. The follow-up period of this study was 6 months. Second, all subjects in this study were breast cancer patients. Thus, our results can only be generalized to this specific population. Third, all participants were enrolled after the completion of chemotherapy. Therefore, we do not know the feasibility of wearable devices to collect PA data during chemotherapy. Fourth, this study was conducted in a single institution, although the clinical characteristics of the patients in this study were similar to those in a nationwide study [22]. Finally, since we cannot figure out how many hours a day patients were wearing the band and carrying the app, we could not evaluate data quality in terms of time span. Future studies should evaluate the time span of each data collection rather than examining the daily collection of PA.

Conclusion

Smartphone apps or smart bands are feasible tools to collect daily PA data in breast cancer survivors. PA data from mobile devices are correlated with participants' distress data, which suggests the potential role of mobile devices in the management of distress in breast cancer survivors. Further research should focus on the interpretation and integration of PA data into clinical practice for the care of breast cancer survivors.

Acknowledgments

This study was supported by a grant from the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (HA17C0047); the Korean Breast Cancer Foundation, Seoul, Republic of Korea (KBCF-2015E003); and the Asan Institute for Life Sciences and Corporate Relations of Asan Medical Center, Seoul, Republic of Korea.

Authors' Contributions

IYC, JWL, YHM, and SC designed the study and provided input throughout the study. YRP, HC, HJP, and ML collected the data. SY, SC, BHS, and SHA provided clinical expertise throughout the study. MJ, DC, SBL, and IYC analyzed portions of the data. IYC and MJ wrote the manuscript along with contributions from all the authors. All authors read and approved the final manuscript.

Conflicts of Interest

HC is chief executive officer of Swallaby Co, Ltd, Seoul, Republic of Korea.

References

1. Mitchell AJ, Ferguson DW, Gill J, Paul J, Symonds P. Depression and anxiety in long-term cancer survivors compared with spouses and healthy controls: a systematic review and meta-analysis. Lancet Oncol 2013 Jul;14(8):721-732. [doi: 10.1016/S1470-2045(13)70244-4] [Medline: 23759376]
2. Zainal NZ, Nik-Jaafar NR, Baharudin A, Sabki ZA, Ng CG. Prevalence of depression in breast cancer survivors: a systematic review of observational studies. Asian Pac J Cancer Prev 2013;14(4):2649-2656 [FREE Full text] [Medline: 23725190]
3. Natioanl Comprehensive Cancer Network. 2018. NCCN guidelines for supportive care: distress management URL: https://www.nccn.org/professionals/physician_gls/pdf/distress.pdf [accessed 2019-10-05]
4. Andersen BL, DeRubeis RJ, Berman BS, Gruman J, Champion VL, Massie MJ, American Society of Clinical Oncology. Screening, assessment, and care of anxiety and depressive symptoms in adults with cancer: an American Society of Clinical Oncology guideline adaptation. J Clin Oncol 2014 May 20;32(15):1605-1619 [FREE Full text] [doi: 10.1200/JCO.2013.52.4611] [Medline: 24733793]
5. Newell S, Sanson-Fisher RW, Girgis A, Bonaventura A. How well do medical oncologists' perceptions reflect their patients' reported physical and psychosocial problems? Data from a survey of five oncologists. Cancer 1998 Oct 15;83(8):1640-1651. [Medline: 9781960]
6. Kim J, Lim S, Min YH, Shin Y, Lee B, Sohn G, et al. Depression screening using daily mental-health ratings from a smartphone application for breast cancer patients. J Med Internet Res 2016 Aug 04;18(8):e216 [FREE Full text] [doi: 10.2196/jmir.5598] [Medline: 27492880]
7. Basch E, Artz D, Dulko D, Scher K, Sabbatini P, Hensley M, et al. Patient online self-reporting of toxicity symptoms during chemotherapy. J Clin Oncol 2005 May 20;23(15):3552-3561 [FREE Full text] [doi: 10.1200/JCO.2005.04.275] [Medline: 15908666]
8. 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. J Clin Oncol 2007 Dec 01;25(34):5374-5380. [doi: 10.1200/JCO.2007.11.2243] [Medline: 18048818]
9. Min YH, Lee JW, Shin Y, Jo M, Sohn G, Lee J, et al. Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. J Med Internet Res 2014;16(5):e135 [FREE Full text] [doi: 10.2196/jmir.3421] [Medline: 24860070]

10. Bidargaddi N, Musiat P, Makinen V, Ermes M, Schrader G, Licinio J. Digital footprints: facilitating large-scale environmental psychiatric research in naturalistic settings through data from everyday technologies. Mol Psychiatry 2017 Dec;22(2):164-169 [FREE Full text] [doi: 10.1038/mp.2016.224] [Medline: 27922603]
11. Sano A, Yu AZ, McHill AW, Phillips AJK, Taylor S, Jaques N, et al. Prediction of Happy-Sad mood from daily behaviors and previous sleep history. Conf Proc IEEE Eng Med Biol Soc 2015;2015:6796-6799 [FREE Full text] [doi: 10.1109/EMBC.2015.7319954] [Medline: 26737854]
12. Wen CKF, 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]
13. 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 1;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
14. Wang JB, Cadmus-Bertram LA, Natarajan L, White MM, Madanat H, Nichols JF, et al. Wearable sensor/device (Fitbit One) and SMS text-messaging prompts to increase physical activity in overweight and obese adults: a randomized controlled trial. Telemed J E Health 2015 Oct;21(10):782-792. [doi: 10.1089/tmj.2014.0176] [Medline: 26431257]
15. Coons SJ, Eremenco S, Lundy JJ, O'Donohoe P, O'Gorman H, Malizia W. Capturing patient-reported outcome (PRO) data electronically: the past, present, and promise of ePRO measurement in clinical trials. Patient 2015 Aug;8(4):301-309 [FREE Full text] [doi: 10.1007/s40271-014-0090-z] [Medline: 25300613]
16. Gwaltney C, Coons SJ, O’Donohoe P, O’Gorman H, Denomey M, Howry C, et al. “Bring Your Own Device” (BYOD). Drug Inf J 2015 Oct 08;49(6):783-791. [doi: 10.1177/2168479015609104]
17. Benze G, Nauck F, Alt-Epping B, Gianni G, Bauknecht T, Ettl J, et al. PROutine: a feasibility study assessing surveillance of electronic patient reported outcomes and adherence via smartphone app in advanced cancer. Ann Palliat Med 2017 Oct 26:1 [FREE Full text] [doi: 10.21037/apm.2017.07.05] [Medline: 29156896]
18. Burstein HJ, Lacchetti C, Anderson H, Buchholz TA, Davidson NE, Gelmon KE, et al. Adjuvant endocrine therapy for women with hormone receptor-positive breast cancer: American Society of Clinical Oncology clinical practice guideline update on ovarian suppression. J Clin Oncol 2016 Dec 10;34(14):1689-1701. [doi: 10.1200/JCO.2015.65.9573] [Medline: 26884586]
19. Giordano SH, Temin S, Davidson NE. Systemic therapy for patients with advanced human epidermal growth factor receptor 2-positive breast cancer: ASCO clinical practice guideline update summary. J Oncol Pract 2018 Aug;14(8):501-504. [doi: 10.1200/JOP.18.00290] [Medline: 29989839]
20. Noone A, Howlader N, Krapcho M, Miller D, Brest A, Yu M. SEER cancer statistics review, 1975-2015. Bethesda: National Cancer Institute; 2018. URL: https://seer.cancer.gov/csr/1975_2015/ [accessed 2019-08-24]
21. Glenn T, Monteith S. New measures of mental state and behavior based on data collected from sensors, smartphones, and the Internet. Curr Psychiatry Rep 2014 Dec;16(12):523. [doi: 10.1007/s11920-014-0523-3] [Medline: 25308392]
22. Park EH, Min SY, Kim Z, Yoon CS, Jung K, Nam SJ, Korean Breast Cancer Society. Basic facts of breast cancer in Korea in 2014: the 10-Year overall survival progress. J Breast Cancer 2017 Mar;20(1):1-11 [FREE Full text] [doi: 10.4048/jbc.2017.20.1.1] [Medline: 28382089]

Abbreviations

DT: Distress Thermometer
ePRO: electronic patient-reported outcome
EQ5D-5L: 5-dimension 5-level health questionnaire
PA: physical activity
PHQ-9: Patient Health Questionnaire–9
PRO: patient-reported outcome

Edited by G Eysenbach; submitted 24.01.19; peer-reviewed by S Lim, M Boukhechba; comments to author 31.05.19; revised version received 22.07.19; accepted 21.08.19; published 06.09.19.

Please cite as:
Chung IY, Jung M, Lee SB, Lee JW, Park YR, Cho D, Chung H, Youn S, Min YH, Park HJ, Lee M, Chung S, Son BH, Ahn SH
An Assessment of Physical Activity Data Collected via a Smartphone App and a Smart Band in Breast Cancer Survivors: Observational Study
J Med Internet Res 2019;21(9):e13463
URL: https://www.jmir.org/2019/9/e13463
doi:10.2196/13463
PMID:31493319

Original Paper

Predicting Inpatient Aggression in Forensic Services Using Remote Monitoring Technology: Qualitative Study of Staff Perspectives

Ben Greer[1], BSc; Katie Newbery[1,2], BSc, MSc; Matteo Cella[1,2], BSc, PhD, DClinPsy; Til Wykes[1,2], MPhil, DPhil

[1]Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[2]South London and Maudsley NHS Foundation Trust, London, United Kingdom

Corresponding Author:
Ben Greer, BSc
Department of Psychology
Institute of Psychiatry, Psychology and Neuroscience
King's College London
Henry Wellcome Building
16 De Crespigny Park
London, SE5 8AF
United Kingdom
Phone: 44 02078480423
Email: ben.greer@kcl.ac.uk

Abstract

Background: Monitoring risk of imminent aggression in inpatient forensic mental health services could be supported by passive remote monitoring technology, but staff attitudes toward the relevance and likelihood of engagement with this technology are unknown.

Objective: This study aimed to explore staff views, specifically potential benefits and implementation barriers, on using this technology for monitoring risk of inpatient aggression.

Methods: We conducted semistructured focus groups with nurses in an inpatient forensic mental health service. We used thematic analysis with two independent raters to identify themes and subthemes related to staff attitudes toward passive remote monitoring. We subsequently checked with members to ensure the validity of the themes identified by the raters.

Results: From January to March 2019, a total of 25 nurses took part in five focus groups. We identified five main themes, one of which concerned the potential benefits that passive remote monitoring could provide for monitoring risk of aggression. Staff suggested it could provide an early warning of impending aggression and enable support to be provided earlier. The remaining themes concerned implementation barriers, including risks to the users' physical and mental well-being; data security concerns and potential access by third parties; the negative impact of a constant stream of real-time data on staff workload; and design characteristics and user awareness of the benefits of passive remote monitoring.

Conclusions: Passive remote monitoring technology could support existing methods of monitoring inpatient aggression risk, but multiple barriers to implementation exist. Empirical research is required to investigate whether these potential benefits can be realized, and to identify ways of addressing these barriers to ensure acceptability and user engagement.

(J Med Internet Res 2019;21(9):e15620) doi:10.2196/15620

KEYWORDS

telemedicine; remote sensing technology; wearable sensors; aggression; risk assessment

Introduction

Background

Infrequent structured risk assessments for inpatient aggression cannot detect rapid, momentary changes in individual risk factors, limiting their ability to predict imminent aggression. For example, purportedly changeable (dynamic) risk factors have shown nonsignificant changes over service users' period of treatment [1]. The limited frequency of these assessments may partially explain why they often show low to moderate predictive accuracy for aggression [2,3]. Acquiring a continuous, real-time measure of how the severity of risk factors changes could provide a more accurate understanding of how these fluctuations relate to the risk of aggressive outcomes [4].

Developments in passive remote monitoring technology (ie, wearable sensors) mean that factors previously associated with aggression cross-sectionally can now be monitored continuously, in real time. For example, many devices monitor indicators of autonomic nervous system activity such as electrodermal activity and heart rate variability, previously associated with aggression in laboratory studies [5-7]. Emerging evidence indicates that passive remote monitoring technology can be used to identify changes in these parameters from 1 to 30 minutes before the observable act of aggression [8,9], and with greater temporal precision than existing risk assessments alone. Monitoring these parameters in real time could identify a reliable psychophysiological signature of when aggression may be more likely, enabling support to be offered before this behavior escalates.

While passive remote monitoring technology has the potential to support risk assessments for aggression [10-13], to our knowledge, the views of frontline staff have not been investigated. Numerous implementation barriers for novel digital health care systems often result in low rates of adoption and adherence [14]. For example, staff have reported a potentially beneficial role for mobile phone–based support for individuals with psychosis, but also highlighted concerns such as infrastructure requirements, data privacy, and the level of digital literacy and skills required of both service users and staff [15,16]. The views of end users, including staff, are therefore critical to understanding the desire and need for digital systems to monitor risk of aggression and the barriers that may be faced.

Objectives

Inpatient aggression is a barrier to rehabilitation that negatively affects service user and staff well-being [17], and passive remote monitoring technology could facilitate appropriate support and de-escalation. However, not taking a user-centered approach or involving staff in the development and introduction of this technology may mean it is not relevant or acceptable for clinical practice [18]. Therefore, the aim of this study was to explore the attitudes of staff toward passive remote monitoring technology for risk of aggression in inpatient forensic mental health services, with a focus on the potential benefits that this technology could provide and barriers to implementation.

Methods

Design

This was an exploratory qualitative study using focus groups following a topic guide. We obtained ethical approval from the Yorkshire & The Humber-Bradford Leeds Proportionate Review Sub-Committee, Jarrow, UK (18/YH/0221) and King’s College London Psychiatry, Nursing and Midwifery Research Ethics Panel, London, UK (LRS-17/18-6715).

Participants

Participants were staff in a medium-secure forensic mental health service in South London, UK, covering a diverse geographical area including areas of high poverty and urban deprivation. Staff were eligible to participate if their role involved direct contact with service users. Nonclinical staff were not eligible, as the aim of this study was to understand attitudes toward passive remote monitoring technology in a clinical context. We conducted recruitment and analysis concurrently, and recruitment stopped when we achieved data saturation, the point at which focus groups stopped yielding new themes [19].

Focus Group Topic Guide

The topic guide was based on previous reports [15,16] and included issues related to perceived utility, safety and security, and data connectivity requirements (see Multimedia Appendix 1). The topic guide was informed by consultation with 2 service user–caregiver advisory groups, by a systematic review of the barriers to and facilitators of remote monitoring for health care [20], and by incorporating suggestions from senior management staff in the recruitment site during the setup phase of this study, to ensure that we covered topics relevant to the forensic setting.

Procedure

We approached ward managers for permission to recruit from their ward, and we conducted 5 focus groups in a private room on the participants’ respective wards during staff handover meetings (2 groups) or at a time convenient for participants (3 groups). Discussions were audio recorded and transcribed verbatim (by BG), with personally identifiable content omitted. Participants were provided £10 (about US $12) in cash after the focus groups in recognition of their time. We conducted member-checking focus groups for the primary themes with the same participants so they could suggest any amendments they felt were appropriate [21]. We informed participants that the study was part of a larger project investigating wearable sensors for monitoring the risk of aggression through physical signals. To provide a context for the discussions, we told participants that the focus groups were the first in a series of studies that aimed to investigate whether real-time monitoring of psychophysiological signals could assist in the earlier detection of an increasing risk of inpatient aggression. We presented 2 remote monitoring devices to illustrate the devices. One device (E4; Empatica Srl, Milan, Italy) is worn around the wrist, and the other (Everion; Biovotion Ltd, Zurich, Switzerland) is worn around the upper arm. Although participants were familiar with commercially available wearable devices, the 2 devices presented were novel to them.

Thematic Analysis

NVivo 12 software (QSR International) facilitated thematic analysis by 2 independent raters (BG and KN). Both read and reread the transcripts, producing a list of initial codes, and then independently collated the codes into a list of candidate themes and subthemes. Both raters’ initial identification of individual codes and overall themes were compared, resulting in an initial agreement rate of 59% and 72% for individual codes and overall themes, respectively. Where there were discrepancies (eg, 1 rater identifying a code or theme that the other had not), both raters discussed these ratings until they reached a consensus, and themes were revised into their final structure.

Results

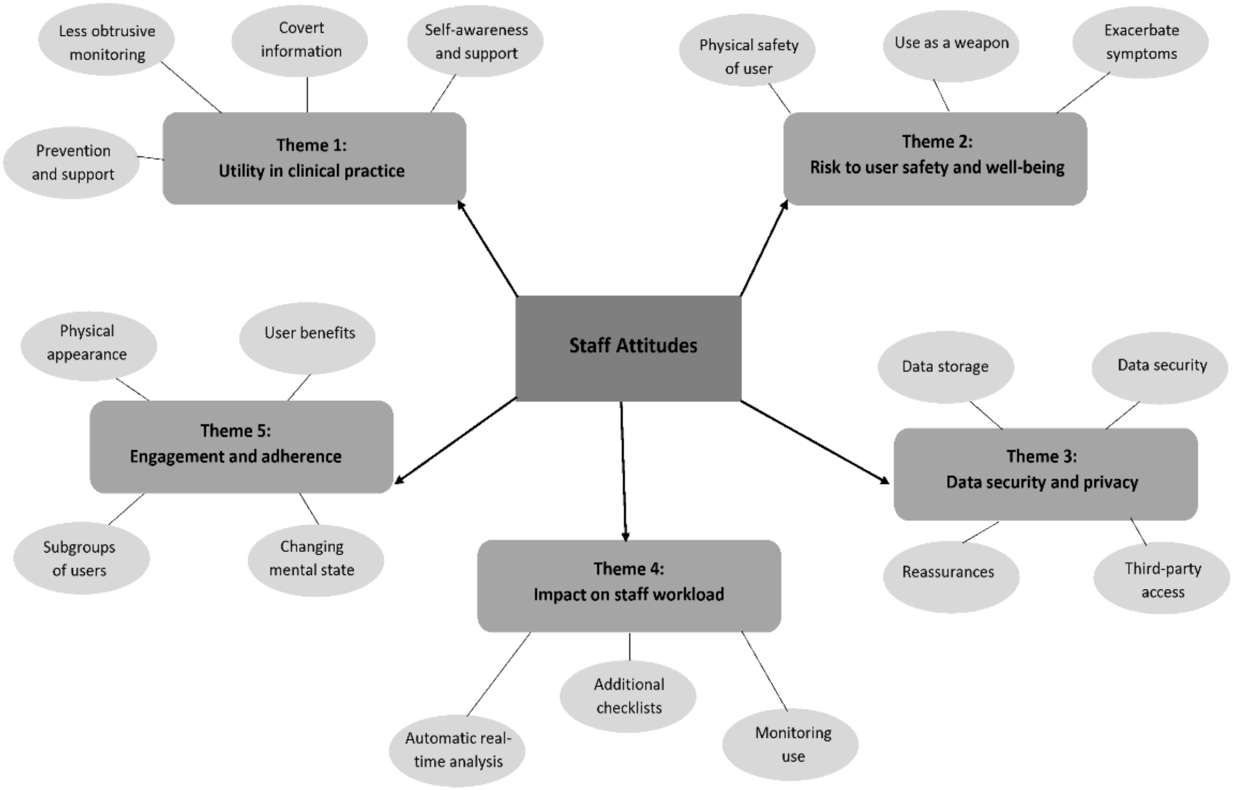
Participant Demographics

From January to March 2019, we approached 43 staff, and 25 of them took part in the focus groups. Those who declined did so because of the focus group timing (n=9), or they were required to remain on the ward to maintain minimum staff numbers and carry out clinical duties (n=4); 5 did not specify a reason. A total of 18 participants were also available to take part in the member-checking focus groups. Table 1 presents participants' demographics. We identified 5 primary themes, which we discuss below in addition to subthemes. Figure 1 provides a visual overview of these themes and subthemes.

Table 1. Participants' demographic characteristics.

Characteristics	Focus group					Total (N=25)	Member-checking groups (n=18)
	1 (n=6)	2 (n=6)	3 (n=4)	4 (n=5)	5 (n=4)		
Age (years)							
Mean (SD)	37.8 (12.4)	39.5 (11.3)	37 (10.8)	55.4 (7.4)	44.5 (5.1)	42.7 (11.6)	44.4 (12.8)
Range	22-54	25-57	25-51	44-64	41-52	22-64	22-64
Sex, n							
Female	3	6	3	3	1	16	12
Male	3	0	1	2	3	9	6
Ethnicity, n							
Black African	4	0	4	4	4	16	15
Black Caribbean	1	3	0	1	0	5	1
White British	1	3	0	0	0	4	2
Job title, n							
Staff nurse	5	6	1	4	4	20	16
Student nurse	1	0	2	0	0	3	1
Ward manager	0	0	1	1	0	2	1
Highest educational attainment, n							
Higher-level qualification (eg, university degree, professional qualification)	5	6	2	5	4	22	17
Secondary (A-level equivalent)	1	0	2	0	0	3	1
Time in post, mean (SD)	5 years, 5 months (3 years, 5 months)	5 years (4 years, 9 months)	4 years (5 years, 8 months)	10 years, 9 months (4 years, 8 months)	4 years, 7 months (2 years, 1 month)	6 years, 6 months (3 years, 5 months)	6 years, 5 months (4 years, 9 months)

Figure 1. Themes and subthemes identified through thematic analysis.



Theme 1: Utility in Clinical Practice

In every focus group, participants identified numerous ways in which using these devices could augment their practices. One area of discussion related to the devices' capacity for *prevention and support* to be offered to users:

> *Yeah it would be helpful, it's like an early-warning sign so, when you know that they are coming to be anxious, you find a way of intervening before it escalates* [Participant 023, nurse]

> *Because sometimes by the time they express it, it means it's, it's already you know, so if we can see ahead of time and monitor it, I think it's good.* [Participant 009, nurse]

Participants also suggested that these devices could facilitate a *less obtrusive monitoring* approach, enabling assessments without the need for staff to be in physical contact with users:

> *Because if a patient is wearing this device even if they are in their bedroom, and they're out of sight of the staff, with the device working you should be able to tell that maybe something has gone wrong...if you just see them physically, they might be in their room they're anxious, they're agitated without you seeing them, you won't be able to tell.* [Participant 024, nurse]

Reflecting on the seemingly unprovoked nature of some aggressive incidents, participants felt that these devices could provide staff with *covert information* that may not otherwise be expressed by the users or observable by staff:

> *...because we did not see it we think it's unprovoked…but with these devices maybe we will know that there's something happening...before the incident, later maybe attacking somebody or something.* [Participant 025, nurse]

> *...not all our patients will be able to say "oh well I feel agitated" or be able to come out and say it, but within themselves all the physical, you know, changes are taking place so I think it's good, it will help us to see the covert, you know, things that are not outward that the patients cannot express.* [Participant 009, nurse]

As a result of identifying this covert information, participants felt that this could be used to foster *self-awareness and support* among users, and augment an anger management program offered to users:

> *...we have taught some of them [service users] who did anger management, but some of them are still not withdrawing when they get to the point of.* [Participant 017, ward manager]

> *...[to the] point of anger exactly.* [Participant 015, nurse]

> *...this would be a different way of reminding them, maybe this would be a reminder. Because, um, if somebody has done anger management and he knows that by the time he starts breathing heavily, or by the time he feels he's getting, um, a bit sweaty or agitated and getting wound up...they should disengage. And they have not been doing that because they still don't have the capacity to do that because they don't, um,*

> *how can I say, they can't get themselves to...take themselves away from aggression. So, what this would do is to then remind them that this is what they need to do, for some of them who have done anger management.* [Participant 017, nurse]

Theme 2: Risk to User Safety and Well-Being

Participants in 4 focus groups discussed the impact of the devices in relation to the *physical safety of the user*. One focus group cited the risk of the device being used as a ligature as a concern, due to the elastic armband of 1 of the devices:

> *...how far does it stretch, can you put it round your neck?* [Participant 008, nurse]
>
> *Oh yeah, you could I reckon, you could stretch it.* [Participant 011, nurse]
>
> *Well that might be an issue, you know, ligatures.* [Participant 008, nurse]

The possibility that the devices could be *used as a weapon* was another risk that participants raised, with 1 focus group discussing the potential implications of 1 of the devices having an elastic armband:

> *...should be something that they cannot use as a weapon, like, there shouldn't be any metal or something that they can use to self-harm.* [Participant 025, nurse]
>
> *You could use this [referring to device] as a weapon like a slingshot.* [Participant 011, nurse]

Participants suggested that users' mental well-being would need to be considered in addition to their physical safety when wearing the devices. Specifically, participants raised concerns that continuous monitoring could *exacerbate symptoms* of paranoia among some users:

> *They might think that you are monitoring, that you are controlling their mind, controlling their mental state, all of this, so it might make more paranoia.* [Participant 003, nurse]
>
> *When you give this to a paranoid patient they will think you are monitoring them. It will be so difficult to explain it to them to understand it that this is what you're monitoring....This paranoia could also lead to them not even wearing this.* [Participant 015, nurse]

Theme 3: Data Security and Privacy

Across 3 focus groups, participants discussed the measures in place to ensure that data collected by the devices would be kept secure and confidential. Participants wanted to know in advance specific details concerning *data storage*: where the information collected by the devices would be stored and who would have access to it; *data security*: what protections would be in place to keep the data private and confidential; and *access by third parties*: whether data would be shared with other individuals or companies:

> *...you know I'm gonna need to know, um, what [inaudible] they do, how, even if you say the data is gonna be stored, how secure is the storage, can it be hacked, you know, cos this is like really private, um, information.* [Participant 002, nurse]
>
> *Yeah I think we'd want reassurance, wouldn't we, that the information we give is protected confidential and it'd be the same for the patients, know how it's going to be used and...just make sure it's anonymized the data.* [Participant 008, nurse]
>
> *You know is it gonna be sold to a third-party like we've seen with social media now where obviously data protection is like a lot* [Participant 002, nurse]

Because of these concerns, participants felt that *reassuring users* that their data would be kept secure would be necessary, particularly due to the risk of exacerbating symptoms of paranoia discussed in theme 1:

> *Just reassuring them that their data will not go anywhere, it's just for the ward, because some of them will be paranoid.* [Participant 025, nurse]

One participant made a practical suggestion as to how users could be reassured, drawing parallels with a ward policy whereby staff model appropriate eating during mealtimes. They suggested that staff members wearing the devices themselves could reassure users:

> *...each time they are eating you need to have 2 staff that will come and model and sit with them...so I'm just thinking that maybe they can try to [inaudible]. If they are wearing it that they see staff wearing it as well, they won't be thinking about confidentiality, maybe they are trying to take their information or do something else.* [Participant 024, nurse]

Theme 4: Impact on Staff Workload

While identifying ways in which these devices could be implemented in clinical practice, participants in 3 focus groups also highlighted that this should not increase staff members' already high workload. Participants stressed the need for *automatic real-time analysis* of the data collected by the device, to ensure that information can be acted on without additional burden on staff time:

> *Yeah I mean that's the only way I can think it'd be useful, without that real-time information, we're gonna have to take the watch and then upload the data and see what's going on.* [Participant 002, nurse]
>
> *Yeah [if] it's automatic, and we don't have to put in a lot more to get the data and to analyze the data, then it will be good, yeah. But if we have to put in a lot more to measuring and analyzing the data and doing deductions for ourselves, that means additional work to do.* [Participant 009, nurse]

Participants also questioned whether *additional checklists* would accompany the devices, increasing the level of ongoing input required from staff:

> *So does this come with a pack or a checklist or something that you've got to fill it out every day during the monitoring, or will you really only attend to it when you, it gives you any signals or something that there are any changes? How is it done, I'm just*

> *thinking if it's something that's supposed to be monitored every now and then and every day it means additional work, isn't it, you feel, more boxes to tick.* [Participant 009, nurse]

Responsibility for ensuring that users are wearing the devices was also discussed, with participants feeling that it would fall to staff to spend additional time *monitoring use*, therefore taking time away from their other duties:

> *Because it's going to be more [inaudible] on staff now. Now they are wearing it they are gonna say, oh, gonna spend a couple of minutes encouraging them to put it on or go put it on, so it's going to take valuable time out of your working day, so, it's going to be time consuming in a way.* [Participant 001, nurse]

Theme 5: Engagement and Adherence

All focus groups discussed the numerous factors that may affect the likelihood of users engaging with the devices. The *physical appearance* of the device, including overall size and possible resemblance to a tracking device, was 1 factor:

> *And you know this one [referring to device] is so conspicuous it looks so much like a tracking device, you know* [Participant 009, nurse]
>
> *...even if they have the reservations about, "oh we don't want to be monitored" and things like that, if they see something that looks a bit stylish they might be more prone to wear it.* [Participant 018, nurse]

Participants also suggested that users would be more likely to engage with the devices if there was a clear *benefit to the user*:

> *If there's nothing for them they won't take it.* [Participant 017, nurse]

One benefit that participants felt would appeal to users was whether wearing the devices would positively affect their leave status:

> *But then I'm thinking it's one thing, how is it gonna directly benefit them, like, "what are you gonna tell if I'm a patient, and you wanna give me this I need to know, like...is it gonna make my leave better* [Participant 002, nurse]

A total of 4 focus groups discussed the impact that *changing mental state* would have on users' engagement, suggesting that *subgroups of users* may be most likely to engage:

> *...that would be a problem, getting them to volunteer for it and, um, making sure they understand completely, cos some people are more paranoid on days...than other days so it could be they're fine for 5 days then the sixth day they're really paranoid* [Participant 008, nurse]
>
> *The most settled patients there, they will cooperate, some of them are, so most of them that are eager to go out they'll cooperate, but this ones, like, very paranoid like you said, you will have a tough time.* [Participant 001, nurse]

Discussion

Principal Findings

To our knowledge, this is first study to investigate the attitudes of frontline staff members in inpatient forensic mental health toward passive remote monitoring for risk of aggression. Staff suggested this technology could benefit their assessments, identifying changes in risk factors that would otherwise not be identified. The real-time stream of information provided by these devices could facilitate targeted support before behavior escalated into aggression. However, staff also raised numerous implementation barriers, including the physical safety of the user and security of their personal data, negative impact on staff workload, and engagement barriers.

Advantages of Passive Remote Monitoring

Participants suggested that the covert information monitored by these devices may account for the seemingly unprovoked nature of some aggressive incidents. Current risk assessments are rated by staff on the basis of observable characteristics (eg, irritability and following instructions [22]); therefore, passive remote monitoring could provide a more complete clinical picture, consistent with previously hypothesized benefits of digital technology for managing aggression [13]. These additional objective data may also circumvent limitations of structured risk assessments such as rater bias [23] and incomplete or inaccurate ratings [24].

Participants suggested that the information provided through passive remote monitoring technology could equip them with prior knowledge of when users may be experiencing difficulties, thereby facilitating appropriate prevention and support. Participants discussed this in relation to staff-initiated de-escalation procedures, but also identified an opportunity to foster self-awareness and support to enable users themselves to de-escalate. This is a novel suggestion for the role of passive remote monitoring in managing aggression, with previous literature typically discussing only how this technology could enable staff to manage risk [13]. Passive remote monitoring could therefore enhance users' ability to identify and manage their unique risk factors for aggression, consistent with the UK national guideline's [25] calls for greater emphasis on individual self-regulation [26].

Enabling users to be monitored without the need for in-person observation was considered less obtrusive than current observation practices. Enhanced observations (eg, eyesight and arm's-length observations) are experienced negatively by both service users and staff [27]; therefore, passive remote monitoring may enable monitoring with fewer physical restrictions. This would need to be balanced with the accuracy and range of clinical observations that can be made with passive remote monitoring alone. For example, peer interactions and negative attitudes are relevant risk factors for aggression [28] but cannot be assessed through actigraphy or biosensors alone, highlighting the need for multiple sources of observation data. Overreliance on technology could also limit the opportunity for physical service user–staff interactions and dialogue, an integral component of therapeutic relationships [29]. As mentioned in 1 focus group, however, providing staff with feedback on changes in psychophysiological parameters of individuals in their care could also facilitate dialogue and staff–service user interaction ("I could see that it could be useful, because, um, it could just be a point of engagement for staff" [Participant 017, nurse]).

The suggestion that passive remote monitoring would be suitable only for subgroups of users is consistent with a previous evaluation of a passive remote monitoring system (a global positioning system [GPS] tracker) in a forensic mental health service [30]. The system in that study was primarily used for subgroups at the early stages of their leave period or during specific periods of transition. Participants felt that these subgroups would be based on users' current mental state and paranoid ideation. Passive remote monitoring may not be suitable for everyone and reflects the need for a personalized approach, which balances the potential benefits to the user (eg, improved understanding of changes in risk state) and challenges (eg, difficulties with engagement).

While highlighting the potential benefits of passive remote monitoring technology, no participants suggested that technology should replace the practice of staff-completed risk assessments. This is consistent with previous research reporting universal agreement among staff that digital health care technologies should be an adjunct to traditional care rather than a replacement, as replacement could be detrimental to user well-being and therapeutic relationships [16]. Future use of passive remote monitoring should therefore be considered as a component of a blended approach that complements, but does not wholly replace, staff-completed structured risk assessments.

The issues discussed above are hypothesized benefits, and while there is potential for passive remote monitoring technology to support risk management for aggression, this needs to be supported by high-quality empirical evidence. Key issues that need to be addressed include the feasibility and acceptability of this technology for end users, whether a reliable psychophysiological signature for aggression exists, and the accuracy of the data provided by this technology, including the ability to correctly identify changes related to aggression and to rule out those that are unrelated.

Implementation Barriers

Participants identified numerous issues that are likely to affect successful implementation of passive remote monitoring systems. The physical safety concerns raised by participants, relating to ligature risk and use as a weapon, appeared to be linked to a specific design characteristic (elastic armband) of a device presented during the focus groups. This highlights the importance of considering the physical design of passive remote monitoring devices intended for use in inpatient services, where physical safety and risk of self-harm are management priorities. Concerns were raised that continuous passive remote monitoring may exacerbate symptoms of paranoia; therefore, establishing trust with users beforehand is likely to be integral to successful implementation. While previous research of passive remote monitoring in the community indicates the acceptability of passive remote monitoring for individuals with psychosis [31-33], it may be a pertinent issue for those involuntarily admitted to inpatient services and experiencing loss of control and restrictive practices [34].

Consistent with previous research among staff [35,36] and service users [37], participants expressed data privacy concerns. Rather than voicing a general unspecified concern, participants specified 3 areas of assurance they would require to be comfortable with passive remote monitoring, relating to data storage, security, and accessibility by third parties. Addressing these concerns in the long term will require digital health companies to be transparent about the procedures in place for handling user data, and to ensure that users have access to this information. In the shorter term, participants reported that staff could play an important role in providing reassurance to users. The suggestion that staff could lead by example by trialing the devices themselves reflects the role of staff as positive role models in inpatient services [38] and is a practical approach to alleviating user concerns.

Participants expressed concern that incorporating passive remote monitoring into their working practice might negatively affect their existing workload. Like Bucci et al [16], participants were concerned with the potential burden of handling and analyzing large volumes of real-time data, and emphasized that these devices need to be complemented by automatic real-time analysis. While this would address the process of analyzing the raw data into an actionable format, it is possible that a constant stream of processed data could still prove overwhelming. It will therefore be important to establish an appropriate format for presenting the data, balancing the frequency of data, level of detail, and staff capacity to act on this information. Participants also highlighted practical considerations such as the introduction of additional checklists with passive remote monitoring and questioned where responsibility for monitoring use would lie. Future use of passive remote monitoring devices will therefore need to balance potential clinical benefits with practical implementation issues, to ensure that they support and do not hinder clinical care. Machine learning algorithms to process data in real time and present it in a user-friendly and actionable format would be an option, and participants discussed presenting the data as a visual display in the ward's central nursing office. Embedding a dedicated technology specialist within the clinical team could also be considered [39]. While not raised in the focus groups, there are likely to be numerous practical implementation issues to using remote monitoring technology, such as the financial cost of acquiring and maintaining the technology, and any ongoing training requirements for staff.

Achieving the hypothesized benefits of passive remote monitoring devices requires adequate levels of user engagement. The influence of a device's physical appearance on engagement highlights the importance of involving users in the choice of devices, as individual preferences may vary. User involvement in the design of these devices could also be an effective way of ensuring that they are considered acceptable. Engaging in passive remote monitoring also needs to have discernible benefits to the user, which in the context of managing risk of aggression could include less restrictive management practices. Communicating these potential benefits could therefore have a positive effect on engagement. User engagement will also entail staff enthusiasm to work with novel digital technologies.

Barriers to organizational change within mental health services include poor job satisfaction among staff, burnout, and lower levels of experience [40,41], and these may also be barriers to the successful introduction of passive remote monitoring technology.

Strengths and Limitations

This study took place in a hospital where passive remote monitoring technology, GPS tracking devices, has been in use for several years to monitor leave [30,42]. Participants' responses may have been influenced by this prior experience and so may not necessarily reflect the views of those unfamiliar with these systems. However, some familiarity might also have facilitated discussion, with the issues and recommendations raised reflecting participants' applied experience of passive remote monitoring. Because this is a medium-secure forensic mental health service, the themes identified in this study may not necessarily generalize to forensic services of higher levels of security (eg, where more stringent policies for patient access to digital equipment exist), or to non–forensic mental health services, where inpatient aggression also occurs.

Senior nursing and managerial staff, whose views might have provided greater organizational context, were present in smaller numbers and so we did not successfully capture their views. Including only members of the nursing staff may have limited the emergence of new themes. For example, the technical infrastructure requirements for remote monitoring technology could be clarified by seeking the views of support service staff (eg, technicians and engineers). The time for focus groups was restricted, and this might also have limited the emergence of new themes. However, the replication of themes across the different groups makes this unlikely.

Conclusion

Passive remote monitoring technology offers potential benefits to monitoring risk of aggression in inpatient forensic mental health services. However, novel digital systems in mental health have generated substantial hype [43], and these potential benefits have yet to be realized through high-quality empirical research. Future research should therefore investigate whether passive remote monitoring is able to achieve the benefits suggested by staff and reliably identify increased risk of aggression. Specifically, future research should investigate the areas of concern identified by staff in this study to determine acceptability and feasibility of passive remote monitoring, such as whether subgroups based on diagnosis are more likely to engage, and the impact of real-time monitoring on staff workload. Determining whether a reliable psychophysiological signature of imminent aggression exists is also critical, and robust methods for analyzing these data, such as machine learning, will need to be developed and evaluated to make this determination.

Acknowledgments

This paper presents independent research funded by the UK National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. TW would also like to acknowledge the support of her NIHR Senior Investigator award. The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR, or the UK Department of Health and Social Care. This study and its written materials were reviewed by a team with personal experience of mental health problems and their caregivers who have been specially trained to advise on research proposals and documentation through the Feasibility and Acceptability Support Team for Researchers (FAST-R) and Service User Advisory Group (SUAG), a free, confidential service in England provided by the National Institute for Health Research Maudsley Biomedical Research Centre at King's College London and South London and Maudsley NHS Foundation Trust. We also thank Richard Stott, who provided support in the development phase of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Focus group topic guide.
[PDF File (Adobe PDF File)54 KB - jmir_v21i9e15620_app1.pdf]

References

1. Mastromanno B, Brookstein DM, Ogloff JRP, Campbell R, Chu CM, Daffern M. Assessing change in dynamic risk factors in forensic psychiatric inpatients: relationship with psychopathy and recidivism. J Forensic Psychiatry Psychol 2017 Oct 11;29(2):323-336. [doi: 10.1080/14789949.2017.1377277]
2. Fazel S, Singh JP, Doll H, Grann M. Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis. BMJ 2012 Jul 24;345:e4692 [FREE Full text] [doi: 10.1136/bmj.e4692] [Medline: 22833604]
3. Chan MKY, Bhatti H, Meader N, Stockton S, Evans J, O'Connor RC, et al. Predicting suicide following self-harm: systematic review of risk factors and risk scales. Br J Psychiatry 2016 Dec;209(4):277-283. [doi: 10.1192/bjp.bp.115.170050] [Medline: 27340111]

4. Coid JW, Kallis C, Doyle M, Shaw J, Ullrich S. Shifts in positive and negative psychotic symptoms and anger: effects on violence. Psychol Med 2018 Dec;48(14):2428-2438. [doi: 10.1017/S0033291718000077] [Medline: 29482669]
5. Armstrong T, Wells J, Boisvert DL, Lewis R, Cooke EM, Woeckener M, et al. Skin conductance, heart rate and aggressive behavior type. Biol Psychol 2019 Dec;141:44-51. [doi: 10.1016/j.biopsycho.2018.12.012] [Medline: 30584895]
6. Murray-Close D, Holterman LA, Breslend NL, Sullivan A. Psychophysiology of proactive and reactive relational aggression. Biol Psychol 2017 Dec;130:77-85. [doi: 10.1016/j.biopsycho.2017.10.005] [Medline: 29055714]
7. Puhalla AA, Kulper DA, Fahlgren MK, McCloskey MS. The relationship between resting heart rate variability, hostility, and in vivo aggression among young adults. J Aggress Maltreat Trauma 2019 Jan 24:1-17. [doi: 10.1080/10926771.2018.1558324]
8. de Looff P, Noordzij ML, Moerbeek M, Nijman H, Didden R, Embregts P. Changes in heart rate and skin conductance in the 30 min preceding aggressive behavior. Psychophysiology 2019 Jun 11:e13420. [doi: 10.1111/psyp.13420] [Medline: 31184379]
9. Goodwin MS, Mazefsky CA, Ioannidis S, Erdogmus D, Siegel M. Predicting aggression to others in youth with autism using a wearable biosensor. Autism Res 2019 Aug;12(8):1286-1296. [doi: 10.1002/aur.2151] [Medline: 31225952]
10. Gulati G, Cornish R, Al-Taiar H, Miller C, Khosla V, Hinds C, et al. Web-based violence risk monitoring tool in psychoses: pilot study in community forensic patients. J Forensic Psychol Pract 2016 Jan 01;16(1):49-59 [FREE Full text] [doi: 10.1080/15228932.2016.1128301] [Medline: 26924945]
11. Patel MX, Sethi FN, Barnes TR, Dix R, Dratcu L, Fox B, et al. Joint BAP NAPICU evidence-based consensus guidelines for the clinical management of acute disturbance: de-escalation and rapid tranquillisation. J Psychiatr Intensive Care 2018 Sep 01;14(2):89-132. [doi: 10.20299/jpi.2018.008]
12. Ramesh T, Igoumenou A, Vazquez Montes M, Fazel S. Use of risk assessment instruments to predict violence in forensic psychiatric hospitals: a systematic review and meta-analysis. Eur Psychiatry 2018 Aug;52:47-53 [FREE Full text] [doi: 10.1016/j.eurpsy.2018.02.007] [Medline: 29626758]
13. Tully J, Larkin F, Fahy T. New technologies in the management of risk and violence in forensic settings. CNS Spectr 2015 Jun;20(3):287-294. [doi: 10.1017/S1092852915000279] [Medline: 25928733]
14. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
15. Berry N, Bucci S, Lobban F. Use of the internet and mobile phones for self-management of severe mental health problems: qualitative study of staff views. JMIR Ment Health 2017 Nov 01;4(4):e52 [FREE Full text] [doi: 10.2196/mental.8311] [Medline: 29092809]
16. Bucci S, Berry N, Morris R, Berry K, Haddock G, Lewis S, et al. “They are not hard-to-reach clients. We have just got hard-to-reach services.” Staff views of digital health tools in specialist mental health services. Front Psychiatry 2019;10:344 [FREE Full text] [doi: 10.3389/fpsyt.2019.00344] [Medline: 31133906]
17. Renwick L, Stewart D, Richardson M, Lavelle M, James K, Hardy C, et al. Aggression on inpatient units: clinical characteristics and consequences. Int J Ment Health Nurs 2016 Aug;25(4):308-318. [doi: 10.1111/inm.12191] [Medline: 26892149]
18. Simon GE. Why the nails should boss the hammers. Psychiatr Serv 2019 Aug 01;70(8):642-643. [doi: 10.1176/appi.ps.201900218] [Medline: 31138055]
19. Fusch PI, Ness LR. Are we there yet? Data saturation in qualitative research. Qual Rep 2015;20(9):1408-1416 [FREE Full text]
20. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
21. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member checking: a tool to enhance trustworthiness or merely a nod to validation? Qual Health Res 2016 Jun 22;26(13):1802-1811. [doi: 10.1177/1049732316654870] [Medline: 27340178]
22. Ogloff JRP, Daffern M. The dynamic appraisal of situational aggression: an instrument to assess risk for imminent aggression in psychiatric inpatients. Behav Sci Law 2006;24(6):799-813. [doi: 10.1002/bsl.741] [Medline: 17171770]
23. Zapf PA, Kukucka J, Kassin SM, Dror IE. Cognitive bias in forensic mental health assessment: evaluator beliefs about its nature and scope. Psychol Public Policy Law 2018 Feb;24(1):1-10. [doi: 10.1037/law0000153]
24. Ho A, Lesneskie E, Hsu KH. What we don’t know may hurt us: an examination of systematic bias in offender risk assessments. Deviant Behav 2017 Dec 04;39(12):1566-1577. [doi: 10.1080/01639625.2017.1410620]
25. National Institute for Health and Care Excellence. Violence and aggression: short-term management in mental health, health and community settings NICE guideline NG 10. London, UK: NICE; 2015 May. URL: https://www.nice.org.uk/guidance/ng10 [accessed 2019-09-11]
26. Kuivalainen S, Vehviläinen-Julkunen K, Louheranta O, Putkonen A, Repo-Tiihonen E, Tiihonen J. De-escalation techniques used, and reasons for seclusion and restraint, in a forensic psychiatric hospital. Int J Ment Health Nurs 2017 Oct;26(5):513-524. [doi: 10.1111/inm.12389] [Medline: 28960738]

27. Cox A, Hayter M, Ruane J. Alternative approaches to 'enhanced observations' in acute inpatient mental health care: a review of the literature. J Psychiatr Ment Health Nurs 2010 Mar;17(2):162-171. [doi: 10.1111/j.1365-2850.2009.01507.x] [Medline: 20465762]
28. Papadopoulos C, Ross J, Stewart D, Dack C, James K, Bowers L. The antecedents of violence and aggression within psychiatric in-patient settings. Acta Psychiatr Scand 2012 Jun;125(6):425-439. [doi: 10.1111/j.1600-0447.2012.01827.x] [Medline: 22268678]
29. McAndrew S, Chambers M, Nolan F, Thomas B, Watts P. Measuring the evidence: reviewing the literature of the measurement of therapeutic engagement in acute mental health inpatient wards. Int J Ment Health Nurs 2014 Jun;23(3):212-220. [doi: 10.1111/inm.12044] [Medline: 24103061]
30. Tully J, Cullen AE, Hearn D, Fahy T. Service evaluation of electronic monitoring (GPS tracking) in a medium secure forensic psychiatry setting. J Forensic Psychiatry Psychol 2015 Dec 24;27(2):169-176. [doi: 10.1080/14789949.2015.1122823]
31. Meyer N, Kerz M, Folarin A, Joyce DW, Jackson R, Karr C, et al. Capturing rest-activity profiles in schizophrenia using wearable and mobile technologies: development, implementation, feasibility, and acceptability of a remote monitoring platform. JMIR Mhealth Uhealth 2018 Oct 30;6(10):e188. [doi: 10.2196/mhealth.8292] [Medline: 30377146]
32. Cella M, He Z, Killikelly C, Okruszek Ł, Lewis S, Wykes T. Blending active and passive digital technology methods to improve symptom monitoring in early psychosis. Early Interv Psychiatry 2019 Feb 28. [doi: 10.1111/eip.12796] [Medline: 30821079]
33. Cella M, Okruszek Ł, Lawrence M, Zarlenga V, He Z, Wykes T. Using wearable technology to detect the autonomic signature of illness severity in schizophrenia. Schizophr Res 2018 Jun;195:537-542. [doi: 10.1016/j.schres.2017.09.028] [Medline: 28986005]
34. McGuinness D, Murphy K, Bainbridge E, Brosnan L, Keys M, Felzmann H, et al. Individuals' experiences of involuntary admissions and preserving control: qualitative study. BJPsych Open 2018 Nov;4(6):501-509 [FREE Full text] [doi: 10.1192/bjo.2018.59] [Medline: 30564446]
35. Breedvelt JJ, Zamperoni V, Kessler D, Riper H, Kleiboer AM, Elliott I, et al. GPs' attitudes towards digital technologies for depression: an online survey in primary care. Br J Gen Pract 2019 Mar;69(680):e164-e170 [FREE Full text] [doi: 10.3399/bjgp18X700721] [Medline: 30559111]
36. Huckvale K, Torous J, Larsen ME. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. JAMA Netw Open 2019 Apr 05;2(4):e192542 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.2542] [Medline: 31002321]
37. Boonstra TW, Nicholas J, Wong QJ, Shaw F, Townsend S, Christensen H. Using mobile phone sensor technology for mental health research: integrated analysis to identify hidden challenges and potential solutions. J Med Internet Res 2018 Jul 30;20(7):e10131 [FREE Full text] [doi: 10.2196/10131] [Medline: 30061092]
38. Marshall LA, Adams EA. Building from the ground up: exploring forensic mental health staff’s relationships with patients. J Forensic Psychiatry Psychol 2018 Aug 13;29(5):744-761. [doi: 10.1080/14789949.2018.1508486]
39. Noel VA, Carpenter-Song E, Acquilano SC, Torous J, Drake RE. The technology specialist: a 21st century support role in clinical care. NPJ Digit Med 2019;2:61 [FREE Full text] [doi: 10.1038/s41746-019-0137-6] [Medline: 31388565]
40. Laker C, Callard F, Flach C, Williams P, Sayer J, Wykes T. The challenge of change in acute mental health services: measuring staff perceptions of barriers to change and their relationship to job status and satisfaction using a new measure (VOCALISE). Implement Sci 2014 Feb 20;9(1):23 [FREE Full text] [doi: 10.1186/1748-5908-9-23] [Medline: 24555496]
41. Laker C, Cella M, Callard F, Wykes T. Why is change a challenge in acute mental health wards? A cross-sectional investigation of the relationships between burnout, occupational status and nurses' perceptions of barriers to change. Int J Ment Health Nurs 2019 Feb;28(1):190-198. [doi: 10.1111/inm.12517] [Medline: 29993168]
42. Murphy P, Potter L, Tully J, Hearn D, Fahy T, McCrone P. A cost comparison study of using global positioning system technology (electronic monitoring) in a medium secure forensic psychiatric service. J Forensic Psychiatry Psychol 2016 Dec 09;28(1):57-69. [doi: 10.1080/14789949.2016.1261172]
43. Torous J, Firth J. Bridging the dichotomy of actual versus aspirational digital health. World Psychiatry 2018 Feb;17(1):108-109 [FREE Full text] [doi: 10.1002/wps.20464] [Medline: 29352531]

Abbreviations

GPS: global positioning system

Edited by G Eysenbach; submitted 25.07.19; peer-reviewed by M Goodwin, S Rush; comments to author 17.08.19; revised version received 21.08.19; accepted 21.08.19; published 19.09.19.

Please cite as:
Greer B, Newbery K, Cella M, Wykes T
Predicting Inpatient Aggression in Forensic Services Using Remote Monitoring Technology: Qualitative Study of Staff Perspectives
J Med Internet Res 2019;21(9):e15620
URL: https://www.jmir.org/2019/9/e15620
doi:10.2196/15620
PMID:31538943

Original Paper

Association Between Institutional Social Media Involvement and Gastroenterology Divisional Rankings: Cohort Study

Austin Lee Chiang[1,2], MD, MPH; Loren Galler Rabinowitz[3], MD; Akhil Kumar[4], BS; Walter Wai-Yip Chan[1,2], MD, MPH

[1]Brigham and Women's Hospital, Boston, MA, United States
[2]Harvard Medical School, Boston, MA, United States
[3]Columbia University Medical Center, New York, NY, United States
[4]Washington University in St Louis, St. Louis, MO, United States

Corresponding Author:
Walter Wai-Yip Chan, MD, MPH
Brigham and Women's Hospital
75 Francis Street
Boston, MA, 02115
United States
Phone: 1 617 732 6389
Email: wwchan@bwh.harvard.edu

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2021/3/e29042

Abstract

Background: Patients often look to social media as an important tool to gather information about institutions and professionals. Since 1990, United States News and World Report (USNWR) has published annual rankings of hospitals and subspecialty divisions. It remains unknown if social media presence is associated with the USNWR gastroenterology and gastrointestinal (GI) surgery divisional rankings, or how changes in online presence over time affects division ranking.

Objective: The objective of this study was to determine if social media presence is associated with USNWR gastroenterology and GI surgery divisional rankings and to ascertain how changes in online presence over time affect division rankings.

Methods: Social media presence among the top 30 institutions listed in the 2014 USNWR gastroenterology and GI surgery divisional rankings were assessed using Pearson’s correlation coefficients and multivariate analysis, controlling for covariates. Linear and logistic regression using data from 2014 and 2016 USNWR rankings were then used to assess the association between institutional ranking or reputation score with any potential changes in numbers of followers over time. Sensitivity analysis was performed by assessing the area under the receiver operating characteristic curve to determine the follower threshold associated with improved or maintained ranking, which was done by dichotomizing changes in followers at values between the 7000 and 12,000 follower mark.

Results: Twitter follower count was an independent predictor of divisional ranking (β=.00004; *P*<.001) and reputation score (β=–.00002; *P*=.03) in 2014. Academic affiliation also independently predicted USNWR division ranking (β=5.3; *P*=.04) and reputation score (β=–7.3; *P*=.03). Between 2014 and 2016, Twitter followers remained significantly associated with improved or maintained rankings (OR 14.63; 95% CI 1.08-197.81; *P*=.04). On sensitivity analysis, an 8000 person increase in Twitter followers significantly predicted improved or maintained rankings compared to other cutoffs.

Conclusions: Institutional social media presence is independently associated with USNWR divisional ranking and reputation score. Improvement in social media following was also independently associated with improved or maintained divisional ranking and reputation score, with a threshold of 8000 additional followers as the best predictor of improved or stable ranking.

(J Med Internet Res 2019;21(9):e13345) doi:10.2196/13345

KEYWORDS

social media; Twitter; hospital ranking

Introduction

Social media refers to a variety of web-based platforms where individuals create and share information, experiences, and ideas. In recent years, social media has transformed into an important form of communication that has led to significant changes across the globe, including political uprisings. The Office of Health Information Technology within the United States Department of Health and Human Services has recognized the significance of social media in healthcare and the opportunities that lie therein for both patients and providers [1]. Medical professionals and organizations utilize social media to promote health through education and advocacy, announce news and events, recruit employees, and communicate with other providers [2]. Similarly, patients use social media to gain medical knowledge, share experiences, and provide feedback [3]. A reported 24-42% of consumers have used social media to access health-related consumer reviews, including reviews of hospitals [3,4].

Twitter is one social media platform that allows individuals or organizations to post messages limited to 280 characters or less. According to Twitter's official website, there are more than 300 million monthly active users around the world and over 500 million tweets per day [5]. Users may also follow the accounts of peers, public figures, and organizations, whose posts will automatically be displayed in a live, chronological feed. In recent years, more gastroenterologists have joined Twitter and engaged in academic discussions on a daily basis. Social media presence is modifiable and increased professional engagement could potentially impact institutional recognition and reputation. In a case study of a corporate brand, Twitter has been shown to amplify positive brand sentiments, and not only shape brand awareness and recognition but also reflect curiosity surrounding a brand [6]. Increased activity on Twitter has been shown to be correlated with Twitter following among US hospitals [7].

United States News and World Report (USNWR) boasts a reach of over 300 million readers [8]. Since 1990, USNWR annually publishes a Best Hospitals ranking of institutions and subspecialties, with 27.5% of an institution's overall score determined by reputation score [6]. These rankings are often used as a marketing tool by hospital leaders to attract patients through traditional methods of advertising, as well as on social media [9-11]. USNWR hospital rankings have been shown to have a significant impact on consumer decisions in the form of a greater than 5% change in nonemergency Medicare patients, and an estimated $76 million in revenue transferred from lower ranked hospitals to higher ranked hospitals [9]. Among 42 urology departmental Twitter accounts, there was a significant correlation between Twitter followers with USNWR reputation score [12].

The aim of this study was to determine the degree to which institutional social media involvement, based on number of followers on Twitter, may be associated with USNWR gastroenterology and gastrointestinal (GI) surgery divisional rankings (all rankings henceforth are divisional rankings). A secondary aim was to determine whether social media presence is a predictor of improvement in USNWR rankings, and to attempt to define the growth necessary to achieve improved division rankings.

Methods

This was a cohort study of the top 30 institutions listed in the 2014 and 2016 USNWR gastroenterology and GI surgery subspecialty rankings. Data from the 2014 rankings were used to address the first aim, and to address the secondary aims a follow-up study compared data based on the 2014 and the 2016 USNWR rankings. Both the numeric ranking and the institutional reputation score were publicly available on the USNWR webpage. The number of Twitter followers corresponding to the 2014 and 2016 USNWR rankings were accessed in June 2015 and November 2016 at time of analysis, respectively, as displayed on each institution's Twitter account. The official institutional Twitter accounts were identified using standard internet search tools, and hospital-specific Twitter accounts were included. Twitter accounts that appeared to combine hospital with medical school affairs were included. Medical school–specific, or university-specific accounts distinct from a hospital-based Twitter account, were excluded.

Spearman's rank correlation coefficient was calculated to evaluate the association between institutional ranking and social media following. Pearson's correlation coefficients were calculated to determine the association between social media presence and divisional reputation score. Multivariate analysis using linear regression was performed, controlling for number of hospital beds, academic affiliation, city population, and the number of full-time gastroenterologists on staff. SAS software (Cary, North Carolina) was used for statistical analyses. Among the covariates, the number of hospital beds was available online as it was published in the USNWR rankings. Other covariates such as academic affiliation (defined as an associated teaching hospital with gastroenterology fellowship training program) and number of full-time gastroenterologists on staff were obtained from the institutions' official websites. Local population size of the city in which the institution is located was obtained from the latest United States Census Bureau figures.

For our secondary aim, two-tailed, two-sample *t* tests were performed to assess the association between 2014 and 2016 institutional rankings, or reputation scores, with changes in numbers of followers between 2014 and 2016. Multivariate analysis was performed using logistic and linear regression to adjust for the aforementioned covariates. Sensitivity analysis was performed by assessing the area under the receiver operating characteristic curve (AUROC) to determine the follower threshold associated with improved or maintained ranking, which was done by dichotomizing changes in followers at the 7000, 8000, 9000, 10,000, and 12,000 follower mark.

Results

Of the top 29 institutions in the 2014 USNWR gastroenterology and GI surgery rankings, there were 28 distinct institutional Twitter accounts and one institution without a Twitter account (John Muir Medical Center). Among these Twitter accounts, the number of followers ranged from 0 to 1.7 million as of June 2015. Spearman’s rank correlation coefficient demonstrated a negative correlation between the number of Twitter followers and the numeric divisional ranking (ρ=–0.582; *P*<.001). Similarly, there was a positive correlation between Twitter followers and the divisional reputation score (ρ=0.91; *P*<.001) (Table 1). Twitter follower count was an independent predictor of divisional ranking on multivariate analysis controlling for the number of hospital beds, academic affiliation, local population, and number of full-time gastroenterologists on staff (β=0.00004, *P*<.001) (Table 1). Quantity of Twitter followers was also an independent predictor of USNWR reputation score after controlling for the same covariates (β=–0.00002, *P*=0.03) (Table 2 and Figure 1). Academic affiliation was an independent predictor of division ranking (β=5.3; *P*=0.04) and reputation score (β=–7.3; *P*=0.03) on USNWR (Table 2 and Table 3).

For the second aim, 39 hospitals were included in the analysis (Figure 2). Of the 2014 ranked institutions, five hospitals were no longer ranked in 2016 and four had rankings that had slipped under 30. Ten hospitals were newly ranked in the top 30 gastroenterology and GI surgery programs, and there were two institutions that were tied at rank #23 and #27. Of note, by 2016 the Mayo Clinic appeared three times in the top 30 ranking despite there being separate entries for the Rochester (MN), Phoenix (AZ), and Jacksonville (FL) campuses. Overall, 12 hospitals showed improved or maintained institutional ranking. Institutions with improved or maintained rankings had a significantly higher increase in followers over 1 year, with a difference of 89,287 versus 2271 (*P*=.03). On logistic regression, change in Twitter followers remained significantly associated with improved or maintained rankings (OR 14.63; 95% CI 1.08-197.81; *P*=.04). On sensitivity analysis, an increase of 8000 Twitter followers was associated with significantly higher area under the curve (AUC) (0.765) for prediction of improved or maintained ranking compared to other cutoffs (Figure 3). On multivariate analyses of 2016 rankings controlling for potential confounders using linear regression, the current number of Twitter followers for an institution was significantly associated with higher rank (β=–0.000017; *P*=.03) and reputation score (β=0.000036; *P*<.001). Academic status was also a significant predictor (β=3.446; *P*=.03) for reputation score.

Table 1. Rankings and characteristics of ranked gastroenterology and GI[a] institutions.

Hospital with Gastroenterology and GI Surgery Program	Rank 2014	Rank 2016	Followers 2014	Followers 2016	Reputation score 2014	Reputation score 2016	Beds 2014	Beds 2016	Academic	Local Population 2014	Local Population 2015	GI faculty
Mayo Clinic	1	1	1,170,000	1,360,000	60.5	52.6	1132	1243	1	110,742	112,225	94
Cleveland Clinic	2	2	387,000	751,000	38.6	36	1268	1278	1	390,113	388,072	52
Massachusetts General Hospital	3	4	21,600	35,100	19.1	18.6	947	999	1	645,966	667,137	37
Johns Hopkins Hospital	4	3	301,000	392,000	24.3	23.5	951	998	1	622,104	621,849	46
UCLA[b] Medical Center	5	5	19,300	25,000	11.1	8.6	466	466	1	3,884,000	3,971,883	44
Cedars-Sinai Medical Center	6	9	6655	9631	8.9	6.9	865	882	1	3,884,000	3,971,883	26
University of Pittsburgh Medical Center	7	6	6440	9246	13.3	9.2	1528	1517	1	305,841	304,391	66
New York Presbyterian Columbia and Cornell	8	14	22,300	28,200	7.7	5.7	2262	2328	1	8,406,000	8,550,405	84
Mount Sinai Hospital	9	7	46,300	63,400	16.5	11.8	1048	1183	1	8,406,000	8,550,405	63
Hospitals of the University of Pennsylvania	10	12	6746	9767	7.9	9.3	784	789	1	1,553,000	1,567,442	56
Northwestern Memorial Hospital	11	17	9533	13,100	6.8	5.7	885	885	1	2,719,000	2,720,546	53
Houston Methodist Hospital	12	11	1,1400	14,100	2.1	2	839	856	1	2,196,000	2,296,224	11
University Hospitals Cleveland Medical Center	13	27	10,600	13,700	1.6	2.3	771	790	1	390,113	388,072	27
Baylor University Medical Center	14	16	16,600	1680	4.6	3.4	876	844	1	1,258,000	1,300,092	40
St. Francis Hospital	15	19	136	304	0	0.1	306	306	0	2791	2791	32
Yale New Haven Hospital	16	23	13,800	17,400	2.6	2.6	1571	1576	1	130,660	130,322	48
Oschner Medical Center	17	27	3810	4905	1.9	1.6	771	789	1	23,319	23,319	10
Florida Hospital	18	—[c]	14,100	18,200	0.7	1.2	2338	2478	0	255,483	270,934	17
Beaumont Hospital	19	18	6113	7752	0.4	0.9	1070	1070	1	58,946	59,008	24
Lehigh Valley Hospital	20	38	6550	8847	0.6	0.1	793	784	0	118,577	120,207	22
St. Alexius Hospital	21	—	794	939	0.3	0	305	280	0	52,398	52,138	4
Brigham and Women's Hospital	22	45	21,100	29,000	4.1	5	779	757	1	645,966	667,137	44

Hospital with Gastroenterology and GI Surgery Program	Rank 2014	Rank 2016	Followers 2014	Followers 2016	Reputation score 2014	Reputation score 2016	Beds 2014	Beds 2016	Academic	Local Population 2014	Local Population 2015	GI faculty
University of Kansas Hospital	23	35	2331	3881	0.5	0.6	623	713	1	148,483	151,306	13
UCSF[d] Medical Center	24	15	25,000	37,100	6.7	8.2	650	650	1	837,442	864,816	46
NYU[e] Langone Hospitals	25	13	7697	12,100	1.9	2.4	791	718	1	8,406,000	8,550,405	54
University of Washington Medical Center	26	31	14,900	17,800	2.5	2.4	450	428	1	652,405	684,451	19
Hackensack University Medical Center	27	—	2627	3864	0	0.1	685	710	0	44,113	44,834	52
Bethesda North Hospital	28	—	1706	2523	0	0	375	342	0	297,517	298,550	19
John Muir Medical Center	29	—	0	207	0	0.1	367	383	0	66,900	68,910	29
Mayo Clinic Phoenix	—	8	1,170,000	1,360,000	—	—	—	268	1	1,513,000	1,563,025	27
Mayo Clinic Jacksonville	—	10	1,170,000	1,360,000	—	—	—	249	1	842,583	868,031	28
Thomas Jefferson University Hospital	—	19	—	9377	—	2.9	—	937	1	1,553,000	1,567,442	38
Stanford Health Care	—	21	—	24,000	—	3.5	—	481	1	66,642	66,853	47
University of Colorado Hospital	—	22	—	1518	—	2.3	—	648	1	345,803	359,407	19
University of Michigan Hospitals	—	23	—	21,900	—	6.5	—	962	1	117,025	117,070	51
Indiana University Health Medical Center	—	25	—	22,800	—	3	—	1243	1	852,866	853,173	76
Tampa General Hospital	—	26	—	8406	—	1.9	—	1011	0	352,957	369,075	18
Barnes-Jewish Hospital	—	28	—	7790	—	4.6	—	1323	1	318,416	315,685	23
University of Wisconsin	—	30	—	17,000	—	0.5	—	544	1	243,344	248,951	30

[a] GI: Gastrointestinal

[b] UCLA: University of California, Los Angeles

[c] Not applicable.

[d] UCSF: University of California, San Francisco

[e] NYU: New York University

Table 2. Multivariate analysis for 2014 Division Ranking per United States News and World Report.

Covariates	β coefficient	*P* value
Twitter followers	0.00004	<.001
Academic affiliation	5.3	.04
Total number of beds	0.0011	.59
Local population	–0.0000004	.58
Number of GI[a] staff	0.066	.22

[a]GI: gastrointestinal

Figure 1. Twitter followers versus 2014 United States News and World Report gastrointestinal reputation score. GI: gastrointestinal. USNWR: United States News and World Report.

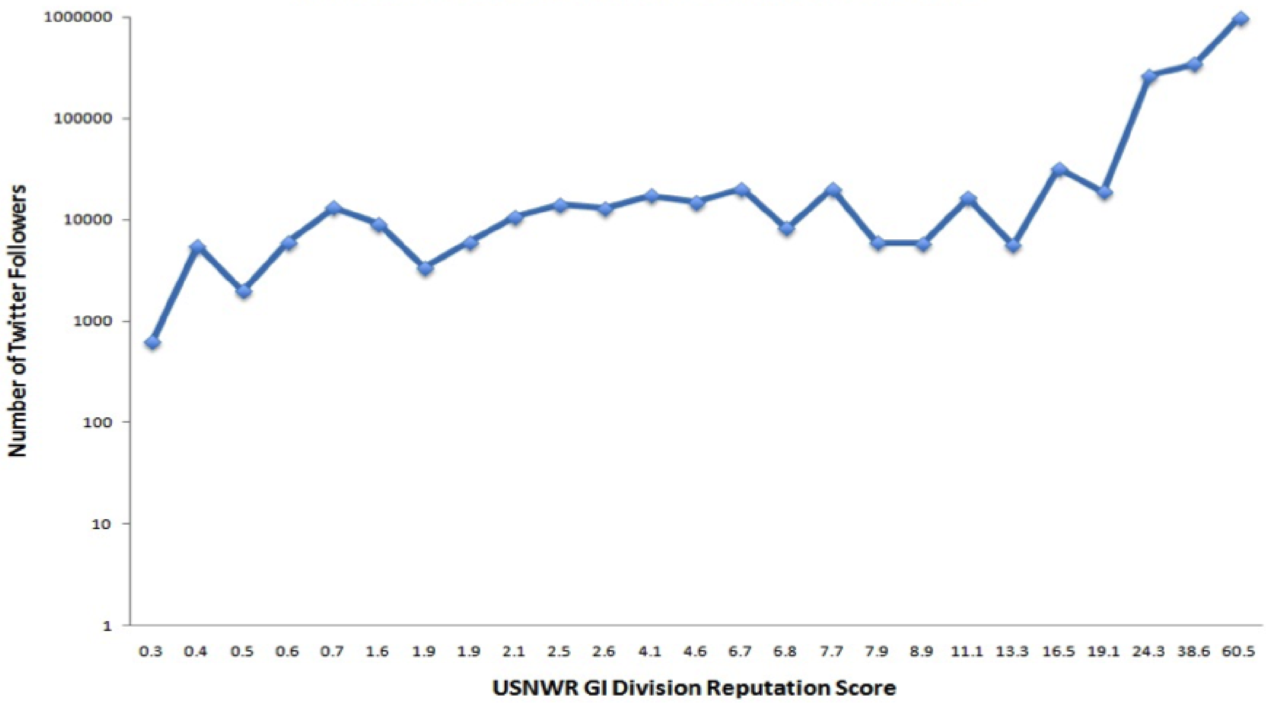


Table 3. Multivariate analysis for 2014 Reputation Score per United States News and World Report.

Covariates	β coefficient	*P* value
Twitter followers	–0.00002	.03
Academic affiliation	–7.3	.03
Total number of beds	–0.004	.12
Local population	–0.0000006	.51
Number of GI[a] staff	0.042	.55

[a]GI: gastrointestinal

Figure 2. Twitter followers by 2016 United States News and World Report ranking. GI: gastrointestinal. USNWR: United States News and World Report.

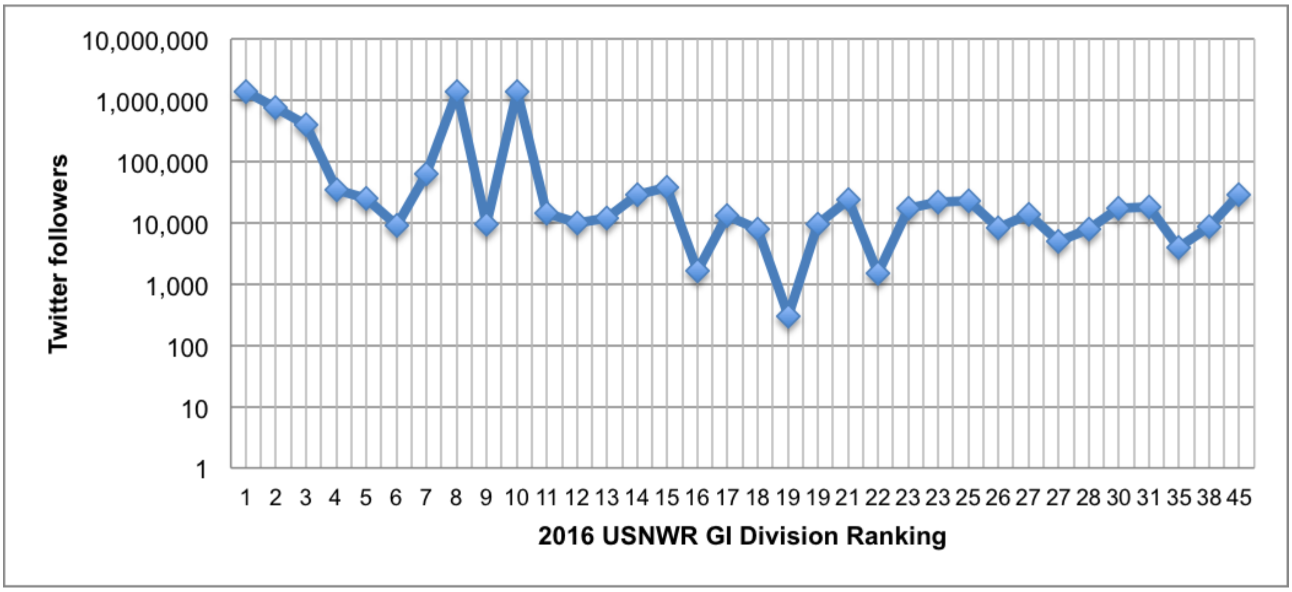
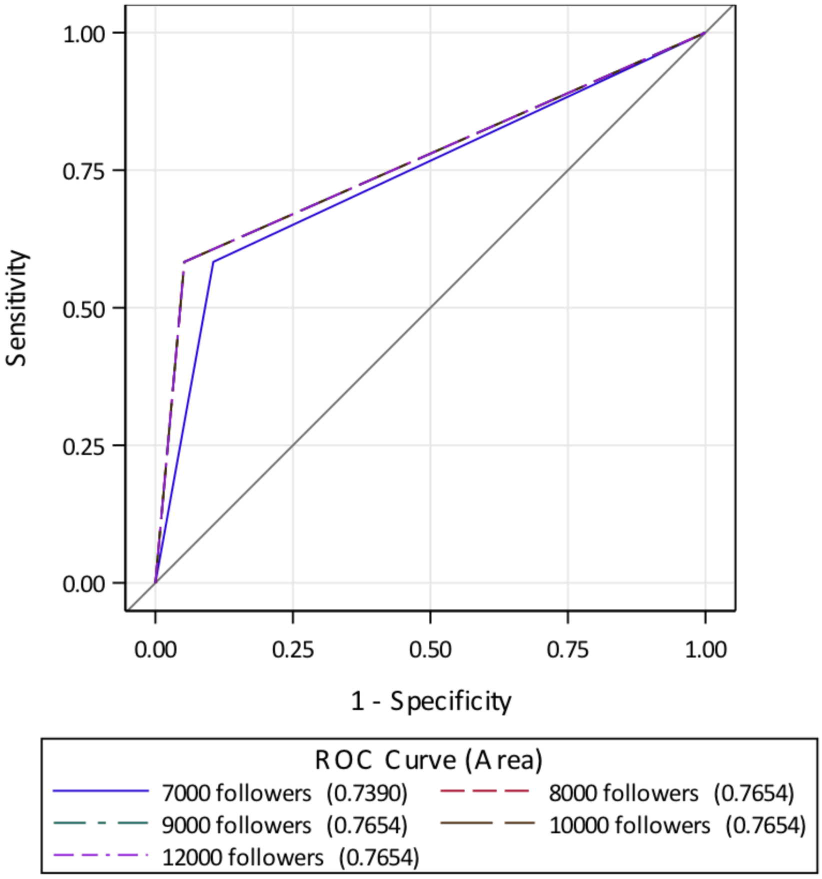


Figure 3. Receiver operating characteristic curve for different follower cutoffs. ROC: receiver operating characteristic.



Discussion

Key Results

Based on the 2014 USNWR figures, both institutional Twitter following and academic affiliation are independent predictors of gastroenterology, and GI surgery subspecialty, rankings and reputation scores. Furthermore, gaining Twitter followers between 2014 and 2016 was significantly associated with improved or maintained divisional ranking, with 8000 more followers as the greatest predictor of improved or stable divisional ranking as noted on sensitivity analysis.

According to the Pew Research Center, social media usage has seen a nearly 10-fold increase over the past more than ten years since 2005, with 65% of all adults using a social networking site [13]. According to the Mayo Clinic Health Care Social

Media List compiled in 2012, 1583/6562 hospitals in the United States have accounts on one or more social media platforms (eg, Twitter, YouTube, Facebook, LinkedIn, Foursquare), of which 1010 have Twitter accounts [14]. Thus far, gastroenterologist reception of the use of social media has been lukewarm, with only 57.5% of physicians perceiving social media as a method of obtaining "current, high-quality information" [15]. Only 47.7% of gastroenterologists reported having ever used any form of social media [16]. A better understanding of how social media might impact health-related and monetary outcomes might help galvanize use among medical professionals.

More importantly, many consumers use social media to access health-related consumer reviews, including hospital reviews [3,4]. Boosting a hospital's social media presence could be a simple and economical way of potentially elevating institutional ranking, attracting more patients, and increasing hospital revenue. By using the principles of behavioral economics, hospitals would see that it seems that simplified ranking systems drive decision making for patients over complex measures of hospital quality [9,11]. The Twitter presences of United States universities, including an extended followers network, have been shown to reflect institutional rankings, like that of UNSWR [17]. Geotagged Twitter dialogue has also been shown to reflect population-level sentiments as well as disease states and changes by geography [18,19]. Pope et al report a shift from a lower hospital rank to a higher rank is associated with both an increase of nonemergency Medicare patients and an estimated $76 million in additional revenue [9]. Furthermore, rankings influenced an estimated 15,000 hospital choices made by Medicare patients over the course of a decade, resulting in a $750 million shift in revenue [9].

The academic community may question the validity and transparency of the USNWR rankings as an objective representation of hospital quality. One study found USNWR ranked institutions had reportedly demonstrated better clinical outcomes in cardiac care and mortality for acute myocardial infarction when compared to their unranked counterparts [20]. However, Hota et al noted discrepancies between certain patient safety indicators (PSIs), as established by the Centers for Medicare and Medicaid Services (CMS), and patient safety scores as reported in USNWR [21]. Since these concerns have been raised (after our analysis), USNWR have announced changes to their ranking methodology in June 2016 that include reducing the weight of the patient safety score and eliminating one of these discrepant PSIs [22]. Even with objective measures, an institutional ranking might not accurately represent the academic or technical capabilities of an individual practitioner. In nonobjective measures, our study also demonstrated significant associations between social media presence and reputation scores. Some researchers have also noted disproportionate influence of reputation score on subspecialty ranking over objective measures [23]. However, others have defended the objectivity of reputation score, citing strong correlations with research productivity, which could impact professional perception of a division [24]. Though our study demonstrated academic affiliation as an independent predictor of ranking, some hospitals may be ranked highly for good performance despite having no academic affiliation. The converse is also true, as there are large, reputable gastroenterology divisions with relatively robust social media presences that were unranked by USNWR at the time of this study.

Limitations

The limitations of our study included using a single social media platform and using a single source for institutional rankings (USNWR), both of which may limit the generalizability of our results. Moreover, only institutional Twitter accounts were included in the study. At the time of data collection, there were too few gastroenterology division–specific accounts (eg, @Duke_GI) and subspecialty- or disease-specific Twitter accounts (eg, @UChicagoIBD, @UCLAIBD, among others). Furthermore, any causality cannot be directly established from our results. The differences in hospital rankings could also potentially affect social media following. Further research would therefore be required to elucidate the mechanism of how greater social media presence impacts hospital ranking, or if a higher divisional ranking may lead to a greater social media following. While reputation score determines 27.5% of the overall USNWR specialty score, there is also no existing data evaluating the number of Doximity voters who are active social media users or have significant exposure to dialogue occurring on Twitter [22]. Moreover, annual fluctuations in ranking might not reflect long-term social media impact, as previous studies have shown that hospital specialties vary by an average 5.49 spots per year [9]. There may have also been temporal differences that may have affected social media following at time of data collection, as data was collected at different times of year (June and November of either year). Finally, other factors could affect social media following but may not impact institutional rankings, such as significant news stories and exposure through other media outlets.

Conclusion

Institutional social media following is independently associated with USNWR divisional ranking and reputation score. Moreover, an increase in social media following was also independently associated with improved or stable divisional ranking and reputation score with a threshold of 8000 additional followers as the best predictor of improved or stable ranking. Institutions hoping to boost their overall and divisional rankings may benefit from strengthening their social media presence by both engaging the public and increasing online visibility through platforms such as Twitter.

Authors' Contributions

CL and CW were responsible for data collection, data analysis, and drafting the manuscript. RG and KA worked on data collection.

Conflicts of Interest

None declared.

References

1. Ricciardi L, Mostashari F, Murphy J, Daniel JG, Siminerio EP. A national action plan to support consumer engagement via e-health. Health Aff (Millwood) 2013 Feb;32(2):376-384. [doi: 10.1377/hlthaff.2012.1216] [Medline: 23381531]
2. Thaker SI, Nowacki AS, Mehta NB, Edwards AR. How U.S. hospitals use social media. Ann Intern Med 2011 May 17;154(10):707-708. [doi: 10.7326/0003-4819-154-10-201105170-00021] [Medline: 21576547]
3. Pricewaterhouse CHRI. PwC Health Research Institute. 2012. Social media "likes" healthcare-From marketing to social business URL: https://www.pwc.com/us/en/health-industries/health-research-institute/publications/pdf/health-care-social-media-report.pdf [accessed 2019-08-13]
4. Fox S. Pew Internet & American Life Project. 2011 May 12. The social life of health information, 2011 URL: https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf [accessed 2019-08-13]
5. Twitter. Investor Relations. 2019. Quarterly Results URL: https://investor.twitterinc.com/financial-information/quarterly-results/default.aspx [accessed 2019-08-13]
6. Jansen B, Zhang M, Sobel K, Chowdury A. Twitter power: Tweets as electronic word of mouth. J. Am. Soc. Inf. Sci 2009 Nov;60(11):2169-2188 [FREE Full text] [doi: 10.1002/asi.21149]
7. Griffis HM, Kilaru AS, Werner RM, Asch DA, Hershey JC, Hill S, et al. Use of social media across US hospitals: descriptive analysis of adoption and utilization. J Med Internet Res 2014 Nov 27;16(11):e264 [FREE Full text] [doi: 10.2196/jmir.3758] [Medline: 25431831]
8. U.S. News & World Report. 2019. Our Audience URL: https://mediakit.usnews.com/audience [accessed 2019-08-13]
9. Pope DG. Reacting to rankings: evidence from "America's Best Hospitals". J Health Econ 2009 Dec;28(6):1154-1165. [doi: 10.1016/j.jhealeco.2009.08.006] [Medline: 19818518]
10. Sehgal AR. The role of reputation in U.S. News & World Report's rankings of the top 50 American hospitals. Ann Intern Med 2010 Apr 20;152(8):521-525. [doi: 10.7326/0003-4819-152-8-201004200-00009] [Medline: 20404383]
11. Richter JP, Muhlestein DB, Wilks CEA. Social media: how hospitals use it, and opportunities for future use. J Healthc Manag 2014;59(6):447-460. [Medline: 25647968]
12. Ciprut S, Curnyn C, Davuluri M, Sternberg K, Loeb S. Twitter Activity Associated With U.S. News and World Report Reputation Scores for Urology Departments. Urology 2017 Oct;108:11-16. [doi: 10.1016/j.urology.2017.05.051] [Medline: 28669746]
13. Perrin A. Pew Research Center Internet & Technology. 2015 Oct 8. Social Media Usage: 2005-2015 URL: https://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/ [accessed 2019-08-13]
14. Bennett E. Mayo Clinic Social Media Network. 2012 Sep 12. Introducing the Health Care Social Media List URL: https://socialmedia.mayoclinic.org/2012/09/12/introducing-the-health-care-social-media-list/ [accessed 2019-08-13]
15. McGowan BS, Wasko M, Vartabedian BS, Miller RS, Freiherr DD, Abdolrasulnia M. Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. J Med Internet Res 2012 Sep 24;14(5):e117 [FREE Full text] [doi: 10.2196/jmir.2138] [Medline: 23006336]
16. Davis ED, Tang S, Glover PH, Wang AY. Impact of social media on gastroenterologists in the United States. Dig Liver Dis 2015 Mar;47(3):258-259. [doi: 10.1016/j.dld.2014.11.008] [Medline: 25499064]
17. McCoy C, Nelson M, Weigle M. Mining the Web to approximate university rankings. Information Discovery and Delivery 2018 Aug 20;46(3):173-183 [FREE Full text] [doi: 10.1108/IDD-05-2018-0014]
18. Gore RJ, Diallo S, Padilla J. You Are What You Tweet: Connecting the Geographic Variation in America's Obesity Rate to Twitter Content. PLoS One 2015;10(9):e0133505 [FREE Full text] [doi: 10.1371/journal.pone.0133505] [Medline: 26332588]
19. Mitchell L, Frank MR, Harris KD, Dodds PS, Danforth CM. The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. PLoS One 2013;8(5):e64417 [FREE Full text] [doi: 10.1371/journal.pone.0064417] [Medline: 23734200]
20. Philibert I. Can hospital rankings measure clinical and educational quality? Acad Med 2009 Feb;84(2):177-184. [doi: 10.1097/ACM.0b013e3181939034] [Medline: 19174660]
21. Hota B, Webb TA, Stein BD, Gupta R, Ansell D, Lateef O. Consumer Rankings and Health Care: Toward Validation and Transparency. Jt Comm J Qual Patient Saf 2016;42(10):439-446. [Medline: 27712602]
22. Olmsted M, Powell R, Murphy J, Bell D, Morley M, Stanley M, et al. Methodology: U.S. News & World Report 2018-19 Best Hospitals: Specialty Rankings. 2018. RTI International URL: https://www.rti.org/sites/default/files/resources/rti-publication-file-38f5185c-ddd5-4e7e-8f86-ef8fbbf848ff.pdf [accessed 2019-08-13]
23. Cua S, Moffatt-Bruce S, White S. Reputation and the Best Hospital Rankings: What Does It Really Mean? Am J Med Qual 2017;32(6):632-637. [doi: 10.1177/1062860617691843] [Medline: 28693335]

24. Prasad V, Goldstein JA. US News and World Report cancer hospital rankings: do they reflect measures of research productivity? PLoS One 2014;9(9):e107803 [FREE Full text] [doi: 10.1371/journal.pone.0107803] [Medline: 25247921]

Abbreviations

AUC: area under the curve
AUROC: area under the receiver operating characteristic curve
CMS: Centers for Medicare and Medicaid Services
GI: gastrointestinal
PSI: patient safety indicator
USNWR: United States News and World Report

Edited by G Eysenbach; submitted 08.01.19; peer-reviewed by F Timimi, B Hoffman, R Gore; comments to author 01.05.19; revised version received 26.06.19; accepted 29.06.19; published 06.09.19.

Please cite as:
Chiang AL, Rabinowitz LG, Kumar A, Chan WWY
Association Between Institutional Social Media Involvement and Gastroenterology Divisional Rankings: Cohort Study
J Med Internet Res 2019;21(9):e13345
URL: http://www.jmir.org/2019/9/e13345/
doi:10.2196/13345
PMID:31493321

Original Paper

Underage JUUL Use Patterns: Content Analysis of Reddit Messages

Yongcheng Zhan[1*], BM; Zhu Zhang[2,3*], PhD; Janet M Okamoto[4], PhD; Daniel D Zeng[2,3], PhD; Scott J Leischow[5], PhD

[1]Department of Management Information Systems, University of Arizona, Tucson, Tucson, AZ, United States

[2]State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China

[3]Shenzhen Artificial Intelligence and Data Science Institute (Longhua), Shenzhen, China

[4]Department of Research, Mayo Clinic, Phoenix, AZ, United States

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

[*]these authors contributed equally

Corresponding Author:
Yongcheng Zhan, BM
Department of Management Information Systems
University of Arizona, Tucson
1130 E Helen St
McClelland Hall 430
Tucson, AZ, 85721-0108
United States
Phone: 1 5205763741
Fax: 1 5206218105
Email: yongchengzhan@email.arizona.edu

Abstract

Background: The popularity of JUUL (an e-cigarette brand) among youth has recently been reported in news media and academic papers, which has raised great public health concerns. Little research has been conducted on the age distribution, geographic distribution, approaches to buying JUUL, and flavor preferences pertaining to underage JUUL users.

Objective: The aim of this study was to analyze social media data related to demographics, methods of access, product characteristics, and use patterns of underage JUUL use.

Methods: We collected publicly available JUUL-related data from Reddit. We extracted and summarized the age, location, and flavor preference of subreddit UnderageJuul users. We also compared common and unique users between subreddit UnderageJuul and subreddit JUUL. The methods of purchasing JUULs were analyzed by manually examining the content of the Reddit threads.

Results: A total of 716 threads and 2935 comments were collected from the subreddit UnderageJuul before it was shut down. Most threads did not mention a specific age, but ages ranged from 13 years to greater than 21 years in those that did. Mango, mint, and cucumber were the most popular among the 7 flavors listed on JUUL's official website, and 336 subreddit UnderageJuul threads mentioned 7 discreet approaches to circumvent relevant legal regulations to get JUUL products, the most common of which was purchasing JUUL from other Reddit users (n=181). Almost half of the UnderageJuul users (389/844, 46.1%) also participated in discussions on the main JUUL subreddit and sought information across multiple Reddit forums. Most (64/74, 86%) posters were from large metropolitan areas.

Conclusions: The subreddit UnderageJuul functioned as a forum to explore methods of obtaining JUUL and to discuss and recommend specific flavors before it was shut down. About half of those using UnderageJuul also used the more general JUUL subreddit, so a forum still exists where youths can attempt to share information on how to obtain JUUL and other products. Exploration of such social media data in real time for rapid public health surveillance could provide early warning for significant health risks before they become major public health threats.

(J Med Internet Res 2019;21(9):e13038) doi:10.2196/13038

KEYWORDS

electronic nicotine delivery systems; social media; minors

Introduction

Background

JUUL, the most popular e-cigarette brand in the United States, currently holds 71% of the market share, according to data from Wells Fargo in July 2018 [1]. JUUL's popularity is due to several characteristics. First, JUUL's appearance is similar to a universal serial bus (USB), which makes it easy to carry and hide as a portable device. Compared with traditional cigarette appearance, USB-like appearance has less of a negative perceived social image in public locations. Second, nicotine delivery via JUUL is effective; it is sufficiently high to warrant concern about development of nicotine addiction [2]. Third, JUUL offers a wide array of flavor options. JUUL's manufacturer claims "JUUL is for adult smokers seeking a satisfying alternative to cigarettes." However, a recent survey reported 8% of Americans aged 15 to 24 years used JUUL in the 30 days before the survey [3]. The popularity of JUUL among youth has also recently been reported in news media [4-6]. Substantial and growing evidence shows that e-cigarette use by youth and young adults increases the risk of ever using conventional cigarettes, according to a 2018 report by the National Academies of Sciences, Engineering, and Medicine [7]. In addition, as social media has become an emerging channel for information diffusion, JUUL-related discussions on platforms such as Twitter and Reddit have also experienced a sharp increase since 2015 [8,9]. Thus, this sharp increase in teen JUUL usage, and the potential health impacts, has raised great public health concerns.

With the rapid increase in JUUL use, there have been a few studies about JUUL. Several studies have been conducted to investigate the prevalence of JUUL use among adolescents [10], youth [11,12], and college students [13]. Moran et al studied sources of awareness of JUUL e-cigarettes in 2 surveys of American adults [14]. Willett et al conducted a national Web-based survey to investigate youth and young adults' knowledge, attitude, and use of JUUL [3]. Recently, Russel et al conducted a Web-based survey to assess transitions in cigarette smoking associated with the use of JUUL vaping device among 18,799 adults in the United States. [15]. Laestadius et al characterized the sale of JUUL products on eBay before a Food and Drug Administration (FDA) request to remove them, documented the impact of the request, and explored how eBay vendors bypassed FDA's effort [16]. In addition, social media data have also been used to study JUUL-related communications and their impact. Huang et al studied the extent of JUUL's growth and its marketing strategies using retail data form Nielsen and social media data from Twitter, Instagram, and YouTube [8]. JUUL-related tweets have been used to understand the public's early experiences with JUUL [17], JUUL information sharing among adolescents [18], and identified patterns of communication around JUUL use and users on Twitter [19].

Objective

Reddit has been extensively used for e-cigarette studies, which shows its value in timely public health surveillance. For example, Zhan et al identified and analyzed topics related to e-cigarettes from multiple social media platforms, including Twitter, Reddit, and JuiceDB [20]. Kavuluru et al described observations and analysis of recent JUUL-related messages on Twitter, Reddit, and in traditional media [9]. Brett et al used Reddit threads to understand reasons of JUUL use and public attitudes [21]. However, little research has been conducted on the age distribution, geographic distribution, approaches to buying JUUL, and flavor preferences pertaining to underage JUUL users. Until early January 2018, a subreddit entitled *UnderageJuul* existed, where these and other topics related to use of JUUL by youth were discussed. It was discontinued and removed by Reddit. Most casual users and analysts of Reddit thus no longer had access to the conversations in the UnderageJuul subreddit as of that date. However, we obtained part of the posting and user information of this subreddit for analyses pertaining to user age and location distributions, as well as self-reported usage patterns of JUUL. This study was aimed to inform the tobacco control community and those responsible for tobacco regulatory decision making as well as demonstrate the value of Reddit data for Web-based surveillance of public health issues.

Methods

Data Collection

Reddit data were collected from pushshift.io [22], which is a website that stores all publicly available Reddit threads and comments. Data were collected from 716 threads and 2935 comments from the subreddit UnderageJuul by the application programming interface (API) of this website. Note that these data were publicly accessible on Reddit and that no personally identifiable information is included in this study. The dataset is publicly available per request.

The first thread on the *UnderageJuul* subreddit was published on July 9, 2017, and the last was published on January 7, 2018, the day just before this subreddit was removed by Reddit. There were 844 unique users who posted threads and comments during this time period.

Data Analysis

This dataset was analyzed from 3 perspectives. First, age and location distributions described characteristics of the Redditors and provided details describing who they were. Second, because some of these Redditors were self-reported underage JUUL users, we examined how they acquired JUUL products. Third, we focused on a specific use pattern, JUUL flavor choice, and analyzed the popularity of different JUUL flavors among these Redditors. For the analyses of these 3 perspectives, we used regular expression-enhanced keyword search, which is a method for locating specific character strings embedded in text [20], to determine the most relevant threads and comments. Python is used for these analyses.

First, age and geographic distribution of *UnderageJuul* users was summarized. Threads or comments containing numerical numbers or keywords *teen*, *twenty*, *thirty*, *forty*, *fifty*, *sixty*, *seventy*, *eighty*, and *ninety* were selected for further manual examination. If the thread or comment author explicitly mentioned that "I am # years old" or used similar statements, we assigned an age label to the Redditor. This age information was further aggregated to an age distribution table. Then, we analyzed the subreddit preferences of these Redditors. For each of them, we identified a list of subreddits they had posted to. Then, the cumulative occurrence of subreddits of all UnderageJuul Redditors was calculated.

To assess the geographic distribution of *UnderageJuul* Redditors, only locations that were explicitly mentioned were included. For example, location information could be a state, an area (eg, *bay area*), a county, a city, or a zip code. These locations were manually extracted by examining the content of the threads. The extracted text location information such as *Bergen county* were converted to a standard form *Bergen County, New Jersey, United States of America* by using the Python Package GeoPy [23]. The standardized location information was then used to generate the corresponding latitude and longitude, which were further fed into a Google Maps plot.

Next, purchase approaches were analyzed by manually examining the content of threads. Purchase approaches discussed in *UnderageJuul* were classified into 7 categories. The categories were generated by an inductive category development method [24]. During this process, 2 authors individually reviewed threads consecutively and determined whether a thread contained purchase approach information. If so, based on the text information, the reviewer determined whether to assign the purchase approach to an existing category or to generate a new category. After reviewing 50% of the materials, the 2 reviewers summarized the categories and reached consensus on all categories. The coding then continued with the second half of the data; no new categories were generated during this second phase of coding. After determining the categories, the 2 reviewers independently labeled the whole dataset; 93.7% (671/716) of threads had the same labeling. The Cohen kappa was 0.906, which is considered a good agreement. The 2 reviewers reached consensus for all records after discussion.

Owing to concerns about the potential that specific JUUL flavors might increase use by and appeal to youth, we studied the flavor distribution among *UnderageJuul* threads. We identified different flavor types from the official JUUL website and counted the occurrence of these flavors in the collected threads and comments by using the regular expression search.

Results

The age distribution for the *UnderageJUUL* subreddit is shown in Table 1. We counted a Reddit user in the general categories of less than 18 years or over 21 years, if the user stated in their threads or comments that they were under 18 or greater than 21 years old; 67 out of 844 (7.9%) *UnderageJuul* Redditors reported their age in the threads or comments.

We analyzed the subreddit preferences of all *UnderageJuul* Redditors. Among 844 *UnderageJuul* Redditors, 437 had also posted threads on the *JUUL* subreddit, which is the general discussion subreddit of JUUL products. The other 9 subreddits in the top 10 most popular with *UnderageJUUL* users were as follows: *AskReddit* (general Q&A, n=250), *trees* (marijuana, n=145), *FortNiteBR* (computer game, n=125), *Drugs* (n=123), *pics* (interesting pictures sharing, n=106), *mildlyinteresting* (general discussion, n=103), *Showerthoughts* (general discussion, n=102), *teenagers* (n=96), and *funny* (general discussion, n=86). We also provided the full list of subreddit interests in Multimedia Appendix 1.

We extracted 74 explicitly mentioned locations from the thread text. We summarized the mentioned locations and labeled them on Google Maps API, as shown on Figure 1. Among these locations, New York City and the Bay Area were the most mentioned areas. The detailed location count summary to each state can be found in Multimedia Appendix 2.

Figure 1. Explicitly mentioned location distributions.

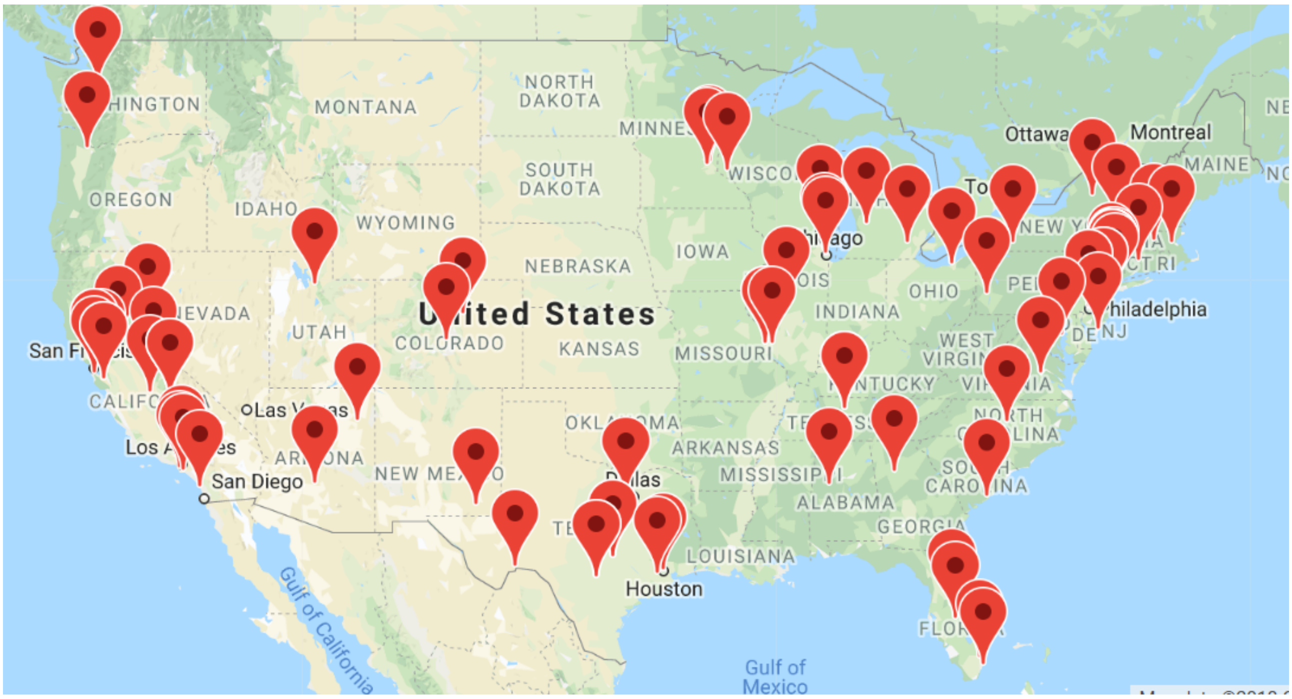


Table 1. The extracted age distribution of the *UnderageJuul* Redditors.

Age (years)	Count (n=67), n (%)
<18[a]	5 (7)
13	2 (3)
14	2 (3)
15	9 (13)
16	10 (15)
17	12 (18)
18	17 (25)
19	4 (6)
20	2 (3)
21	1 (1)
23	1 (1)
>21[a]	2 (3)

[a] These are broad categories of age where the user did not give a specific age but stated generally that they were under age (<18) or a legal adult (>21).

Next, we analyzed the approaches that were used to purchase JUUL by *UnderageJuul* users. By examining purchase requests and completed transactions in detail, we observed that there were 7 approaches that users employed to buy JUUL. These approaches are listed as follows:

1. JUUL's official website (n=34): Although JUUL changed their age policy and raised the minimum purchase age to 21 years in August, 2018 [25], JUUL required its users to be at least 18 years to order from its official website when the dataset was collected, but some *UnderageJuul* Redditors (n=24) used identity documents (IDs), date of birth, or credit cards of parents or friends to buy JUUL products. Using fake IDs (n=3) was another approach to buy JUUL products from the JUUL website. Some Redditors (n=7) were selling verified accounts on JUUL's official website, which could be used to order JUUL products as well.
2. Other vaping-related websites (n=13): Some of the other vaping-related websites sold JUUL devices and pods without age verification. Some *UnderageJuul* Redditors exchanged information about these websites. For example, 1 user asked if a particular website does an age check for buying JUUL products. Currently, this website does not sell JUUL, but it does sell e-cigarette products without age verifications. Other sites were shared, including 1 selling JUUL that has since been removed by law enforcement.
3. Peer-to-peer purchase and shipping (n=181): Some *UnderageJuul* Redditors bought JUUL products from other Redditors. The buyer usually sent money by PayPal and then the seller shipped the products to the buyer's address. A noteworthy observation was that some of the Redditors claimed they were scammed. For example, 1 user said in a thread, "don't trust ***, he scammed me and will scam you."
4. Through vendors from Amazon and eBay (n=27): This was another case of Web-based purchase and shipping. The vendor had an Amazon or eBay selling page from which the buyers purchased JUUL products. For example, 1 user shared the experience that he or she bought a JUUL from eBay and had it shipped to a mailbox of a for-sale house.
5. Face-to-face transactions (n=22): Some tried to meet up with other Reddit users in the same city and then finished transactions face to face.
6. Replacement codes of JUUL (n=47): Some bought JUUL products by using JUUL replacement codes or warranty serial numbers shared by other Redditors. The replacement codes or warranty serial numbers are provided by the JUUL company and can be used to get a brand new JUUL devices as a replacement. For example, 1 user mentioned that he wanted to sell a replacement code for $30 that could be used for a free basic kit.
7. Local stores without checking IDs (n=35): Some *UnderageJuul* Redditors did not try to buy JUUL products from other individuals. Instead, they inquired about vape shops or smoke shops in certain locations that sold JUUL but did not check IDs. For example, 1 user asked for vaping stores in NYC that did not check ID card of the buyers.

Finally, we analyzed the flavors mentioned in the subreddit. There were 7 flavors from JUUL's official website: mango, cucumber, mint, menthol, tobacco, fruit medley, and crème brûlée. We found that mango was the most mentioned flavor (n=38). Mint (n=22) and cucumber (n=21) were popular as well. The other 4 flavors, including fruit medley (n=6), crème brûlée (n=6), menthol (n=5), and tobacco (n=4), were not frequently mentioned.

Discussion

In this study, we obtained the following main findings pertaining to demographics, methods of access, product characteristics, and use patterns of underage JUUL use. First, most threads or comments did not mention specific age, but ages ranged from 13 years to greater than 21 years in those that did. Second, mango, mint, and cucumber were the most popular among the 7 flavors from JUUL's official website, and 336 subreddit *UnderageJuul* threads mentioned 7 discreet approaches to circumvent relevant legal regulations to get JUUL products, the most common of which was purchasing JUUL from other Reddit users (n=181). Third, almost half of the *UnderageJuul* users (389/844, 46.1%) also participated in discussions on the main *JUUL* subreddit and were seeking information across multiple Reddit forums. Finally, most (64/74, 86%) posters were from large metropolitan areas defined by the United States Office of Management and Budget [26]. These findings could help public health community and policy makers understand the current situation of underage JUUL use and provide them with insights on how to leverage social media for timely surveillance of emerging public health issues.

We found that only a small number of Redditors (67/844, 7.9%) reported their ages in a subreddit for underage JUUL users. After the FDA's deeming rule, the minimum age to purchase e-cigarettes for all states in the United States is 18 years. If we set 18 years as the legal age threshold, 60% (40/67) of the Reddit users in Table 1 were underage JUUL users. Alabama, Alaska, and Utah have set 19 years as the legal age to buy e-cigarettes and vaping products. In California, Hawaii, Maine, Oregon, and New Jersey, the legal age is 21 years. There are also more than 330 states, cities, and counties that have raised the minimum age to 21 years. Thus, if we raise the legal age threshold to 21 years, 94% (63/67) of the Reddit users in Table 1 were underaged. As purchasing JUUL is not legal for underage Redditors, they were more likely to avoid discussing their ages or to report fake ages. Therefore, the age distribution we obtained from explicit mentions was biased. The actual age distribution could probably be even younger. Given subreddit *UnderageJuul* was an active forum oriented for underage JUUL users before it was banned, it is still a valuable data source for analyzing underage JUUL users and their usage patterns.

We found that there was an overlap of users between the *JUUL* subreddit and the *UnderageJuul* subreddit. Specifically, 46.1% (389/844) of UnderageJuul Redditors also participated in discussions on the *JUUL* subreddit. Given the fact that most of the JUUL-related discussions on Reddit are on the *JUUL* subreddit, it is reasonable to infer that even though the *UnderageJuul* subreddit was removed by Reddit, underage JUUL users still have other forums for discussing JUUL with other users. This finding indicates that public health interventions, such as social media campaigns, are needed to target underage e-cigarette users.

We also analyzed the interests of the *UnderageJuul* users. These interests were extracted from their Reddit profiles and indicated some of these users contributed to teenager-related topics on Reddit. This finding suggests that some *UnderageJuul* Redditors have a high probability to be real underage JUUL users, though most of them did not report their ages in the *UnderageJuul* subreddit.

This study found that most of the geographic locations mentioned in threads were from large cities in the United States. In particular, New York City and the Bay Area were the most mentioned areas. The location information was collected via user-posted JUUL Web-based vendors, shipping information, or simply inquiring about local JUUL stores. As many of the inquiries were about local vape or smoke shops without age verification, the findings suggest that social media could be a useful tool for federal and state regulatory authorities to use for assessing whether vape or smoke shops are meeting regulatory requirements pertaining to e-cigarette use, especially in large cities.

Our study found that mango, mint, and cucumber flavors were more popular than other JUUL flavors among youth and young adult Reddit users, which is similar to previous findings that mint, cucumber, and mango JUUL flavors were most mentioned in tweets [9]. From previous research about e-cigarette flavors [27], strawberry was the most mentioned fruit flavor for other e-cigarette products. Instead, mango or cucumber flavors were hardly mentioned. However, these 2 flavors turned out to be successful for JUUL, suggesting that more effort is needed to understand if some flavors are more appealing to underage e-cigarette users or if a product's appeal is more important than a specific flavor availability. Tobacco flavor was widely mentioned in previous e-cigarette studies, but it was not popular in youth and young adults. One possible reason is that many of the underage JUUL users are nonsmokers, so the tobacco flavor is not appealing. The finding suggests that more research is needed for studying the impact of e-cigarette flavors on underage e-cigarette use.

Our study found that users of the *UnderageJuul* subreddit used different discreet approaches to circumvent relevant legal regulations to get JUUL. Some of these approaches (eg, purchase via eBay) are also reported in previous studies [9,16]. This study and the findings contribute to the literature by providing a detailed category dictionary for purchase approaches. On the basis of the well-established coding method, we developed 7 different categories for purchase approaches [24]. Our finding suggests that more regulation is needed to make sure that online or offline e-cigarette vendors strictly perform age verification and that legal e-cigarette users cannot provide ways of buying e-cigarettes to teenagers. In addition, parents may need to assess their teenagers' access to credit cards and IDs of legal adults.

Note that JUUL has established their age policy and required the minimum purchase age on their official website to be 21 years since August 2018. The JUUL product replacement request is also subject to the age verification. However, the identified 7 categories of workaround purchase approaches still work. Ordering the products or replacements from the official website just needs a fake ID or a preverified account. Not to mention those approaches provided by third parties such as eBay or PayPal. As the prevalence of teen JUUL use [28] and the high conversion rate from e-cigarette users to smokers [29], we call for more research and effort on this domain and take action to monitor and control these approaches.

Although this Web-based forum facilitates the JUUL transactions under the table, we still observed Redditors who tried to persuade teenagers not to start the vaping habit or not to get addicted to nicotine, especially for those who were extremely young, such as 13- or 14-year-old teens. This phenomenon represented the goodwill of the community. These comments also create additional questions about the specific population using *UnderageJuul*. Were those who made comments recommending against the use of ENDS products adults or youth? If they were adults, who are they, and why are they contributing to this particular subreddit? We cannot answer these questions given the limited information on this online forum, but we believe it is a promising future research direction to study the community and environment of underage e-cigarette users.

This study has some limitations. First, we could not verify if the self-reported ages in the data were real or fake owing to the anonymity of Reddit users. Most of these ages were general, and these Reddit users often stated that they were underage or legal adults. However, as this subreddit was specifically created for discussion of underage JUUL users, we think the mentioned ages were fairly reliable. In addition, the number of mentioned ages was relatively small compared with the number of the subreddit users, and only a small number of users shared an age. Second, the location distribution in this study faces a similar limitation as the age distribution, because we cannot get the exact addresses of Redditors from Reddit. In this study, we did not distinguish buyers' locations and sellers' locations, as we focused on the locations related to JUUL transactions. Third, the number of different flavors mentioned in the data was relatively small, but we can see their relative popularity among these Redditors. Fourth, we could not determine whether some messages were posted by e-cigarette store owners, others who have an interest in selling e-cigarettes, or even robots because the demographics of Reddit users were not available.

Despite the above limitations, the findings of this study have some important implications in terms of research and policy. It is clear that discussions about preferred flavors was 1 reason that users used the UnderageJuul subreddit and also that many users used this venue to explore ways to obtain access to JUUL. At the same time, in the 6 months that this venue existed, it was not heavily used. However, because about half of the *UnderageJuul* users also used the *JUUL* subreddit, there are alternative venues whereby youth could obtain information on flavors and methods to obtain JUUL. Our analysis of the subreddit for underage JUUL users suggests that social media data are a valuable data source for rapid public health surveillance, especially for emerging products such as JUUL, as the popularity of such products is often accompanied by discussions and promotions of the products on social media [8]. There are several limitations of these data, and they should not be the only source of information, but they provide a unique and potentially real-time source of data to better understand behavioral intentions, potential demographic variables, purchase behaviors, social networks (which were not assessed in this analysis), temporal changes across a variety of variables, and other data that could be valuable to the public health, research, and regulatory communities.

In addition, our findings indicate that social media could be a valuable surveillance tool for federal and state authorities responsible for regulating stores that sell e-cigarettes. Each state in the United States has a robust, FDA-funded process for tracking and regulating such stores, and social media data could be used to identify stores where additional assessment or observation and enforcement may be needed. When used along with other sources of data, social media data can point to emerging issues that need to be addressed in more depth.

Acknowledgments

This work was partially supported by the NIH Grant 5R01DA037378-05, NNSFC Innovative Team Grant 71621002, NNSFC Grants 71602184 and 71472175, and CAS Grant ZDRW-XH-2017-3. The work was also supported by the Ministry of Health of the People's Republic of China under Grant 2017ZX10303401-002, Research Foundation of SKL-MCCS for Young Scientists under Grants 20190212 and 20190204. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Authors' Contributions

YZ, ZZ, JO, DZ, and SL conceived the idea for this study. YZ and ZZ designed the study, conducted the data analysis, and drafted the manuscript. JO, DZ, and SL provided critical feedback, helped interpret the analysis of results, and revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

SL is the co-owner and editor-in-chief of the journal Tobacco Regulatory Science.

Multimedia Appendix 1
Subreddit interest.
[PDF File (Adobe PDF File), 1MB - jmir_v21i9e13038_app1.pdf]

Multimedia Appendix 2
Location count.
[PDF File (Adobe PDF File), 18KB - jmir_v21i9e13038_app2.pdf]

References

1. Bach L. Campaign for Tobacco-Free Kids. 2018. JUUL and Youth: Rising E-Cigarette Popularity URL: https://dev-craft.tobaccofreekids.org/assets/factsheets/0394.pdf [accessed 2019-08-04]
2. Goniewicz ML, Boykan R, Messina CR, Eliscu A, Tolentino J. High exposure to nicotine among adolescents who use Juul and other vape pod systems ('pods'). Tob Control 2018 Sep 7:- (epub ahead of print). [doi: 10.1136/tobaccocontrol-2018-054565] [Medline: 30194085]
3. Willett JG, Bennett M, Hair EC, Xiao H, Greenberg MS, Harvey E, et al. Recognition, use and perceptions of Juul among youth and young adults. Tob Control 2019 Jan;28(1):115-116. [doi: 10.1136/tobaccocontrol-2018-054273] [Medline: 29669749]
4. Ibarra AB. CNN International. 2018. JUUL E-cigarettes and Teens: 'Health Problem of the Decade'? URL: https://www.cnn.com/2018/03/15/health/juul-e-cigarette-partner/index.html [accessed 2019-08-04]
5. Dadourian M. Fox News. Students Are Sneaking Discreet New E-cigarette That Looks Like Thumb Drive URL: https://www.foxnews.com/health/students-are-sneaking-discreet-new-e-cigarette-that-looks-like-thumb-drive [accessed 2018-12-05] [WebCite Cache ID 74R5ovwh3]
6. Bui L. The Washington Post. 2018. Juuling: If You Don’t Know What It is, Ask Your Kids URL: https://www.washingtonpost.com/local/public-safety/juuling-if-you-dont-know-what-it-is-ask-your-kids/2018/05/09/37e2f026-4d65-11e8-84a0-458a1aa9ac0a_story.html?utm_term=.3986617c920c [accessed 2019-08-04]
7. The National Academies of Sciences, Engineering, and Medicine. 2018. Public Health Consequences of E-Cigarettes URL: http://nationalacademies.org/hmd/Reports/2018/public-health-consequences-of-e-cigarettes.aspx [accessed 2018-12-05] [WebCite Cache ID 74R6Ziy0J]
8. Huang J, Duan Z, Kwok J, Binns S, Vera LE, Kim Y, et al. Vaping versus Juuling: how the extraordinary growth and marketing of Juul transformed the US retail e-cigarette market. Tob Control 2019 Mar;28(2):146-151 [FREE Full text] [doi: 10.1136/tobaccocontrol-2018-054382] [Medline: 29853561]
9. Kavuluru R, Han S, Hahn EJ. On the popularity of the USB flash drive-shaped electronic cigarette Juul. Tob Control 2019 Jan;28(1):110-112 [FREE Full text] [doi: 10.1136/tobaccocontrol-2018-054259] [Medline: 29654121]
10. McKeganey N, Russell C. Prevalence of awareness and use of Juul e-cigarettes in a national probability sample of adolescents in the United States. Am J Health Behav 2019 May 1;43(3):591-605. [doi: 10.5993/AJHB.43.3.13] [Medline: 31046889]
11. Hammond D, Wackowski OA, Reid JL, O'Connor RJ, International Tobacco Control Policy Evaluation Project (ITC) Team. Use of Juul e-cigarettes among youth in the United States. Nicotine Tob Res 2018 Oct 27:- (epub ahead of print). [doi: 10.1093/ntr/nty237] [Medline: 30371838]
12. Vallone DM, Bennett M, Xiao H, Pitzer L, Hair EC. Prevalence and correlates of Juul use among a national sample of youth and young adults. Tob Control 2018 Oct 29:- (epub ahead of print). [doi: 10.1136/tobaccocontrol-2018-054693] [Medline: 30377241]
13. Ickes M, Hester JW, Wiggins AT, Rayens MK, Hahn EJ, Kavuluru R. Prevalence and reasons for Juul use among college students. J Am Coll Health 2019 Mar 26:1-5 (epub ahead of print). [doi: 10.1080/07448481.2019.1577867] [Medline: 30913003]
14. Moran MB, Chen-Sankey JC, Tan AS, Soneji S, Lee SJ, Choi K. Sources of awareness of Juul e-cigarettes in 2 surveys of adults in the United States. Am J Health Behav 2019 Mar 1;43(2):279-286. [doi: 10.5993/AJHB.43.2.5] [Medline: 30808468]
15. Russell C, McKeganey N, Haseen F. RealClearPublicAffairs. 2018. Transitions in Cigarette Smoking Associated with Use of the Juul Vaping Device Among 18,799 Adults in the United States URL: https://www.realclearpublicaffairs.com/public_affairs/2019/04/25/transitions_in_cigarette_smoking_associated_with_use_of_the_juul_vaping_device_among_18799_adults_in_the_united_states.html [accessed 2019-08-18]
16. Laestadius L, Wang Y. Youth access to Juul online: eBay sales of Juul prior to and following FDA action. Tob Control 2018 Sep 5:- (epub ahead of print). [doi: 10.1136/tobaccocontrol-2018-054499] [Medline: 30185531]
17. Allem JP, Dharmapuri L, Unger JB, Cruz TB. Characterizing Juul-related posts on Twitter. Drug Alcohol Depend 2018 Sep 1;190:1-5. [doi: 10.1016/j.drugalcdep.2018.05.018] [Medline: 29958115]
18. Chu KH, Colditz JB, Primack BA, Shensa A, Allem JP, Miller E, et al. Juul: spreading online and offline. J Adolesc Health 2018 Nov;63(5):582-586. [doi: 10.1016/j.jadohealth.2018.08.002] [Medline: 30348280]
19. Malik A, Li Y, Karbasian H, Hamari J, Johri A. Live, love, Juul: user and content analysis of Twitter posts about Juul. Am J Health Behav 2019 Mar 1;43(2):326-336. [doi: 10.5993/AJHB.43.2.9] [Medline: 30808472]
20. Zhan Y, Liu R, Li Q, Leischow SJ, Zeng DD. Identifying topics for e-cigarette user-generated contents: a case study from multiple social media platforms. J Med Internet Res 2017 Jan 20;19(1):e24 [FREE Full text] [doi: 10.2196/jmir.5780] [Medline: 28108428]
21. Brett EI, Stevens EM, Wagener TL, Leavens EL, Morgan TL, Cotton WD, et al. A content analysis of Juul discussions on social media: using Reddit to understand patterns and perceptions of Juul use. Drug Alcohol Depend 2019 Jan 1;194:358-362. [doi: 10.1016/j.drugalcdep.2018.10.014] [Medline: 30472576]
22. pushshift.io: Reddit Statistics. URL: https://pushshift.io/ [accessed 2019-08-18]
23. GeoPy's Documentation! — GeoPy 1.20.0. URL: https://geopy.readthedocs.io/en/stable/ [accessed 2018-12-05]

24. Mayring P. Social Science Open Access Repository. 2014. Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution URL: https://www.ssoar.info/ssoar/handle/document/39517 [accessed 2019-08-20]
25. Craver R. Winston-Salem Journal. 2018. Juul Labs Agrees to Support Age-21 Restrictions for Buying Its E-cigs URL: https://www.journalnow.com/news/local/juul-labs-agrees-to-support-age--restrictions-for-buying/article_9a50fcd6-488c-11e8-aa75-9f12af40387b.html [accessed 2019-08-18]
26. The White House. 2015 Jul 15. Revised Delineations of Metropolitan Statistical Areas, Micropolitan Statistical Areas, and Combined Statistical Areas, and Guidance on Uses of the Delineations of These Areas URL: https://obamawhitehouse.archives.gov/sites/default/files/omb/bulletins/2015/15-01.pdf [accessed 2019-08-18]
27. Wang L, Zhan Y, Li Q, Zeng DD, Leischow SJ, Okamoto J. An examination of electronic cigarette content on social media: analysis of e-cigarette flavor content on Reddit. Int J Environ Res Public Health 2015 Nov 20;12(11):14916-14935 [FREE Full text] [doi: 10.3390/ijerph121114916] [Medline: 26610541]
28. Centers for Disease Control and Prevention. 2019. Quick Facts on the Risks of E-cigarettes for Kids, Teens, and Young Adults URL: https://www.cdc.gov/tobacco/basic_information/e-cigarettes/Quick-Facts-on-the-Risks-of-E-cigarettes-for-Kids-Teens-and-Young-Adults.html [accessed 2019-08-18]
29. National Institute on Drug Abuse (NIDA). 2016. Teens and E-cigarettes URL: https://www.drugabuse.gov/related-topics/trends-statistics/infographics/teens-e-cigarettes [accessed 2019-08-18]

Abbreviations

API: application programming interface
FDA: Food and Drug Administration
ID: identity document
USB: universal serial bus

Edited by G Eysenbach; submitted 06.12.18; peer-reviewed by JP Allem, L Laestadius, A Davoudi; comments to author 28.04.19; revised version received 22.06.19; accepted 21.07.19; published 09.09.19.

Please cite as:
Zhan Y, Zhang Z, Okamoto JM, Zeng DD, Leischow SJ
Underage JUUL Use Patterns: Content Analysis of Reddit Messages
J Med Internet Res 2019;21(9):e13038
URL: http://www.jmir.org/2019/9/e13038/
doi:10.2196/13038
PMID:31502542

Original Paper

Social Media Outrage in Response to a School-Based Substance Use Survey: Qualitative Analysis

Ruth Ann Gassman[1*], PhD; Tapati Dutta[2*], PhD; Jon Agley[1*], MPH, PhD; Wasantha Jayawardene[1*], MD, PhD; Mikyoung Jun[1*], MPH, PhD

[1]Prevention Insights, Institute for Research on Addictive Behavior, School of Public Health, Indiana University Bloomington, Bloomington, IN, United States

[2]Rural Center for AIDS/STD Prevention, School of Public Health, Indiana University Bloomington, Bloomington, IN, United States

[*]all authors contributed equally

Corresponding Author:
Ruth Ann Gassman, PhD
Prevention Insights, Institute for Research on Addictive Behavior
School of Public Health
Indiana University Bloomington
501 N Morton St
Suite 110
Bloomington, IN, 47404
United States
Phone: 1 812 855 1237
Email: rgassman@indiana.edu

Abstract

Background: School-based alcohol, tobacco, and other drug use (ATOD) surveys are a common epidemiological means of understanding youth risk behaviors. They can be used to monitor national trends and provide data, in aggregate, to schools, communities, and states for the purposes of funding allocation, prevention programming, and other supportive infrastructure. However, such surveys sometimes are targeted by public criticism, and even legal action, often in response to a lack of perceived appropriateness. The ubiquity of social media has added the risk of potential online firestorms, or digital outrage events, to the hazards to be considered when administering such a survey. Little research has investigated the influence of online firestorms on public health survey administration, and no research has analyzed the content of such an occurrence. Analyzing this content will facilitate insights as to how practitioners can minimize the risk of generating outrage when conducting such surveys.

Objective: This study aimed to identify common themes within social media comments comprising an online firestorm that erupted in response to a school-based ATOD survey in order to inform risk-reduction strategies.

Methods: Data were collected by archiving all public comments made in response to a news study about a school-based ATOD survey that was featured on a common social networking platform. Using the general inductive approach and elements of thematic analysis, two researchers followed a multi-step protocol to clean, categorize, and consolidate data, generating codes for all 207 responses.

Results: In total, 133 comments were coded as oppositional to the survey and 74 were coded as supportive. Among the former, comments tended to reflect government-related concerns, conspiratorial or irrational thinking, issues of parental autonomy and privacy, fear of child protective services or police, issues with survey mechanisms, and reasoned disagreement. Among the latter, responses showed that posters perceived the ability to prevent abuse and neglect and support holistic health, surmised that opponents were hiding something, expressed reasoned support, or made factual statements about the survey. Consistent with research on moral outrage and digital firestorms, few comments (<10%) contained factual information about the survey; nearly half of the comments, both supportive and oppositional, were coded in categories that presupposed misinformation.

Conclusions: The components of even a small online firestorm targeting a school-based ATOD survey are nuanced and complex. It is likely impossible to be fully insulated against the risk of outrage in response to this type of public health work; however, careful articulation of procedures, anticipating specific concerns, and two-way community-based interaction may reduce risk.

(J Med Internet Res 2019;21(9):e15298) doi:10.2196/15298

KEYWORDS

social media; ATOD; survey; firestorm; digital

Introduction

School-Based Public Health Surveys

School-based alcohol, tobacco, and other drug (ATOD) surveys are internationally recognized as one of the most important ways of understanding ATOD use and related risk behaviors among young people [1]. In the United States, such surveys often include questions on attitudes (eg, perceived harm of substance use) and behaviors (eg, bullying); they are also used to monitor national trends [2] and provide aggregated information to school districts, states, and communities [3]. Administration of these surveys continues to be important, especially in the midst of a national addictions crisis that affects youth and adolescents [4] and the continually changing patterns of substance use and risk behaviors in this population, such as the recent surge in e-cigarette use [5]. Survey data and reports are also valuable in supporting continued efforts to design effective state and local education standards for substance use prevention [6] and evidence-based prevention strategies [7].

Though these surveys are designed for public health surveillance and prevention, there is a risk that they will be misunderstood as experimental research being conducted on youth, which occupies a more controversial space in public discourse [8]. Such misunderstandings, in conjunction with potentially contentious questionnaire topics like sexual health and substance use, have led to federal lawsuits against schools for administering surveys; however, two notable federal appellate cases had been resolved prior to 2007 in favor of the school districts [9]. Nonetheless, public health surveys continue to be attacked in traditional and online media [10-12], although such instances are rarely documented in academic literature. At times, clusters of incidents have also been captured in blog posts advocating for a specific position on school-based surveillance surveys [13]. School responses to such events, documented in these stories and posts [10-13], have ranged from a simple apology to cessation of survey administration.

Nature and Impact of Public Controversy

Although little research has focused on the outcomes of public controversy on school-based surveys, such negative exposure is known to have a “chilling effect” on public health scholars in general, leading to self-censorship, modification of agendas, and even career changes [14]. Controversy can also serve as an informal social control on the types of information that is *acceptable* to collect, even for the purposes of harm reduction [15]. When criticisms emerge from outside of academia, they tend to involve social media, blog posts, and occasional *extreme* reactions, such as threats of violence or demands that a particular scholar be terminated [16]. In fact, the emergence of social media has fundamentally changed how such controversies play out, with new potential for “huge waves of outrage” to emerge rapidly in an “online firestorm” [17].

Rather than stemming from fundamentally irrational behavior, online firestorms often manifest as aggressive forms of *sousveillance* [18] that “use the available masses of weak ties in social media to publicly enforce social-political norms” [19]. While similarities have been drawn between firestorms and moral panics, a recent study suggested that “social appropriateness of attacking the denounced actor” was an important antecedent of firestorm participation, as opposed to a predetermined moral stance [20]. In fact, the *piling on* behavior observed in firestorms is a fundamental component of moral grandstanding—wanting to be perceived as moral by others [21]. Although firestorms may not promote long-standing change in public discourse [22], they can result in significant short-term consequences. This can be observed, for example, in the case of Justine Sacco, a communications executive who was publicly terminated after a firestorm erupted in response to a tweet she sent in 2014 [23]. These consequences can even accrue when the firestorm is based on misleading or false information, as with Dominique Moran who was fired in response to social media outrage accusing her of racism, though she was offered her job back after the full spectrum of evidence was presented [24].

The Firestorm

In 2015, an annual ATOD survey was administered by local school officials to students in grades 6 through 12 in several hundred schools in a midwestern state. Participation in the survey by any school was voluntary. Schools that elected to participate were told that they had the choice of either electronic or paper administration formats; they were informed of their responsibility to seek either active or passive consent from parents and guardians for their children to participate in the survey. More than 25% of all schools in the state participated, and the aggregate student body of participating schools was predominantly white (approximately 70%), generally reflecting the state population.

As part of this process, the finalized survey instrument was made fully and openly available on a website for parental review for several months prior to survey administration, and schools were instructed to direct parents to that website. Schools were also instructed that at the time of survey administration, students should be told that completing the survey, as well as responding to individual survey items, was voluntary, even if their parent or guardian consented to their participation. At the beginning of the survey instrument, participants were prompted to indicate their date of birth, the initials of their first and last names, and the first letter of the school they attended at the beginning of the first grade. These data were to have been used to generate a confidential survey ID for the purposes of matching data between years, but were not designed to identify individuals, nor did the survey administrators have the ability to do so. This was the first year that this long-standing survey attempted to ask these questions, the purpose of which was to support confidential longitudinal data analyses. This information would have been used, for example, in identifying a particular grade or *cohort* with unique risks, which is commonly done in many areas of public health, for example, in Willets, 2004 [25]. Once the pilot test was fielded, the item asking for date of birth and initials raised parental curiosity. One of the families contacted a traditional media outlet, which aired a local television story about the survey.

The focus of the news story was not the new questions, though that was what had sparked interest. Instead, parents were mostly upset with several standardized risk and protective factor questions from the family domain of the Communities that Care (CTC) Youth Survey [26] having to do with "arguing about the same things in my family over and over," "having serious arguments," and "often insulting or yelling at each other." Immediately following this story, the new questions, but not the CTC items, were removed from the final survey tool. Despite this, a minor firestorm erupted on social media in the comments section of the local news station's page and a vitriolic debate ensued.

This Study

School-based surveys support harm reduction and substance use prevention efforts. The concomitant risk of public controversy and, more recently, of a digital firestorm may make school officials, epidemiologists, and public health scholars wary of advancing this important work. To better understand the nature of digital firestorms targeting school-based ATOD surveys, this study analyzes the qualitative comments made on social media during a small firestorm targeting a school-based ATOD survey administered by an independent organization and funded by a state-level government agency. The study also uses that information to make recommendations for how those who administer school-based surveys might proactively set in place protective structures that minimize the likelihood of controversy.

Methods

Data Collection

Data were captured by archiving all public comments posted to social media in response to the news story described above (ie, *the firestorm*) about a school-based public health survey. Local media featured this story on television then published the story—the only story covering this topic—to their website and two social media outlets. On one of the social media platforms, the post had no active engagement of any type. Thus, study data were drawn exclusively from the other social media platform. The thread was active (ie, the time from posting the story to social media until the final comment was posted) for approximately 35 hours. Because of the nature of online debate, at the point of archiving, several of the more incendiary posts had either been made private or had been redacted prior to capture and were, therefore, not included in the dataset. Although all data were posted in a digital public forum, to protect individuals' privacy, this study did not identify the social media platform in question; quotes that might potentially be used to identify individuals were paraphrased or edited for publication, without changing the tone or meaning, to remove colloquialisms that would render the strings unique.

This work was granted *Exempt* status by the Indiana University Institutional Review Board (number: 1511842027).

Theoretical Orientation

More than a decade ago, researchers identified a "crisis of trust" in how the public at-large views science [27]. As lay exploration of science related to all topics, including public health, has moved to the Internet, it is probable that the level and type of interactivity matters in how content is understood and the ways that the public reacts to the content [28]. Further, comment sections in online science articles often veer from the initial story content, particularly in the direction of moral and ethical claims and community-driven narrative [29].

Because these comments were made in a public digital forum, they offered a naturalistic opportunity to explore and categorize the components of a digital firestorm targeting a public health survey tool. For this reason, we used a general inductive approach [30] informed by theoretical thematic analysis [31] to extract individual and overlapping categories from individual comments.

Data Analysis

Data were transcribed into Dedoose and cleaned by the lead author (RAG), who then segmented comments into two broad categories: oppositional and supportive. For comments that were brief or unclear (eg, only the word "Agreed"), opposition or support was determined by context. Comments by the same person that were broken into multiple *posts* on the thread were counted as single *comments* when they focused on the same message, though this occurred very infrequently since this was a small firestorm.

A second reviewer validated comment sorting before the study moved forward. Then, within each of the two categories, the second researcher (1) completed multiple *deep reads* of the dataset, (2) generated preliminary themes and prepared a codebook, (3) assigned one or more codes to each social media comment, (4) revised the themes, and (5) recoded each individual social media comment. The lead author then (6) received the coded data and codebook, (7) reviewed each individual comment, (8) suggested revisions to the codebook, and (9) recoded each comment, highlighting any potential areas of disagreement. Finally, the two researchers met in person to discuss the data and continued shared analysis until 100% agreement on coding for each comment was achieved.

Results

Negative or Oppositional Comments

Overview

A total of 133 individual comments were coded as negative or oppositional. Within this segment, researchers identified seven categories, as outlined in Table 1.

The majority of comments (105/133, 78.9%) fell within a single categorical element. The most frequently shared categories were *government related* + *conspiratorial or irrational* (10/133, 7.5%) and *reasoned disagreement* + *survey mechanisms* (6/133, 4.5%). Most comments were easy to parse for categorical content, though a few posed conceptual difficulty and required lengthier discussion. For example, we coded a comment about children belonging to the state, and the potential for *resistance*, as *government related* and *conspiratorial or irrational*, but cases could be made for additional categories depending on how the intention of the comment is interpreted. Frequencies and percentages of categories are provided in Table 2.

Table 1. Negative or oppositional comments.

Comment category	Description
Government related	This code was used to indicate comments referencing government overreach, as distinct from school overreach; criticisms of government in any form; or references to a political agenda. It specifically excluded conspiratorial thinking.
Conspiratorial or irrational	This code was used to indicate comments displaying conspiratorial thinking or creation of implications that demonstrably surpass a *reasonable person* standard (eg, comparison of a survey to actions taken by Nazi Germany). Importantly, simply expressing concern about the unique identifier did not qualify a comment for this code.
Parental autonomy and privacy	This code was used to indicate comments focusing on the roles or rights of parents, especially vis a vis the school; the appropriate role of the school versus other entities; or criticism of the school's involvement in students' home lives, including privacy concerns.
Child protective services and police	This code was used to indicate comments suggesting that the survey was a tool that could result in a law enforcement agency investigating parents.
Reasoned disagreement	This code was used to indicate criticism of the survey that attempted to make a logical case against it.
Survey mechanisms	This code was used to indicate concerns about survey mechanisms, including perceptions about informed consent, confidentiality, and the voluntary nature of its completion.
Other or miscellaneous	This code was used to categorize trolling within negative threads (eg, spelling correction), affirmation of negativity (eg, simply writing "Agreed"), or other outlier comments.

Table 2. Amount of negative or oppositional social media comments containing identified categorical elements.

Comment category[a]	Number of comments (N=133), n (%)
Government related	22 (16.5)
Conspiratorial or irrational	19 (14.3)
Parental autonomy and privacy	53 (39.8)
Child protective services and police	7 (5.3)
Reasoned disagreement	16 (12.0)
Survey mechanisms	15 (11.3)
Other or miscellaneous	33 (24.8)

[a] Categories were not mutually exclusive. The most common overlapping codes were *government related* + *conspiratorial or irrational* (10/133, 7.5%) and *reasoned disagreement* + *survey mechanisms* (6/133, 4.5%).

Government Related

Most respondents who discussed the government used this firestorm as a place to insert specific criticisms of the state or federal government, sometimes without any direct mention of the survey.

> *The state can't accomplish anything positive...* [Participant #4]
>
> *That's the public government indoctrination school system...* [Participant #177]

Other respondents enmeshed criticism or concern about the government with other thematic areas, including conspiratorial thinking and parental issues. Some participants argued that the survey was an attempt to violate civil rights and control children's upbringing while *pretending* to be a process intended to protect youth. Others simply asserted government intrusiveness.

> *Government is...increasingly intrusive on people's lives...* [Participant #62]

Conspiratorial or Irrational

Nearly all responses fitting this category referenced Nazi ideology or fascism, perhaps in loose observance of Godwin's law [32]. One participant, for example, referenced Nazi Germany and then indicated that this type of questionnaire historically has been a tool used against parents by oppressive governments. Other responses used memes or articulated more specific, if unrealistic, concerns related to large-scale arrests of those whose children responded to the surveys in certain ways.

> *[Meme used as part of response: A picture of Anne Frank saying, "hiding away up in the attic, are we Anne Frank? Why don't you just obey the law?"]* [Participant #0]
>
> *Soon they will [ask] children [whether their parents criticize] the school board...or other elected officials.*

> *Then [the parents will be sent] to jail just like...dictatorships [in the past]...* [Participant #62]

Parental Autonomy and Privacy

A number of respondents focused on defining what the school's role should be in society, both in general and in distinction to other organizations or groups.

> *It's not their place to ask!* [Participant #201]
>
> *That looks like a doctor's office survey so leave it [there].* [Participant #192]

Other responses indicated a feeling of having their privacy invaded by the school, and implied that the schools were on a slippery slope in terms of the questions they were asking. One participant, for example, began listing sex positions and referenced the Kama Sutra in the context of stating that none of those things are anyone's business, though no questions on the survey pertained to sexuality. Comments also sometimes suggested that this was an attempt for schools to gain authority over youth.

> *We need to take back control of our children...* [Participant #1]
>
> *What goes on in the home [isn't] their business...schools already overreach...* [Participant #19]

Child Protective Services and Police

There was a small cluster of responses that specifically expressed concern that the survey would trigger legal action to seize children.

> *...my fear is [the survey] may also have [the government and child protective services]...causing trouble where none exists by taking answers out of context...* [Participant #52]

Reasoned Disagreement

Some respondents did not support or agree with the survey being conducted, but they approached their concerns in a measured manner. One participant suggested calmly instructing youth first to decline participation and then to notify the parents if they received pushback against that decision. Others posed theoretical disagreement or questions about the survey.

> *If the school [were concerned], they [wouldn't necessarily have to ask] all of the kids [whether their parents argue all of the time].* [Participant #67]

Survey Mechanisms

A number of comments indicated concerns with the survey protocol, though it is not possible to tell whether these reflect actual lapses in protocol at individual schools or respondent concerns that are not based on actual events.

> *...students at <redacted> were not informed participation was voluntary, and were told they had to fill out the survey...* [Participant #91]

Positive or Supportive Comments

Overview

A total of 74 individual comments were coded as positive or supportive. Within this segment, researchers identified six categories, as discussed in Table 3.

The majority of comments (48/74, 65%) fell within a single categorical element. The most frequently shared categories were *prevent abuse or neglect + hiding something* (8/74, 11%) and *factual statements + reasoned support* (7/74, 9%).

Frequencies and percentages of categories are provided in Table 4.

Table 3. Positive or supportive comments.

Comment category	Description
Prevent abuse or neglect	This code was used to indicate comments suggesting that the survey would serve to prevent or intervene with specific cases of abuse or neglect.
Holistic health	This code was used to indicate comments suggesting that youth are affected by home life or environment or justifying the school's role in supporting home life.
Hiding something	This code was used to indicate comments suggesting that the only reason to be upset with youth completing the survey is that the parents have something to hide.
Factual statements	This code was used to indicate comments demonstrating a factual understanding of how the ATOD[a] survey works.
Reasoned support	This code was used to indicate support for the survey that attempted to make a logical case for it.
Other or miscellaneous	This code was used to categorize trolling within supportive threads (eg, spelling correction), affirmation of positivity (eg, simply writing "Agreed"), or other outlier comments.

[a] ATOD: alcohol, tobacco, and other drug.

Table 4. Amount of positive or supportive social media comments containing identified categorical elements.

Comment category[a]	Number of comments (N=74), n (%)
Prevent abuse or neglect	17 (23)
Holistic health	17 (23)
Hiding something	18 (24)
Factual statements	15 (20)
Reasoned support	16 (22)
Other or miscellaneous	23 (31)

[a] Categories were not mutually exclusive. The most common overlapping codes were *prevent abuse or neglect* + *hiding something* (8/74, 11%) and *factual statements* + *reasoned support* (7/74, 9%).

Prevent Abuse or Neglect

A number of participants supported the survey because they believed that it could be used to identify specific youth who were experiencing adverse events (eg, abuse) and, in some cases, to intervene with those individuals.

> *...it is important for the truth to [be revealed so that youth can be saved]. I would be happy to [clarify my own actions with child protective services in order to help protect] other kids that are being abused or neglected...* [Participant #51]
>
> *Who had a problem, abusive parents?* [Participant #99]

Holistic Health

Other participants justified administration of the survey by asserting that home life often affects educational performance or that parents cannot or should not be inconsistent in requests that the school invest in their children's lives.

> *[Anyone who says that home life has] nothing to do with education must be joking!* [Participant #24]
>
> *...You want the school to be involved [and not involved at the same time]. [Parents need to] figure this out...[it's] the first step to ensure kids' safety...* [Participant #74]

Hiding Something

Some individuals indicated that there was nothing to worry about as long as others had nothing to hide; others took this concept one step further, asserting that those who were upset about the survey specifically had something to hide from authorities or school personnel.

> *If there is nothing to hide it should not matter...* [Participant #60]
>
> *...[if you don't want your own child to answer these questions], is it because you are scared of what they will say...? If you answered yes to either [of the questions about home life], you are the reason they have to ask these questions.* [Participant #65]

Factual Statements

A few participants provided statements outlining factual information about the ATOD survey. This included both comments related to the mechanics of survey administration as well as statements related to the implications of the survey (eg, what is intended to be done with the data).

> *The survey results are [used] to gather data...when you see stories that [describe the prevalence of youth smoking or drug use]...This is where that information comes from.* [Participant #68]
>
> *These surveys are given to every student...They are completely anonymous. Teachers [cannot] administer them [to preserve privacy].* [Participant #30]

Reasoned Support

Although there was a meaningful amount of overlap between factual understanding of the survey and reasoned support, there were also other supportive statements from individuals who did not assert knowledge about the survey but who attempted to reason through the debate.

> *[It would be hard to find a family that could honestly say they didn't argue]...That's part of family life...we don't always say warm and fuzzy things...but [we] still [care for one another].* [Participant #61]
>
> *Schools...cannot win. [It's unfortunate]...[what if we could potentially] offer support and programs [to] guardians who may need assistance?* [Participant #82]

Discussion

Principal Findings

This study reviewed the full spectrum of social media comments posted as part of a localized social media firestorm surrounding administration of a school-based ATOD survey. As one might expect, such discourse is primed for a certain amount of controversy [33]. Analyses found that the responses were split at the broadest levels of *opposition* or *support*, though nearly two-thirds (133/207, 64.3%) opposed or criticized the survey. Further, most comments focused on the survey itself, the perceived context in which it was offered, or broader claims about society; few responses displayed individually directed *flaming* behavior (eg, "negative violations of...interactional norms") [34].

Notably, only 7.2% (15/207) of the total responses contained factual information about the ATOD survey, and many of the other comments specifically relied on inaccurate assumptions about the survey. It is objectively true that the survey responses could not have been linked to individual students by the survey administrators; however, many commenters who wrote oppositional messages, especially those related to child protective services and privacy, either were unaware of this fact or did not believe it to be true. Likewise, many conspiratorial or seemingly irrational comments presupposed a broader, sinister agenda, components of which would require the ability to link survey data to individual students. At the same time, it is important to avoid “white hat bias” [35] by noting that some comments supportive of the survey were also advanced based on the same incorrect premises; in particular, supportive statements suggested that the ATOD survey could prevent individual cases of abuse or maltreatment or that parents only disliked the CTC items because they might be caught hiding harmful behavior; however, neither of those scenarios were possible.

In total, many (99/207, 47.8%) coded comments were in the following categories: *parental autonomy and privacy*, *child protective services and police*, *conspiratorial or irrational*, *prevent abuse and neglect*, or *hiding something*. Although there was no overarching code for misinformation, comments in each of these categories, both positive and negative, logically presupposed inaccurate information about the purpose of the survey or the survey methods. In the context of digital moral outrage [36], one interpretation of the high volume of this type of comment is that the negative comments were indicating that the survey’s perceived invasions of privacy or question content violated social or moral norms in varying ways. On the other hand, the positive comments were presenting oppositional norms in which child safety was more highly valued than privacy or preferences about the kinds of questions that should be posed to youth.

Consistent with prior research [29], some of the comments not directly related to privacy, both positive and negative, also relied on a priori personal or community-based normative beliefs. This was most commonly observed in the *government related* and *holistic health* categories, manifesting as general mistrust or disdain for the government in the former and an underlying belief or overt assertion that home life is interlinked with school life in the latter. Many comments in these categories could have been posted as template responses to any news story focusing on government activity or the role of the school in raising youth.

One important general finding from this study is that while this social media firestorm occurred in the context of a school-based ATOD survey, much of the online activity was driven by factors external to the survey itself, though the consequences to the survey administration process could easily have been substantial. In fact, organizations administering school-based ATOD surveys have a vested interest in avoiding controversy, especially in online formats. Based on the data from this study, it is tempting simply to suggest that more specific, directed information provision about the nature of the survey and the purpose of administration would serve as *firestorm protection* for school-based ATOD surveys. This is an especially compelling argument given that this particular firestorm included a high density of conflict (ie, prosurvey and antisurvey) stemming from norms unrelated to the actual survey itself. At the same time, such an approach likely must be nuanced and should not minimize or trivialize these norms. For example, if there is widespread pre-existing concern about privacy, in general, in a segment of the population, then it may be beneficial to avoid the appearance of violating privacy, even when such a thing could not occur in practice.

Recommendations

The data in this study cannot determine whether misunderstandings about the survey stemmed from a lack of information availability or from a lack of belief in available information. It is plausible that the reality is a mixture of the two. Researchers have long acknowledged that “public understanding of science, and of public risk perceptions, are not so much about public capabilities...but about the trust and credibility they are prepared to invest...” [37]. This may extend beyond conspiratorial or government-related concerns to the schools themselves, which may not be trusted [38] or which may, as seen in this study, be viewed solely as an adversarial government apparatus. Thus, while it may be necessary to thoroughly outline survey procedures and protocols for public consumption, it is likely not sufficient. This study also suggests the importance of providing clarity about the survey’s purpose, how the data will be used, how they will be safeguarded, and who oversees the survey at each level. Further, there may be value in facilitating some level of community member coconstruction of the messaging, both to engender trust and to identify potential *pain points* that might spark outrage [39].

There is also a documented disconnection between public outrage and objective level of risk, especially when risks are perceived as involuntary or unfamiliar [40]. It is important to acknowledge the community’s perceived risks associated with data collection, even when they may objectively be minimal, as meaningful concerns [41] and to take additional steps to mitigate or address them. The most obvious step is to avoid collecting anything that might appear to link data to participants, even when safeguards are in place to prevent that from occurring. Other strategies might also include additional or redundant layers of student protection, for example, messaging written on the survey itself indicating, “even if you have been instructed by your teacher that you must take this survey, you are not required to do so.”

Ultimately, however, prior research [15,18-21,29] suggests that online firestorms are complex and often emerge as a result of many factors outside of the context of the specific *target*. Data from this study do nothing to attenuate that claim, with a significant digital presence in the firestorm allocated to a priori beliefs about the government, child welfare, and society. Thus, it is likely impossible to be 100% protected against the possibility of digital outrage when conducting a school-based survey. At the same time, depending on the scope or intent of a given public health project (eg, global immunization), it may be possible to generate targeted social media messaging to assuage some concerns (eg, the World Health Organization’s syndicated, organized, social media campaigns) [42].

Limitations and Study Context

This study is subject to several limitations. First, although all data points were captured and reviewed, this study was intended to deeply understand this particular incident. One cannot assume that all digital firestorms will include similar content, though we suggest applicability of these findings specifically to school-based ATOD surveys administered for epidemiological purposes. Second, the study team did not compute interrater reliability for the coding, choosing instead to complete iterative review until 100% concordance was reached. Interrater reliability for qualitative research has long been a subject of discussion among scholars, who have noted that agreement on basic themes can easily be achieved, but interpretation of those themes can diverge [43]. Such divergence has been observed even in highly rigorous qualitative studies and tends to result from complex information (ie, multiple themes and subthemes) within a single piece of data [44]. Because of the nature of the data in this study—generally short, written statements rather than transcripts of interviews or focus groups and the a priori determination of the research interest (ie, the survey)—it was feasible to achieve full agreement on themes and interpretation relative to the research topic among the coders. Finally, this study did not have the ability to determine whether comments originated from people in the state where the firestorm occurred, nor was it able to determine whether comments were from online trolls or social bots intending to influence the discussion, for example, in Ferrara et al, 2016 [45].

To the authors' knowledge, there are no extant studies that have investigated digital outrage relative to public health ATOD surveys in schools and few studies that have researched this in public health in general. Much of the writing about online firestorms has examined marketing and branding, including stakeholder relations and online *paracrisis* (ie, "a publicly visible crisis threat that charges an organization with irresponsible or unethical behavior") [46]. This work has included, for example, experimental assessment of a proposed model of paracrisis development in response to a visual ad [47] and nonacademic books designed to help businesses (eg, *How to Protect [or Destroy] Your Reputation Online*) [48]. This study advanced understanding of the online firestorm phenomenon specifically as it pertains to school-based epidemiological surveys, which was heretofore mostly unexplored. At the same time, as noted above, it is limited in scope to the specific incident described and does not address, nor was it intended to address, the more general research literature conceptualizing digital outrage.

Conflicts of Interest

RAG, JA, and MJ were involved with conducting the school-based ATOD survey that was the target of this specific firestorm.

References

1. United Nations Office on Drugs and Crime. Conducting School Surveys on Drug Abuse: Global Assessment Programme on Drug Abuse. Toolkit Module 3. New York, NY: United Nations; 2003. URL: https://www.unodc.org/documents/GAP/GAP%20Toolkit%20Module%203%20ENGLISH.pdf [accessed 2019-08-21]
2. Johnston LD, Miech RA, O'Malley PM, Bachman JG, Schulenberg JE, Patrick ME. University of Michigan Library: Deep Blue. 2018. Monitoring the Future national survey results on drug use, 1975-2017: Overview, key findings on adolescent drug use URL: https://deepblue.lib.umich.edu/handle/2027.42/148123 [accessed 2019-08-21]
3. Gassman R, Jun M, Samuel S, Agley JD, Lee J, Wolf J. Indiana Youth Survey – 2018. Bloomington, IN: Institute for Research on Addictive Behavior; 2018. URL: https://inys.indiana.edu/docs/survey/indianaYouthSurvey_2018.pdf [accessed 2019-08-21]
4. Eisdorfer S, Galinkin J. Opioid use disorder in children and adolescents: Risk factors, detection, and treatment. Clin J Pain 2019 Jun;35(6):521-524. [doi: 10.1097/AJP.0000000000000708] [Medline: 30985401]
5. Barrington-Trimis JL, Berhane K, Unger JB, Cruz TB, Urman R, Chou CP, et al. The e-cigarette social environment, e-cigarette use, and susceptibility to cigarette smoking. J Adolesc Health 2016 Jul;59(1):75-80 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.03.019] [Medline: 27161417]
6. Bruckner TA, Domina T, Hwang JK, Gerlinger J, Carpenter C, Wakefield S. State-level education standards for substance use prevention programs in schools: A systematic content analysis. J Adolesc Health 2014 Apr;54(4):467-473 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.07.020] [Medline: 24016749]
7. Skiba D, Monroe J, Wodarski JS. Adolescent substance use: Reviewing the effectiveness of prevention strategies. Soc Work 2004 Jul 01;49(3):343-353. [doi: 10.1093/sw/49.3.343]
8. Felzmann H. Ethical issues in school-based research. Res Ethics 2009 Sep;5(3):104-109. [doi: 10.1177/174701610900500304]
9. Soronen LE. Surveying students about controversial subjects. J Sch Health 2007 Mar;77(2):92-97. [doi: 10.1111/j.1746-1561.2007.00173.x] [Medline: 17222161]
10. Colias M. Rapid City Journal. 2014 Nov 13. After complaints, school district postpones behavioral surveys URL: https://rapidcityjournal.com/news/local/after-complaints-school-district-postpones-behavioral-surveys/article_415b37e1-9b69-5894-8d97-e41e71044478.html [accessed 2019-08-21]
11. Ward D. NBC4 Washington. 2014 Nov 07. Parents discuss controversial school sex survey URL: https://www.nbcwashington.com/news/local/Parents-Discuss-Controversial-School-Sex-Survey_Washington-DC-281985481.html [accessed 2019-08-21]

12. Kelley D. The Gazette. 2015 Sep 06. Controversial Healthy Kids survey back in Colorado schools this fall URL: https://gazette.com/business/controversial-healthy-kids-survey-back-in-colorado-schools-this-fall/article_f6d932b6-763f-5e61-a4d9-8600c296d917.html [accessed 2019-08-21]
13. School Choice for New Hampshire. 2015 Feb 14. Controversial surveys and questionnaires need opt-out URL: https://www.schoolchoicenh.org/2015/02/14/controversial-surveys-and-questionnaires-need-opt-out/ [accessed 2019-08-21]
14. Kempner J. The chilling effect: How do researchers react to controversy? PLoS Med 2008 Nov 18;5(11):e222 [FREE Full text] [doi: 10.1371/journal.pmed.0050222] [Medline: 19018657]
15. Kempner J, Merz JF, Bosk CL. Forbidden knowledge: Public controversy and the production of nonknowledge. Sociol Forum 2011;26(3):475-500. [doi: 10.1111/j.1573-7861.2011.01259.x]
16. Hoepner JE. 'You Need to Shut Up': Research Silencing and What it Reveals About Academic Freedom [doctoral thesis]. Canberra, Australia: Australian National University; 2017 Jul. URL: https://openresearch-repository.anu.edu.au/bitstream/1885/121823/1/Hoepner%20Thesis%202017.pdf [accessed 2019-08-21]
17. Pfeffer J, Zorbach T, Carley KM. Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. J Mark Commun 2013 Jun 10;20(1-2):117-128. [doi: 10.1080/13527266.2013.797778]
18. Mann S, Ferenbok J. New media and the power politics of sousveillance in a surveillance-dominated world. Surveill Soc 2013 Jul 16;11(1/2):18-34. [doi: 10.24908/ss.v11i1/2.4456]
19. Rost K, Stahel L, Frey BS. Digital social norm enforcement: Online firestorms in social media. PLoS One 2016 Jun 17;11(6):e0155923 [FREE Full text] [doi: 10.1371/journal.pone.0155923] [Medline: 27315071]
20. Johnen M, Jungblut M, Ziegele M. The digital outcry: What incites participation behavior in an online firestorm? New Media Soc 2017 Nov 29;20(9):3140-3160. [doi: 10.1177/1461444817741883]
21. Tosi J, Warmke B. Moral grandstanding. Philos Public Aff 2016 Dec 27;44(3):197-217. [doi: 10.1111/papa.12075]
22. Lamba H, Malik MM, Pfeffer J. A tempest in a teacup? Analyzing firestorms on Twitter. In: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). 2015 Presented at: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM); August 25-28, 2015; Paris, France URL: https://dl.acm.org/citation.cfm?id=2808828 [doi: 10.1145/2808797.2808828]
23. Ronson J. The New York Times Magazine. 2015 Feb 12. How one stupid tweet blew up Justine Sacco's life URL: https://www.nytimes.com/2015/02/15/magazine/how-one-stupid-tweet-ruined-justine-saccos-life.html [accessed 2019-08-21]
24. Blake J. How an internet mob falsely painted a Chipotle employee as racist. 2019. URL: https://www.cnn.com/2019/05/25/us/false-racism-internet-mob-chipotle-video/index.html [accessed 2019-08-21]
25. Willets RC. The cohort effect: Insights and explanations. Br Actuar J 2011 Jun 10;10(4):833-877. [doi: 10.1017/S1357321700002762]
26. Arthur MW, Briney JS, Hawkins JD, Abbott RD, Brooke-Weiss BL, Catalano RF. Measuring risk and protection in communities using the Communities That Care Youth Survey. Eval Program Plann 2007 May;30(2):197-211. [doi: 10.1016/j.evalprogplan.2007.01.009] [Medline: 17689325]
27. Bauer MW, Allum N, Miller S. What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. Public Underst Sci 2016 Aug 18;16(1):79-95. [doi: 10.1177/0963662506071287]
28. Brossard D. New media landscapes and the science information consumer. Proc Natl Acad Sci U S A 2013 Aug 20;110 Suppl 3:14096-14101 [FREE Full text] [doi: 10.1073/pnas.1212744110] [Medline: 23940316]
29. Laslo E, Baram-Tsabari A, Lewenstein BV. A growth medium for the message: Online science journalism affordances for exploring public discourse of science and ethics. Journalism 2011 Sep 08;12(7):847-870. [doi: 10.1177/1464884911412709]
30. Thomas D. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
32. Dictionary.com. 2019. Godwin's law URL: https://www.dictionary.com/e/memes/godwins-law/ [accessed 2019-08-21]
33. Grubbs J, Warmke B, Tosi J, James AS. Moral grandstanding in public discourse: Status-seeking motives as a potential explanatory mechanism in predicting conflict. PsyArXiv 2019 Apr:1-20 [FREE Full text] [doi: 10.31234/osf.io/gnaj5]
34. O'Sullivan PB, Flanagin AJ. Reconceptualizing 'flaming' and other problematic messages. New Media Soc 2016 Jun 30;5(1):69-94. [doi: 10.1177/1461444803005001908]
35. Cope MB, Allison DB. White hat bias: Examples of its presence in obesity research and a call for renewed commitment to faithfulness in research reporting. Int J Obes (Lond) 2010 Jan 1;34(1):84-88; discussion 83 [FREE Full text] [doi: 10.1038/ijo.2009.239] [Medline: 19949416]
36. Crockett MJ. Moral outrage in the digital age. Nat Hum Behav 2017 Nov 18;1(11):769-771. [doi: 10.1038/s41562-017-0213-3] [Medline: 31024117]
37. Wynne B. Misunderstood misunderstanding: Social identities and public uptake of science. Public Understand Sci 2016 Dec 21;1(3):281-304. [doi: 10.1088/0963-6625/1/3/004]
38. Gassman RA, Agley J, Fly AD, Beckmeyer JJ, He K, Aaron Sayegh M, et al. Using biological samples for youth ATOD survey validation. Addict Res Theory 2015 Oct 12;24(3):177-185. [doi: 10.3109/16066359.2015.1094061]
39. Davies SR. Constructing communication. Sci Commun 2008 Mar 27;29(4):413-434. [doi: 10.1177/1075547008316222]

40. De Boer M, Mccarthy M, Brennan M, Kelly A, Ritson C. Public understanding of food risk issues and food risk messages on the island of Ireland: The views of food safety experts. J Food Saf 2005 Nov;25(4):241-265. [doi: 10.1111/j.1745-4565.2005.00020.x]
41. Cook G, Pieri E, Robbins PT. 'The scientists think and the public feels': Expert perceptions of the discourse of GM food. Discourse Soc 2016 Jul 25;15(4):433-449. [doi: 10.1177/0957926504043708]
42. Centers for Disease Control and Prevention. 2018 Oct 18. Global health social media toolkits URL: https://www.cdc.gov/globalhealth/socialmedia/toolkits/index.html [accessed 2019-08-21]
43. Armstrong D, Gosling A, Weinman J, Marteau T. The place of inter-rater reliability in qualitative research: An empirical study. Sociology 2016 Jul 02;31(3):597-606. [doi: 10.1177/0038038597031003015]
44. Moretti F, van Vliet L, Bensing J, Deledda G, Mazzi M, Rimondini M, et al. A standardized approach to qualitative content analysis of focus group discussions from different countries. Patient Educ Couns 2011 Mar;82(3):420-428. [doi: 10.1016/j.pec.2011.01.005] [Medline: 21292424]
45. Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. Commun ACM 2016 Jun 24;59(7):96-104. [doi: 10.1145/2818717]
46. Coombs WT, Holladay JS. The paracrisis: The challenges created by publicly managing crisis prevention. Public Relat Rev 2012 Sep;38(3):408-415. [doi: 10.1016/j.pubrev.2012.04.004]
47. Lim JS. How a paracrisis situation is instigated by an online firestorm and visual mockery: Testing a paracrisis development model. Comput Human Behav 2017 Feb;67:252-263. [doi: 10.1016/j.chb.2016.10.032]
48. David JP. How to Protect (Or Destroy) Your Reputation Online: The Essential Guide to Avoid Digital Damage, Lock Down Your Brand, and Defend Your Business. Wayne, NJ: The Career Press; 2017.

Abbreviations

ATOD: alcohol, tobacco, and other drug
CTC: Communities that Care

Edited by G Eysenbach; submitted 28.06.19; peer-reviewed by JP Allem, E Da Silva, Y Rivera, AJ Greenberg; comments to author 22.07.19; revised version received 26.07.19; accepted 27.07.19; published 12.09.19.

Please cite as:
Gassman RA, Dutta T, Agley J, Jayawardene W, Jun M
Social Media Outrage in Response to a School-Based Substance Use Survey: Qualitative Analysis
J Med Internet Res 2019;21(9):e15298
URL: http://www.jmir.org/2019/9/e15298/
doi:10.2196/15298
PMID:31516129

Original Paper

Family Separation and the Impact of Digital Technology on the Mental Health of Refugee Families in the United States: Qualitative Study

Sayyed Fawad Ali Shah[1], PhD; Julia Meredith Hess[2], PhD; Jessica R Goodkind[3], PhD

[1]Department of Communication, Jacksonville State University, Jacksonville, AL, United States
[2]Department of Pediatrics, University of New Mexico School of Medicine, Albuquerque, NM, United States
[3]Department of Sociology, University of New Mexico, Albuquerque, NM, United States

Corresponding Author:
Jessica R Goodkind, PhD
Department of Sociology
University of New Mexico
MSC05 3080
Albuquerque, NM, 87131
United States
Phone: 1 505 277 2501
Email: jgoodkin@unm.edu

Abstract

Background: Conflicts around the world have resulted in a record high number of refugees. Family separation is a critical factor that impacts refugee mental health. Thus, it is important to explore refugees' ability to maintain contact with family members across the globe and the ways in which they attempt to do so. It is increasingly common for refugees to use information and communication technologies (ICTs), which include mobile phones, the internet, and social media sites, such as Facebook, WhatsApp, Skype, and Viber, for these purposes.

Objective: The aim of this study was to explore refugees' perceptions of the impact of communication through ICTs on their mental health, the exercise of agency by refugees within the context of ICT use, especially their communication with their families, and logistical issues that affect their access to ICTs in the United States.

Methods: We used a constructivist grounded theory approach to analyze in-depth interviews of 290 adult refugee participants from different countries, who were enrolled in a randomized controlled trial of a community-based mental health intervention.

Results: Analyses showed that communication through ICTs had differing impacts on the mental health of refugee participants. ICTs, as channels of communication between separated families, were a major source of emotional and mental well-being for a large number of refugee participants. However, for some participants, the communication process with separated family members through digital technology was mentally and emotionally difficult. The participants also discussed ways in which they hide adversities from their families through selective use of different ICTs. Several participants noted logistical and financial barriers to communicating with their families through ICTs.

Conclusions: These findings are important in elucidating aspects of refugee agency and environmental constraints that need to be further explicated in theories related to ICT use as well as in providing insight for researchers and practitioners involved in efforts related to migration and mental health.

(J Med Internet Res 2019;21(9):e14171) doi:10.2196/14171

KEYWORDS

refugees; family separation; mental health; mobile phone

Introduction

Background

Conflicts in different parts of the world, especially the Middle East and Africa, have resulted in a significant number of people who have experienced forced displacement worldwide. According to the United Nations High Commission for Refugees (UNHCR) [1], the number of displaced people reached an all-time high—70.8 million—at the end of 2018. UNHCR has classified 25.9 million of these displaced people as refugees. UNHCR defines a refugee as someone who has been forced to leave his or her country because of violence, persecution, or war and has a well-founded fear of persecution that prevents them from returning home [2].

Refugees, similar to other forcibly displaced people, often experience traumatic events, such as torture, mass killings, and political victimization. Exposure to these traumatic experiences negatively affects the mental health of refugees [3,4]. In addition to witnessing and experiencing these traumatic events, refugees often leave some or all of their family members behind. Family separation is a critical factor that impacts the mental health of refugees [5]. Often, refugees use information and communication technologies (ICTs), which include mobile phones, the internet, and social media sites, such as Facebook, WhatsApp, Skype, and Viber, to stay in touch with their family [6-8]. This study aimed to understand (1) refugee perceptions of the impact of the use of ICTs to communicate with family members on their mental health and (2) the exercise of agency by refugees, within the context of their imagined affordances of ICTs (users' perceptions of communication technologies that influence how they approach them), especially with respect to their communication with their families.

Family Separation and Mental Health

Several studies have documented the negative impacts of family separation on the mental health of refugees [5,9]. Sudanese refugees resettled in Australia perceived family separation as the most significant factor that impacted their mental well-being. They described their most salient mental health issues as resulting from missing their families and being concerned about the financial situation of their families who were residing in Sudan [10]. Miller et al [9] conducted a mixed-methods study to understand how family separation affected the mental health of refugees who resettled in the United States. The authors found that separation from a family member or loved one was 1 of only 2 traumatic experiences out of the 27 they measured, which explained significant variance in all the 3 components of refugees' mental health included in their study—posttraumatic stress disorder (PTSD) symptoms, emotional distress, and psychological quality of life—even after accounting for refugees' overall level of trauma exposure. They also found that family separation caused mental stress among refugees in different forms. For example, refugees were fearful about the safety of their family members who remained in their home countries, which made them question their decision to resettle in the United States.

Information and Communication Technology Use Among Refugees

ICTs are often very important in the lives of refugees. ICTs can facilitate refugees' well-being by helping them navigate the difficult terrain to their destinations, accumulate information about host countries, and get in touch with friends and family [11-14]. However, ICT use may also endanger refugee lives by making their information available for law and order personnel in the countries of their destinations or by making state or nonstate actors of the countries they are fleeing from aware of their flight or location [15,16]. Studies looking at refugees' use of ICTs are mainly focused on (1) refugees' mobilization from their home countries to their countries of refuge [14,15,17] and (2) the resettlement of refugees in new countries [6,11,13,18-20].

Studies investigating the use of ICTs primarily explore their use during the journey: maps to navigate routes, text and phone to remain in contact with family and their smugglers, and at times to save their lives by calling the border patrol in case of emergency [14,15]. Similarly, several studies examine the ways in which refugees use ICTs to meet several social and economic needs once they resettle in their destinations. For example, Kaufmann [21] found that Syrian refugees in Vienna used smartphones to access information, learn a language, and stay in touch with their families. He suggests that the refugees in his study were emotionally attached to their smartphones. Similarly, Wilding [22] noted that young refugees in Australia used ICTs to maintain their links with several communities at a time, including family and friends in the country of origin; therefore, the ICTs helped them transcend time and space. A study done in the Netherlands found that refugees used ICTs to learn local language and culture, among other things [11]. Similarly, studies have noted that refugees use ICTs to access different resources. For example, a study in Kenya revealed that ICTs offered platforms for refugees to access higher education in refugee camps [23]. The studies that examined the use of digital technologies by refugees during resettlement mainly focus on the use of ICTs for gratifying social needs, such as communicating with other refugees, family and friends, and promotion of health [20,24,25].

Information and Communication Technologies and Family Separation

Given the distress that family separation causes to refugees, it is important to explore refugees' ability to maintain contact with their family members and the ways in which they attempt to do so. An increasingly common way through which refugees are able to stay in touch with family members is through the use of ICTs. ICTs enable *copresence*, a state of being emotionally connected with families that are far away [26].

ICTs can act as a vital lifeline for refugees who have resettled in a new country, and research has revealed that ICTs can help improve the mental health of refugees by enabling them to find displaced family members, communicate with their families and friends and inform them of their needs, and receive and send financial assistance [25]. UNHCR and host country governments are also exploring the use of ICTs to help refugees resettle in host countries and reduce the distress caused by family separation. Some research suggests that ICTs can contribute to bridging the separation gap by making Web-based family reunification possible [10]. Thus, it seems that digital technology can provide separated families the opportunity to retain a communication link with their families.

Recent studies have found that cellular networks are the most common tools used to access the internet among refugees and that refugees living in camps use the internet and social media for several purposes, including connecting with family, building social support networks, and telling their stories [27,28]. A study conducted in a refugee camp in Syria noted that social media, such as WhatsApp and Facebook, were more often used for communication and information-seeking purposes by refugees than mobile voice, short message service text messaging, Skype, and email [29]. Digital technology has not only helped refugees connect with their families, but it has also helped them integrate into their host countries by allowing them the opportunity to access amenities on the Web, such as banking systems and aid agencies, and share their stories with different segments of the society [30]. Overall, research has documented that the primary use of ICTs among refugees is for communicating with their families [31]. However, the literature on the impact of ICTs on the mental health of refugees in the context of family separation is limited.

Refugees’ Agency in the Usage of Information and Communication Technologies

Research shows that refugees’ agency to interact with or through ICTs is dependent on the imagined affordances of the ICTs [16]. Imagined affordances are the users’ approaches to a technology shaped by the users’ expectations from that technology [32]. For example, in the context of mobility to Europe from Syria, refugees used smartphones to navigate their routes and stay in touch with their families and, at the same time, they feared negative consequences, such as being exposed to surveillance by state and nonstate actors [16]. Therefore, these possibilities of benefit and harm shaped how refugees strategically used technology while moving to their destination. However, there is a lack of literature on how refugees make strategic decisions about and exercise their agency to communicate through ICTs with their friends and family, especially in the context of resettlement in host countries.

Barriers to Refugee Usage of Information and Communication Technologies

Some scholars argue that even though ICTs are essential tools of communication among refugees, the limited ability to physically access ICTs and limited skills to use these technologies have resulted in inequities among different refugee groups [7,33]. For example, newly arriving refugees to Australia have shown low ICT literacy, as they were deprived of Web-based communication tools during their displacement [28]. As the number of forcibly displaced people around the world increases, it is essential to understand more about barriers to ICT use, which is faced by diverse refugee populations.

Thus, although there is burgeoning and important literature on ICT use among refugees, there is a need to explore refugee perceptions of the impact of ICTs on their mental health and how refugees’ decisions and actions illustrate their agency in their interaction with ICTs, in the context of resettlement in their host countries, and potential barriers they may face to ICT usage.

Hence, our study was guided by 2 overarching research questions.

Research question 1: What is the range of mental health effects refugees report as a result of communicating with separated family members using ICTs?

Research question 2: Given the potential positive and negative impacts of ICT use on refugee mental health, how do refugees exercise agency (within a context of opportunities and constraints that shape their access to ICTs) to protect their well-being?

Methods

Setting

The data for this study were collected as part of a randomized controlled trial of the Refugee Well-being Project (RWP), from 2013 to 2018. Funded by the National Institute on Minority Health and Health Disparities, the study examined the effectiveness of a community-based advocacy and learning mental health intervention for recently arrived refugees in New Mexico, 1 of 49 states where refugees are resettled in the United States. Participants included refugee adults from Afghanistan, Iraq, Syria, and the Great Lakes region of Africa (Burundi, Democratic Republic of Congo, and Rwanda), who were resettled in the United States for less than 3 years. Participants were randomized by household to an intervention or waitlist control condition. Participants in the intervention were matched with an undergraduate or graduate student (receiving training and supervision through enrollment in a 2-semester course) at the University of New Mexico. Together, the student and refugee partners engaged for 6 months in mutual learning and the mobilization of community resources to address the goals and unmet needs of refugee participants. Those in the waitlist control group were invited to participate in a 1-time stress management session, and they were eligible to participate in the RWP intervention after the study period ended. All participants who reported high PTSD symptoms were invited to receive Narrative Exposure Therapy (NET) at no cost.

Participants

Study participants were enrolled in 4 cohorts (2013-2016). Of the 290 study participants, 31.0% (90/290) were from the Great Lakes Region of Africa, 36.2% (105/290) were Afghan, and 32.8% (95/290) were Middle Eastern (Iraqi and Syrian). A total of 52.4% (152/290) of the participants were women, and they ranged in age from 18 to 71 years (mean 34.6). More than half (58.3%; 169/290) of the participants were married, with 33.1% (96/290) single, 6.6% (19/290) widowed, and 1.7% (5/290) divorced. At the first interview time point, they had been in the United States for an average of 30 weeks. Participants completed mixed-method interviews at 4 time points (pre, mid, post, and follow-up), over the course of 12 months. All participants completed a qualitative interview at the first time point; we purposely sampled 16 participants in each of the 4 cohorts for qualitative interviews at the 3 later time points, selecting participants who presented rich preliminary data in both the NET-eligible (clinically significant PTSD symptoms) and ineligible categories and across gender and national origin groups.

Methodological Approaches

The mixed-method research used a convergent parallel design [34]. The qualitative and quantitative strands were separate; however, we designed each strand with complementarity at the forefront, meaning that we built on the strengths of each strand to answer specific kinds of research questions. The data reported in this study are qualitative, which were collected with an aim to provide an in-depth understanding of processes we were seeking to understand. Constructivist grounded theory guided the qualitative analysis [35]. Constructivist grounded theory acknowledges that analysis is a coconstruction of researchers and participants and, as such, is an interpretivist approach, but one that takes into account the specific place and time in which the research is conducted. Constructivist grounded theory is flexible, but in general, its tenets direct researchers to stay close to the data, attend to processes and actions, and use abduction to make inferences about the social world with which we are engaging.

The qualitative component of each interview involved a semistructured interview guide, with questions in numerous domains, including open-ended questions about the benefits and challenges of life in the United States, work experiences, and impacts of resettlement on participants' health and family. Other domains in the interview included questions about refugees' social networks, cultural maintenance, and changes across all domains since the previous interview. The quantitative strand of the study included collecting information about participants' mental health, social networks, and access to resources. Data from previous RWP research had underscored the importance of refugee participants maintaining communication and receiving and providing support from family members and friends from whom they had been separated during the migration journey. Thus, we added questions about if and how often they contacted family members and others, as well as the impacts these contacts had on their health and well-being.

Interviews were conducted with an interpreter who spoke a common language with the participant (eg, Arabic, Dari, Pashto, Kirundi or Kinyarwanda, Swahili, or French). The average interview length was 50 min, with a range from 12 to 143 min. Participants were provided with gift cards in compensation for their time (amounts increased for each time point from US $20 to US $50). The research was approved by the University of New Mexico Human Research Protections Office.

Data Analysis

English portions of qualitative interviews were checked for accuracy, anonymized, and imported into NVivo 10, a qualitative data analysis software package (QSR International). All qualitative data were analyzed using a multistep process. First, all data were coded according to question, allowing for rapid analysis of research questions. To begin our analysis for this study, we focused on participant responses to the following questions: "Are you still in contact with people from your home country or country of first refuge?" and "How does that impact your well-being?" A second round of coding was applied to the dataset by 2 independent coders. The hierarchical coding tree was collaboratively established around inductive and deductive codes that were primarily descriptive in nature. Given the large number of interviews and resulting number of references coded at prominent categories (eg, health), we used NVivo's querying functions to examine intersecting codes. Thus, we looked across the entire dataset for themes related to mental health that intersected with participant discussions of family in their home country or country of refuge (see Table 1). This was achieved through using the matrix query function of NVivo 10.

Table 1. Number of themes related to mental health, which intersected with participants' discussions of their social networks.

Health nodes and subnodes	Number of health nodes cocoded at the node *Social network—family, friends, neighbors, and colleagues*
Health	14
Healing, recovery	03
Memory, forgetfulness	01
Mental health	21
Afraid	07
Alone, lonely	09
Depressed, depression	11
Happy	42
Homesick, missing family and friends	75
Positive, optimistic outlook	18
Sad	20
Stress	27
Suffering	07
Tired	04
Worry	52
Physical health	14
Treatment	04

Third, focused coding and memo writing [35] were conducted on prominent categories related to communication through ICTs with family and its impact on well-being. Memo writing attempts to (1) determine the definition of a category, explicating its subcategories; (2) examine the conditions under which the category arises, is maintained, and changes; (3) describe its consequences; and (4) show how it relates to other categories. The data were also examined for anomalies and contradictions to explore further in the analysis. Sorting, diagramming, and elucidating connections and relationships among these ideas were also done through memo writing [35]. These memos were the basis of the results shared below.

Results

Refugee participants described a range of mental health effects from communicating with separated family members using ICTs, including positive impacts, such as healing, encouragement, social support, and bridging the physical distance separating them, and negative effects related to distress, elicited by heightened concerns about family's safety and feelings of helplessness, as they could not help their distant family members.

Communication as a Healing Process

Data from the semistructured qualitative interviews revealed that ICTs, as channels of communication between separated families, were a major source of emotional and mental well-being for a large number of refugee participants. Many participants described the process of talking to their families through ICTs as mentally, emotionally, and spiritually healing. They believed that talking to their families and friends through phones, social media, and other services, such as WhatsApp and Viber, made them feel good and relieved stress and depression. For example, an Iraqi man explained the following:

> *There is no day can pass without talking to them. I have to talk to them every day and all the times, a lot of times.*

Thereafter, the interviewer asked the following question:

> *And when you talk with them how does that make you feel?*

The participant replied by saying the following:

> *Of course, I will be comfortable, I will be comforted, because whenever I feel depressed, and a boring time, I have to talk to them, I have to call them.*

Encouragement and Social Support From Family and Friends

Participants discussed the encouragement and moral help they received from talking to friends and family members on several occasions. They described the encouragement they obtained from their family and friends to address the challenges they face in the United States as helpful in managing postresettlement stressors. In part, this was because communication with family and friends prompted the participants to think about the financial and political situations of their home countries. These reminders from families and friends motivated them to face the difficulties of resettlement and living in the United States. Communication with family and friends also provided them with the much-needed social support required to persevere in challenging times. A male participant from Iraq described how talking with people from home affected him:

> *Yeah, I'm still like in contact with my friends in Syria but just as far as saying hi and stuff. My friends are like encouraging me and telling me to be patient, friends from over there tell me to be patient, that nothing comes as simple as you want it to be.*

Then, the interviewer asked the following question:

> *And so, when you talk with your friends and family in Iraq and Syria, how does it make you feel and how does it affect your well-being?*

The participant said the following:

> *In a good way. They make me feel more comfortable spiritually, like deep inside of me. I talk to them over Tango, and I sit with a bunch of them talking, and they just make me feel a lot better and happier.*

The comparison of the financial, political, and educational opportunities available for participants and their dependents in the United States and their families and friends in the home countries also motivated participants to stay in the United States, despite facing many problems associated with resettlement. A male participant from Democratic Republic of Congo said the following:

> *When I talk to them, I always compare my life and their lives, so I remember where I came from. I know they have so many problems. They don't have food. They don't have money to pay rent. They don't have money to take their kids to school. They have all those kinds of problems. Then, I thank God for bringing me here. I don't need food, and I don't need to pay for school for my kids. I thank God for that even though I know that I do have problems. I haven't found a job and other stuff, but there's a difference between them and me.*

Communication with family and friends reminded participants how hard things were back home, which encouraged them to get through tough times in the United States.

Overcoming the Physical Distance

Another prominent theme in our data was how ICTs helped bridge the gap between separated families. Participants described how their communication with relatives through ICTs helped decrease the burden of separation from their families and fulfilled some of their emotional and mental needs by making them feel at home when they were away from their homes. An Afghan woman described this process:

> *I say every night I talk to them. At least like I think they're very close to me. I say we talk on the Skype; they see us like we feel like they are very close and in touch with us.*

An Iraqi woman also echoed these feelings:

> *When I talk with them I feel good. You know, the day passes very happy for me and I feel that I'm home and I feel that everybody's around me. Nowadays the networking system [ICTs] has really brought the people very close to each other.*

Another Iraqi participant echoed the same feelings:

> *Yeah, I have my friends and I usually contact them like to see what things over there are happening. What things changed after we left? Yeah, so we're in continuous contact with them.*

Then, the interviewer asked the following question:

> *And so, how does that make you feel being able to contact them still?*

The participant responded by saying the following:

> *It makes me feel better because when you leave something that you live for a long time, you think that there is something that attracts you. So, you need to know what happened over there. Who is feeling well. Who is feeling not well. Yeah so, when you are in continuous contact with your friends, with your family, it will support you here.*

The comment shows the ways in which communication among families, through ICTs, motivated participants to overcome the difficult circumstances in their host countries, with the help and support of faraway friends and family. However, not all communication with family members was helpful or had a beneficial effect on participants. Some refugees described the process of communicating with their separated family members as depressing and stressful.

Communication as a Source of Depression and Stress

Distress of Leaving Family Members in Conflict Situations

In some interviews, participants stated that they felt stressed and depressed after talking to their family members and hearing about the dangers they faced every day. A female participant from Afghanistan felt that she had done injustice to her family members by leaving them back home.

The interviewer asked the following question:

> *When you talk to them, how does that impact your well-being?*

The participant responded by saying the following:

> *My [daughter-in-law's mother] and family members are crying when I call, and I cry also. I'm the most unlucky woman in the world, so I don't show this to my son because he is here working hard, and I don't want him to know how badly I feel. So, I hide it. When we came from Kyrgyzstan to the U.S., we stopped in Istanbul and then went to Chicago, and I cried the whole way because I had to leave my son there.*

Another female participant from Democratic Republic of Congo spoke about how she felt emotionally drained after talking to her family members through the phone or internet.

The interviewer asked the following question:

> *Yeah? Okay. And is that difficult for you? Is that, sometimes it can be positive and sometimes it can be negative because you don't know how people are doing, and I'm just wondering how that's been for you?*

The participant responded by saying the following:

> *We're not okay because it's hard especially when you leave your family somewhere and you know that place is not good. They call us and tell us to check the social media and we check Facebook and see shootings and people are dying. So, it's not easy.*

These participants hoped that their stress would be relieved after taking refuge in the United States, but talking to their families and seeing and hearing about their ordeals resulted in increased emotional distress. Hearing troubled and fearful voices of their relatives and seeing videos of their houses and streets through apps, such as WhatsApp, Skype, and Viber, reminded many participants of the difficulties of their past and the difficulties their loved ones still faced.

Helplessness

Several participants described how helpless they felt when they talked to their family members. They relayed that talking to their separated family members was disturbing, as they could talk to their families but could not help them out. For example, one of them, a woman from Syria commented.

The interviewer asked the following question:

> *How often are you in contact with your mom and your sister?*

The participant replied by saying the following:

> *If I cannot contact her or call her within two days I'll be crazy. She will be crazy, too.*

Then, the interviewer asked the following question:

> *How does it impact your well-being to be in touch with your family?*

The participant responded by saying the following:

> *Of course, it's helping me a lot contacting her every day. Now she's sick; she cannot leave the bed for two months now. I cannot go to help her. Of course, talking to her is kind of helping, but not going to her, being with her, serving her, taking care of her, as a sick mom, and when I went to her last time, she was sick also, but when she saw me she kind of refreshed herself, she was fine, she was happy. I wish I can go, but I'm really afraid. If I have a little money, I will go, but I'm really afraid that might hurt my status as a refugee or my green card. I don't know. But yeah, that, of course, will help, and of course talking to her every day is really kind of helping.*

Several other participants discussed feelings of helplessness after talking to their relatives. For example, a woman from Iraq said the following:

> *When I speak to my family back home, I really get very worried about my mother especially because my mother is all the time telling me that she is sick because I am away from her and she is really worried about me. All the time, she would tell me that 'You need to come back to Iraq no matter how difficult it is over here in Iraq. You will just live the way we live.' I get concerned about what she's telling me, that I lose my focus. Even when I take my English classes, my mind is all the time with my family members back home, and I all the time think about it.*

The messages communicated in these exchanges were stressful for participants and increased the emotional distress of participants. The participants said that even though the ICTs helped them keep in touch with their families, the technology could not replace face-to-face communication, as their inability to physically care for their families was distressing.

Refugee Agency and Information and Communication Technologies

Given the positive and negative impacts of ICT use on refugee mental health, it is important to understand how refugees exercise agency (within the context of opportunities and constraints that shape their access to ICTs) to protect their well-being.

We found that participants' agency around ICT use was evident, as they satisfied their need to maintain relationships with their family members, including whether to remain in touch and how often. However, the use of ICTs also had unintended effects on refugees' mental health. For example, participants described their own agency with respect to controlling communication from their side, but they could not control messages they received from their family members. Several participants said that the communication process with their separated family members through digital technology was mentally and emotionally draining.

Communication to Hide Adversities

Interestingly, participants discussed the ways in which they used different ICTs to hide adversities from their family members. Some participants spoke about how they did not reveal their actual living conditions to their family members when they used digital technologies to communicate with their separated families. They described relatives thinking that they were living lavishly in the United States. However, family members from the country of origin were not familiar with the everyday struggles the participants faced in the new country. Many refugees covered up their struggles so as to not cause family members to worry. Strategic use of some ICTs over others made it possible for them to hide things they did not want to share. For example, a man from Iraq described how he used different communication methods:

> *Sometimes I talk to them by phone, so they just hear my voice. Sometimes I'm talking with them by the Facebook, as I said before, without any picture, but remember that whenever I'm talking to them like that, I pretend that I'm in the restroom or in the bathroom or something like that, to hide myself. My sister called me a few days ago and by the Viber. Have you heard about the Viber?*

The interviewer asked the following question:

> *The app?*

The participant responded by saying the following:

> *Exactly, because the picture is just like a video record, so I told my wife to talk to my sister and tell her I'm not in the home right now.*

Then, the interviewer added the following:

> *I'm sorry. So how does that impact you and how you feel when you talk to your family?*

The participant responded by saying the following:

> *Actually, I feel sad because they used to see me in a better situation, healthy. And if they see me like that right now I will feel very sad and uncomfortable, so that will affect my situation, my health.*

The same feeling was shared by other participants, who explained that they did not share current difficulties with their families to protect their families and not cause worry, as their families were not able to help them at that time. Participants worried that telling their family members about the hardships they faced in the United States as refugees would further harm the physical and mental health of their faraway family members.

In summary, participants exercised agency to control the communication process with their separated families on the basis of the perceived outcome of the communication. Some participants did not share the problems they were facing in their country of refuge. For example, several participants described avoiding using phone cameras in their communication with their families to make their families believe that they were living a good life and to protect their families from seeing their stress. Agency in the use of ICTs reflects the differing ways participants tried to control the impact of communication with family and friends on their mental health.

Financial and Logistical Barriers Restricting Access to Information and Communication Technologies

Financial and logistical issues limited some refugees' agency with respect to their access to ICTs. Refugees had mixed experiences with ICTs in terms of the impact on their mental health and well-being; in addition, refugees' access to ICTs also varied greatly. In particular, we found that certain financial and logistical issues restrained some refugees' access to ICTs. A Burundian man described how his financial constraints limited his communication with his family:

> *Yeah, we are not in communication yet because we don't have money to buy a phone card (the way that we call over there) because we don't have the money yet.*

Some participants could not contact their families through the internet because of poor internet infrastructure in their home countries.

The interviewer asked the following question:

> *How often do you talk or do you Skype or Viber or whatever?*

The participant from Afghanistan responded by saying the following:

> *Yeah, calling my family members and things it's very hard because we have to make a phone call. I guess a $5 [prepaid calling card lasts for] 15 minutes, something like that you can talk through a phone. But you know, in Afghanistan not everybody got access to the internet. The phone technology...kind of new there. So some of them have...Facebook messenger, and they contact just with one. My brother, he got the internet on his phone. The only other people they have a net connection in their phone, those are interpreters, so I'm in contact with them.*

A participant from Afghanistan said that when she and her family arrived in the United States, they did not know how to contact their family members through internet or phone. However, she subsequently learned how to get in touch with them:

> *Yeah, it's very good for me when I talk to them and the first we didn't know where to buy telephone cards, or we didn't have the internet. It was so difficult, and I cried every night, that I couldn't talk to the people there.*

The time difference between the United States and the refugees' home countries also served to limit communication among family members. A man from Iraq said that the time difference and the time constraints (because of his work) had caused strain between him and his family.

The interviewer asked the following question:

> *And how often do you talk to them?*

The participant replied by saying the following:

> *When I have the time, and before I get to work, I was always in contact with them, and make conversation and ask for them. But with the time, with the work, I get a little bit and little bit and little bit. And they blame me [now for not calling them]. "Why don't you contact us again like before?"*

In summary, our qualitative results demonstrated that refugee participants experienced both positive and negative effects on their mental health from their communication with their families through ICTs; in addition, they attempted to moderate these effects through strategic decisions about their use of ICTs. Furthermore, participants described the lack of financial resources, poor information technology infrastructure in their home countries, and time constraints in the United States as the major hurdles for using ICTs to communicate with their families.

Discussion

Principal Findings

This study used in-depth qualitative interviews with 290 refugees from Afghanistan, Iraq, Syria, and the Great Lakes Region of Africa to understand how participants exercised agency in the usage of ICTs. Furthermore, we explored how their agency was shaped by imagined affordances—users' expectations of ICTs and what actions are available to them as part of the interaction with those technologies that practically shape how users approach ICTs [32]. As such, this study provides important insight into refugees' perceptions of how the use of digital technology or ICTs, such as mobile phones, social media, and other apps, such as WhatsApp and Viber, to keep in touch with family affects their mental health. We also explored the types of logistical problems that refugees face in accessing ICTs.

Our results show that refugees make strategic use of ICTs and that the exercise of their agency was based on the imagined affordances of ICTs or what refugees expected from ICTs. The primary motive for the use of ICTs among refugees was staying in touch with their separated families. ICTs helped them transcend physical boundaries. Our findings also elucidate that ICTs impacted individuals differently. Some participants described communication with their families through ICTs as “relieving,” “relaxing,” “tension-free moments,” “motivating,” and “encouraging”; others described communication with their families through ICTs as “depressing,” “stressful,” and “disheartening.” For some participants, hearing of their families’ struggles caused a feeling of helplessness, which was related to their inability to be there with them. Furthermore, some participants felt the need to keep from family members the stress of their everyday lives to not burden them with the same feeling of helplessness. These dynamics suggest that although ICTs have greatly improved the frequency of contact and the ability to keep in close touch, they may also exacerbate feelings of distress through increasing knowledge of the difficulties family members face, highlighting the feeling that one is not able to intervene to the desired extent. Another interesting aspect of ICTs is the ability participants have to see each other (rather than only hearing each other’s voices). We found that refugees decide whether to make a video call with their families on the basis of the imagined affordances of the video call. For example, we found that for some refugees, this facilitated feelings of closeness and connection, whereas for other refugees, the visual function of some ICTs added a layer of complication to communications when they did not want their family members to see them, as it would make it difficult to hide their poor health, distress, or unfavorable living conditions. In addition, some participants contended that they could not access ICTs because of financial issues and logistical problems.

Theoretical Implications

Our study has implications for an important theory in the field of communication, Uses and Gratification Theory. The theory proposes that media users are active stakeholders in communication processes and that they actively seek media channels and content that gratify their needs [36]. The theory takes into consideration several aspects that are critical to our study, including a focus on the agency of users to choose how often and under what circumstances they communicate through a particular media outlet to gratify their needs. Although social theorists generally recognize agency—the ability to act in response to desire or motivation—as a human universal [37], refugees are often ascribed little or no agency, as the social constraints within which they live can be extreme [22]. Thus, it is important to recognize the spheres in which refugees employ agency in their everyday lives.

Gillespie and colleagues (2018) noted that the current application of Uses and Gratification Theory in the context of refugee migration is too simplistic, and they called for investigation of the fluid relationship between technology affordances and user agency in the context of refugee migration [16]. Our study addresses this call by exploring the relationship between user agency and the perceived outcomes of communication through ICTs. Our findings suggest that the exercise of agency by refugees in their interactions with ICTs was dependent on the imagined affordances of the ICTs [16,32]. Importantly, we found that refugees acted to use ICTs to gratify their needs of staying in touch with separated families and friends in their countries of origin, and we found that ICTs enabled them to transcend physical borders. The feelings of happiness among refugee participants who were able to remain connected with their separated families through ICTs emerged as a major theme. Participants discussed the importance of ICTs to their mental health, as it provided them the opportunity to feel the presence of their family members and feel as if they were getting together with their families. Baldassar [26] notes the following (page 161):

> *This might be defined as an ideal of shared co-presence, where the ideal of being there for each other is met with the actual practice of being in touch and the feeling of being in touch enough…Despite the potential pitfalls, the conditions of polymedia appear to provide the capacity to deliver a sense of adequate “distant co-presence” and adequate “distant care.”*

These descriptions are consistent with previous studies on refugees’ use of ICTs [14,15,25,38]. Communication among separated family members also provides refugees with a coping mechanism for the adversities they face in their adopted countries [39]. Copresence, the idea of being emotionally present with family, through the use of ICTs, has recently received much attention from researchers [18,26,39,40]. Baldassar argues that communication among separated family members increases the urge to be physically present with one’s family members. Among immigrants who are able to travel freely back to their home countries, they are able to fulfill their desire to be physically present with family members and are thus able to both physically and emotionally decrease the distance between themselves and their families.

However, in the case of most refugees, the inability to return home, whether because of ongoing conflict in their home country or policy restrictions that impede travel from their resettlement country, may result in communication through ICTs causing emotional distress. Unlike other documented immigrants who can travel to their home countries by choice, refugees often cannot visit their parents or other family members. They have been forced out of their countries; thus, their urge to be physically present with their family members can increase stress and depression.

In addition, our findings highlight the constraints on refugee agency in the use of ICTs. Although some refugees managed to take refuge in safe countries, others remained in conflict areas. For some participants, communication through the use of ICTs was keeping the memories of the trauma they had been through alive. Their communication was mostly about war and the fear of losing a family member to that war. Refugees employed agency in their choice of ICT channels to communicate with their families and friends. For example, some of our participants narrated how they did not use the *video chat* option while talking to their families to hide their adverse conditions from the families, but they could not control the messages they received from their friends and families. The exchange of messages with their family members was a frequent reminder of the traumatic conditions their families continued to experience, and this kept them immersed, albeit at a distance, in the difficulties they had fled.

Participants also discussed the ways in which ICTs could not replace face-to-face or physical interaction with their families. For example, refugees described how they could not express their feelings of agony and helplessness to their families through ICTs. Baldassar [26] argues that ICTs leave little excuse for not being there emotionally for each other: “the form of co-presence that the constant and varied uses of ICTs deliver has become an expected feature of daily life” (page 156). However, the constant presence of ICTs can also create confusion among separated family members. For example, a person can be online on WhatsApp or Viber 24 hours a day, but he or she might not be able to attend a call or reply to a message because of different reasons, such as a difference in the time zones or obligations, for example, work or school. A participant noted that communicating through ICTs increased the emotional distance between her husband and his family members, as their families could not understand how limited her husband’s time was for communication and as they were not able to understand how busy her husband was in the way they would be if they were in the same geographic locale. The lessening of communication within the family, despite the convenience of ICTs, had an adverse impact on both parties involved in the act of communication.

We found that financial and logistical issues also limited refugees’ ability to communicate with their families through ICTs. Some previous studies have noted that not all refugees have equal access to digital technology [13,28,33]. We did not specifically ask participants about accessing ICTs, but in response to open-ended questions about communication with family and friends, 9 participants talked about difficulties accessing ICTs. A total of 3 participants contended that they did not have enough financial resources to buy mobile phone calling cards, 2 participants said that they were not technologically literate, and 4 participants noted poor ICT infrastructure in their home countries.

Recommendations or Practical Implications

Our study makes significant contributions to the understanding of communication through ICTs for separated refugee families. Our findings suggest some ways that organizations serving refugees in host countries may be able to use ICTs to promote the mental well-being of refugees. First, host countries could arrange tailored ICT usage training for refugee families. These trainings could include teaching them the use of different ICTs. These types of training could help bridge the digital literacy divide among refugees. In addition, refugee service providers and refugees themselves could be made aware of both the potential benefits and challenges of frequent communication with family members through the use of ICTs.

Host organizations should also understand the importance of helping recently arrived refugees access the internet or mobile services, at least during the early phases of their resettlement. The provision of free internet or mobile phones would enable refugees to communicate with their families. The provision of ICTs could also help refugees better integrate into their host countries. As some previous studies have shown, refugees not only use ICTs to communicate with their families but also use these tools to access different social services and learn about the political and cultural environments of their host countries [25,30,31]. The development of a mobile app specifically for refugees might also be helpful. The mobile app could contain features such as social support, information about their new country, and how to get access to different social services. Mobile apps should be created in the languages of refugees. The apps should also have audio and video features. The audio and video features could help refugees whose literacy rates are low and who cannot read in their native languages.

Limitations

Even though our study makes a critical contribution to the existing literature on refugee communication, it has a few limitations. We were able to explore how ICTs affected participants’ mental health. However, as we did not specifically and systematically ask refugees about their use of ICTs, we were not able to compare whether refugees with different background characteristics were more likely to have positive or negative experiences with the use of ICTs in terms of the impact on their mental health. Future researchers could use quantitative tools to investigate the relative impact of ICTs on people of different backgrounds and social locations. Another notable limitation is that our study design focused solely on the use of ICTs among refugees for communication with their separated families. Earlier research shows that refugees also use ICTs to get news about their host countries and their country of origin. Obtaining information through different news websites can also affect refugees’ mental health in several ways. Future research could test the impact of Web-based information available to refugees on their well-being. We also did not measure whether age, sex, or education level had any effect on the medium of choice of the participants—such as Facebook, Skype, WhatsApp, Viber, and Mobile phone. Future research could also test or explore whether the type of ICT platform explains any variance in refugees’ mental health.

Conclusions

ICTs play a critical role in bridging the communication gap among separated refugee families. This study sought to explore how refugee adults from Afghanistan, Iraq, Syria, and the Great Lakes region of Africa (Burundi, Democratic Republic of Congo, and Rwanda), who were resettled in the United States for less than 3 years, perceived the effect of ICT communication with their separated families on their mental health and well-being. We found that the use of ICTs has both positive and negative impacts on the mental health of refugees and that refugees recognize this and thus make strategic decisions about how much, when, and which ICT to use to communicate with distant family members. Several refugees in our study also faced strategic and financial problems in communicating with their separated families through ICTs. These findings highlight the complexity of ICT use among refugees, both in terms of how it can ease or exacerbate the distress caused by family separation for people who are forcibly displaced and in terms of inequities in access to ICTs.

Acknowledgments

The authors thank all research participants and student advocates and the following people for their contributions to this research. Research Team: Deborah Bybee, Brian Isakson, Brandon Baca, Suha Amer, R. Neil Greene, Ryeora Choe, Martin Ndayisenga, Mahbooba Pannah, Suha Amer, Charlisa Christian, Amy Armstrong, and Eric Ndaheba. Community Advisory Council Members: Habiba Alkozai, Malalai Alkozai, Rahilah Alkozai, Hala Al-kurdi, Muslim Al-kurdi, Larry Buelow, Sharief Hadi, Marshall Jenson, Desire Kasi, Antoinette Kasi, Kiri Mathsen, Nina Nahimana, Jerome Ndabirorere, Freedance Nibakiza, Jean Paul Ninziza, Ngerina Nyankundwakazi, Kresta Opperman, Mandy Ortaa Danielle Parker, Rosa Reyes, Seyyed Qasim Sadat, and Sadiqa Sadat. Interviewers/Interpreters/Research Assistants: Safaa Abid, Hadeer Albazzaz, Ebtisam Ali, Abdulamir Alibrahimi, Motasem Aljubury, Mohammed Alkawaz, Sulaf, Alkawaz, Rana Al-khafaji, Hala Al-kurdi, Muslim Al-kurdi, Sahib Dad Ahmadi, Sayyed Fawad Ali Shah, Odai Amer, Salim Ansari, Angelica Archuleta, Mika Armenta, Amy Armstrong, Warood Arrawi, Samia Ayoubi, Karim Bakhsh, Venes Barlas, Jodi Beers, Chanda Begin, Kaveri Bisht, Ariana Burks, Alexandra Cervantes, Mather Cotter, Emily Czajkowski, Yousif Dabbach, Yuka Doherty, Jennifer Espinosa, Mariah Everett, Abdul Fahim, Mallory Fallin, Julieta Ferrera, Ashley Flores, Marilyn Garcia, Kayla Garvey, Julia George-Jones, Noha Ghaly, Sharief Hadi, Alexandra Haigh, Caroline Hanawalt, Alyssa Herrera, Courtney Howe, Yasir Hussain, Parwin Hussaini, Alaa Hussein, Mohammed Ismail, Peace Izabayo, Rick Jackson, Maribel Jauregui, Lauren Jones, Kerstin Kalke, Maggie Kentilitisca, Arif Khan, Haley Kohn, Kelly Kourouma, Sonam Lama, Kaleka Lukusa, Sherry MacKay, Arianna Martinez, Ahmad Manshad, Ahmed Mashhadi, Larissa Messier, Cyntia Mfurakazi, Stephanie Mladinich, Chance Najera, Peter Njagi, Jemima Organ, Lorraine Pacheco, Shafiqa Pannah, Danielle Parker, Jennifer Perry, Lorah Plante, Amber Rateau, Rose Rohr, Angelica Romero, Faustin Rwinga, Sadaf, Babak Shahsiah, Cece Shantzek, Ammar Sheraad, Sahar Shirzada, Sara Siyavoshi, Casey Smith, Claire Stasiewicz, Chonour Varyani, Angelica Velasquez, Sage Vogel, Karen Wang, Sanaa Yaqoob, and Asma Yusufi. This study was funded by a grant to the last author from the National Institute on Minority Health and Health Disparities (R01MD007712).

Conflicts of Interest

None declared.

References

1. United Nations High Commissioner for Refugees. 2016. Global Trends: Forced Displacement in 2015 URL: http://www.unhcr.org/en-us/statistics/country/576408cd7/unhcr-global-trends-2015.html [accessed 2017-02-01]
2. USA for UNHCR. 2019. Refugee Statistics: Global Trends At-a-Glance URL: https://www.unrefugees.org/refugee-facts/statistics/ [accessed 2019-06-26]
3. Hess JM, Isakson BL, Amer S, Ndaheba E, Baca B, Goodkind JR. Refugee mental health and healing: understanding the impact of policies of rapid economic self-sufficiency and the importance of meaningful work. J Int Migr Integr 2018:1-18 epub ahead of print. [doi: 10.1007/s12134-018-0628-3]
4. Silove D, Mares S. The mental health of refugees and people who seek asylum. In: Benjamin R, Haliburn J, King S, editors. Humanising Mental Health Care in Australia: A Guide to Trauma-Informed Approaches. New York, NY: Routledge; 2019.
5. Rousseau C, Mekki-Berrada A, Moreau S. Trauma and extended separation from family among Latin American and African refugees in Montreal. Psychiatry 2001;64(1):40-59. [doi: 10.1521/psyc.64.1.40.18238] [Medline: 11383441]
6. Gifford SM, Wilding R. Digital escapes? ICTs, settlement and belonging among Karen youth in Melbourne, Australia. J Refug Stud 2013;26(4):558-575. [doi: 10.1093/jrs/fet020]
7. Kabbar EF, Crump BJ. The factors that influence adoption of ICTs by recent refugee immigrants to New Zealand. Inform Sciences 2006;9:111-121. [doi: 10.28945/475]
8. Wasserman H, Kabeya-Mwepu P. Creating connections: exploring the intermediary use of ICTs by Congolese refugees at tertiary educational institutions in Cape Town. Afr J of Inf Commun 2005;6:94-103. [doi: 10.23962/10539/19808]
9. Miller A, Hess JM, Bybee D, Goodkind JR. Understanding the mental health consequences of family separation for refugees: implications for policy and practice. Am J Orthopsychiatry 2018;88(1):26-37 [FREE Full text] [doi: 10.1037/ort0000272] [Medline: 28617002]
10. Savic M, Chur-Hansen A, Mahmood MA, Moore V. Separation from family and its impact on the mental health of Sudanese refugees in Australia: a qualitative study. Aust N Z J Public Health 2013 Aug;37(4):383-388. [doi: 10.1111/1753-6405.12088] [Medline: 23895483]
11. Alencar A. Refugee integration and social media: a local and experiential perspective. Inf Commun Soc 2017;21(11):1588-1603. [doi: 10.1080/1369118X.2017.1340500]

12. Andrade AD, Doolin B. Information and communication technology and the social inclusion of refugees. MIS Q 2016;40(2):405-416. [doi: 10.25300/MISQ/2016/40.2.06]
13. Charmarkeh H. Social media usage, tahriib (migration), and settlement among Somali refugees in France. Refuge 2013;29(1):43-52 [FREE Full text]
14. Dekker R, Engbersen G, Klaver J, Vonk H. Smart refugees: how Syrian asylum migrants use social media information in migration decision-making. Soc Media Soc 2018;4(1):1-11. [doi: 10.1177/2056305118764439]
15. Gillespie M, Ampofo L, Cheesman M, Faith B, Iliadou E, Issa A, et al. The Open University: Distance Learning Courses and Adult Education. 2016. Mapping Refugee Media Journeys: Smartphones and Social Media Networks URL: https://www.open.ac.uk/ccig/sites/www.open.ac.uk.ccig/files/Mapping%20Refugee%20Media%20Journeys%2016%20May%20FIN%20MG_0.pdf
16. Gillespie M, Osseiran S, Cheesman M. Syrian refugees and the digital passage to Europe: smartphone infrastructures and affordances. Soc Media Soc 2018;4(1):1-12. [doi: 10.1177/2056305118764440]
17. Alencar A, Kondova K, Ribbens W. The smartphone as a lifeline: an exploration of refugees' use of mobile communication technologies during their flight. Media Cult Soc 2018:1-17 epub ahead of print. [doi: 10.1177/0163443718813486]
18. Robertson Z, Wilding R, Gifford S. Mediating the family imaginary: young people negotiating absence in transnational refugee families. Global Netw 2016;16(2):219-236. [doi: 10.1111/glob.12111]
19. Almohamed A, Vyas D. Designing for the Marginalized: A Step Towards Understanding the Lives of Refugees and Asylum Seekers Provocations and Work-in-Progress DIS. In: Proceedings of the 2016 ACM Conference Companion Publication on Designing Interactive Systems. 2016 Presented at: DIS'16; June 4-8, 2016; Brisbane, Australia. [doi: 10.1145/2908805.2909415]
20. Walker R, Koh L, Wollersheim D, Liamputtong P. Social connectedness and mobile phone use among refugee women in Australia. Health Soc Care Community 2015 May;23(3):325-336. [doi: 10.1111/hsc.12155] [Medline: 25427751]
21. Kaufmann K. Navigating a new life: Syrian refugees and their smartphones in Vienna. Inf Commun Soc 2018;21(6):882-898. [doi: 10.1080/1369118X.2018.1437205]
22. Wilding R. Mediating culture in transnational spaces: an example of young people from refugee backgrounds. Continuum 2012;26(3):501-511. [doi: 10.1080/10304312.2012.665843]
23. Dahya N, Dryden-Peterson S. Tracing pathways to higher education for refugees: the role of virtual support networks and mobile phones for women in refugee camps. Comp Educ 2017;53(2):284-301. [doi: 10.1080/03050068.2016.1259877]
24. Wall M, Otis CM, Janbek D. Syrian refugees and information precarity. New Media Soc 2017;19(2):240-254. [doi: 10.1177/1461444815591967]
25. Leung L. Phoning home. Forced Migr Rev 2011;38:24-25 [FREE Full text]
26. Baldassar L. De-demonizing distance in mobile family lives: co-presence, care circulation and polymedia as vibrant matter. Global Netw 2016;16(2):145-163. [doi: 10.1111/glob.12109]
27. Xu Y, Maitland C. Communication Behaviors When Displaced: A Case Study of Za'atari Syrian Refugee Camp. In: Proceedings of the Eighth International Conference on Information and Communication Technologies and Development. 2016 Presented at: ICTD'16; June 3-6, 2016; Ann Arbor, MI, USA. [doi: 10.1145/2909609.2909642]
28. Leung L. Availability, access and affordability across 'digital divides'. Aust J Telecommun Digit Econ 2014;2(2):X1-X5. [doi: 10.7790/ajtde.v2n2.38]
29. Xu Y, Maitland CF. Communicating Abroad: A Case Study of Za'atari Syrian Refugee Camp. In: Proceedings of the International Telecommunications Society. 2015 Presented at: ITS'15; October 25-28, 2015; Los Angeles, CA.
30. Loh T. Digitizing refugees: the effect of technology on forced displacement. Commun Cult Technol 2016;XVI(II):1-10.
31. Levy I. ThinkIR: The University of Louisville's Institutional Repository. 2013. Talking Across Borders: Information and Communication Technology Use Among Iraqi Refugees in Amman, Jordan, and Louisville, Kentucky URL: https://ir.library.louisville.edu/honors/97/
32. Nagy P, Neff G. Imagined affordance: reconstructing a keyword for communication theory. Soc Media Soc 2015;1(2):1-9. [doi: 10.1177/2056305115603385]
33. Alam K, Imran S. The digital divide and social inclusion among refugee migrants: a case in regional Australia. Inf Technol People 2015;28(2):344-365. [doi: 10.1108/ITP-04-2014-0083]
34. Cresswell JW, Clark VL. Designing and Conducting Mixed Methods Research. Second Edition. London, UK: Sage Publications; 2011.
35. Charmaz K. Constructing Grounded Theory. London, UK: Sage Publications; 2014.
36. Rubin AR. Uses-and-gratifications perspectives on media effects. In: Bryant J, Oliver MB, editors. Media effects: Advances in theory and research. New York: Routledge; 2009.
37. Ruggiero TE. Uses and gratifications theory in the 21st century. Mass communication & society 2000;3(1):3-37. [doi: 10.1207/S15327825MCS0301_02]
38. Zijlstra J, Liempt IV. Smart (phone) travelling: understanding the use and impact of mobile technology on irregular migration journeys. Int J Migr Border Stud 2017;3:174-191. [doi: 10.1504/IJMBS.2017.083245]
39. Brown RH. Multiple modes of care: internet and migrant caregiver networks in Israel. Global Netw 2016;16(2):237-256. [doi: 10.1111/glob.12112]

40. Baldassar L. Missing kin and longing to be together: emotions and the construction of co-presence in transnational relationships. J Intercult Stud 2008;29(3):247-266. [doi: 10.1080/07256860802169196]

Abbreviations

ICTs: information and communication technologies
NET: Narrative Exposure Therapy
PTSD: posttraumatic stress disorder
RWP: Refugee Well-being Project
UNHCR: United Nations High Commission for Refugees

Edited by G Eysenbach; submitted 03.04.19; peer-reviewed by K Alam, M Gillespie; comments to author 08.05.19; revised version received 03.07.19; accepted 07.07.19; published 03.09.19.

Please cite as:
Shah SFA, Hess JM, Goodkind JR
Family Separation and the Impact of Digital Technology on the Mental Health of Refugee Families in the United States: Qualitative Study
J Med Internet Res 2019;21(9):e14171
URL: https://www.jmir.org/2019/9/e14171/
doi:10.2196/14171
PMID:31482853

Original Paper

The #MeToo Movement in the United States: Text Analysis of Early Twitter Conversations

Sepideh Modrek[1,2], PhD; Bozhidar Chakalov[2], MA

[1]Health Equity Institute, San Francisco State University, San Francisco, CA, United States

[2]Economics Department, San Francisco State University, San Francisco, CA, United States

Corresponding Author:
Sepideh Modrek, PhD
Health Equity Institute
San Francisco State University
1600 Hollaway Avenue HSS 386
San Francisco, CA, 94132
United States
Phone: 1 415 405 7556
Email: smodrek@sfsu.edu

Abstract

Background: The #MeToo movement sparked an international debate on the sexual harassment, abuse, and assault and has taken many directions since its inception in October of 2017. Much of the early conversation took place on public social media sites such as Twitter, where the hashtag movement began.

Objective: The aim of this study is to document, characterize, and quantify early public discourse and conversation of the #MeToo movement from Twitter data in the United States. We focus on posts with public first-person revelations of sexual assault/abuse and early life experiences of such events.

Methods: We purchased full tweets and associated metadata from the Twitter Premium application programming interface between October 14 and 21, 2017 (ie, the first week of the movement). We examined the content of novel English language tweets with the phrase “MeToo” from within the United States (N=11,935). We used machine learning methods, least absolute shrinkage and selection operator regression, and support vector machine models to summarize and classify the content of individual tweets with revelations of sexual assault and abuse and early life experiences of sexual assault and abuse.

Results: We found that the most predictive words created a vivid archetype of the revelations of sexual assault and abuse. We then estimated that in the first week of the movement, 11% of novel English language tweets with the words “MeToo” revealed details about the poster’s experience of sexual assault or abuse and 5.8% revealed early life experiences of such events. We examined the demographic composition of posters of sexual assault and abuse and found that white women aged 25-50 years were overrepresented in terms of their representation on Twitter. Furthermore, we found that the mass sharing of personal experiences of sexual assault and abuse had a large reach, where 6 to 34 million Twitter users may have seen such first-person revelations from someone they followed in the first week of the movement.

Conclusions: These data illustrate that revelations shared went beyond acknowledgement of having experienced sexual harassment and often included vivid and traumatic descriptions of early life experiences of assault and abuse. These findings and methods underscore the value of content analysis, supported by novel machine learning methods, to improve our understanding of how widespread the revelations were, which likely amplified the spread and saliency of the #MeToo movement.

(J Med Internet Res 2019;21(9):e13837) doi:10.2196/13837

KEYWORDS

social media; sexual abuse; sexual assault; machine learning; infodemiology; infoveillance

Introduction

Public discourse on sensitive topics, ranging from sexual violence to health (mis)behaviors, increasingly occurs on social networking platforms such as Twitter [1-3]. Public health officials and social scientists alike are turning to Twitter to better understand who participates in these conversations, what new information may be gleaned, and the reach of online messages. In health fields, studies of the content of social media posts have a variety of aims, including producing new data to support surveillance efforts, predicting onset of a variety of conditions, and targeting interventions [4]. Detailed content analyses of posts, above and beyond search frequencies and hashtag analysis, have examined health issues such as influenza, allergies, and a variety of mental health conditions (depression, postpartum depression, eating disorders, etc) [5-10]. Likewise, social scientists have conducted detailed content and network analysis of social media posts to understand salient topics in particular networks (Black Twitter and online Feminism) and the reach of salient topics [11-14].

The recent explosion in the public discourse on sexual violence is an interesting case to consider. Sexual violence, including sexual harassment, abuse, and assault, is highly pervasive and has long-term behavioral and mental health sequelae [15]. In the US, one in three women experience unwanted sexual contact in their lifetime [16]. Although hashtag movements such as #BeenRapedNeverReported had traction and encouraged public disclosures of personal experiences of rape [12], these initial movements were relatively small. The public discourse on sexual violence changed substantially on October 15, 2017, when actress Alyssa Milano called her followers to post “MeToo” if they had ever experienced unwanted or inappropriate sexual contact [5].

Ms Milano’s tweet immediately went viral, with 1,595,453 tweets posted in the first week, and ignited a movement where victims of sexual assault, abuse, and harassment felt empowered to divulge as much, or as little, information as they wanted about their personal experiences. The phrase “MeToo” was coined by Tarana Burke, a civil rights activist, as a way to raise awareness and provide support for survivors of sexual violence. Ms Milano’s use of this phrase in her tweet, as opposed to other more explicit hashtags such as #BeenRapedNeverReported, allowed posters to retain some privacy about the details of the event and still participate. This led to millions of users joining the conversation and subsequent normalizing of the revelations. Meanwhile, the wide use of the generic #MeToo encouraged an outpour of detailed revelations. Because of the massive size of the #MeToo movement, many who may not have experienced sexual violence were confronted with the knowledge that members of their network had.

In this study, we aim to describe the public disclosures of sexual violence within the first week of the #MeToo movement. Here, we document the content, quantify the scale, and present the demographic characteristics of Twitter users who disclose incidents of sexual assault/abuse in the early conversation on Twitter for the #MeToo movement. We use simple machine learning tools to create the archetype of tweets, which often included detailed accounts of sexual assault and abuse and early life experience of such events among women from all walks of life. Next, we categorize the content of individual-level tweets to estimate the proportion of all #MeToo tweets with such revelations in the first week of the movement. Furthermore, we use our categorization to detail the demographic characteristics of posters with revealed events and the reach of the revealed events on the Twitter platform generally.

Methods

Data

The data for this project are tweets, short messages of 140 characters or less, sent from a Twitter user (Twitter handle) to their network of followers. Twitter data are considered existing data in the public domain and therefore exempt from human subject review.

We applied for and were granted access to the Twitter Premium application programming interface (API) platform, which allows users to purchase and query all nondeleted public tweets since the first tweet posted in 2006. We purchased both the counts and full tweets from historical Twitter data between October 14 and 21, 2017 (ie, the first week of the movement). Unlike other social media movements that took several months to take hold, the #MeToo movement had the greatest activity in the first week (Figure 1), which is one reason we limited our data collection to this period. For our analysis, we concentrated only on novel, or user generated, English language tweets with “MeToo” in the text. Novel tweets exclude replies to other’s tweets, retweets without comments, and links to other sites or images. This was chosen to capture posts that were tied to the specific user and that would be available to all of one’s followers. We further limit the tweets to those with geotagged information, placing the tweet in the United States (N=12,337; this count is subject to some variability, as it depends on the day of the query and only includes nondeleted tweets).

Figure 1 shows the daily counts of novel English Language tweets starting from one day prior to the creation of each hashtag on Twitter (July 13, 2013, for BlackLivesMatter and October 14, 2017, for MeToo). The counts derived from the Twitter Premium API vary by the date of query. These counts were extracted on August 14, 2018.

The analytic sample for our content analyses comprised the full text and associated metadata for 97% of novel English language US-based tweets in the study period (N=11,935). Figure 2 presents a flowchart of the filtering process on the sample of tweets captured and analyzed (see Multimedia Appendix 1 for details of the selection process).

Figure 1. Comparison of relative time patterns of novel English tweets including MeToo and BlackLivesMatter. MeToo counts are on the left axis, and BlackLivesMatter counts are on the right axis. BLM: BlackLivesMatter.

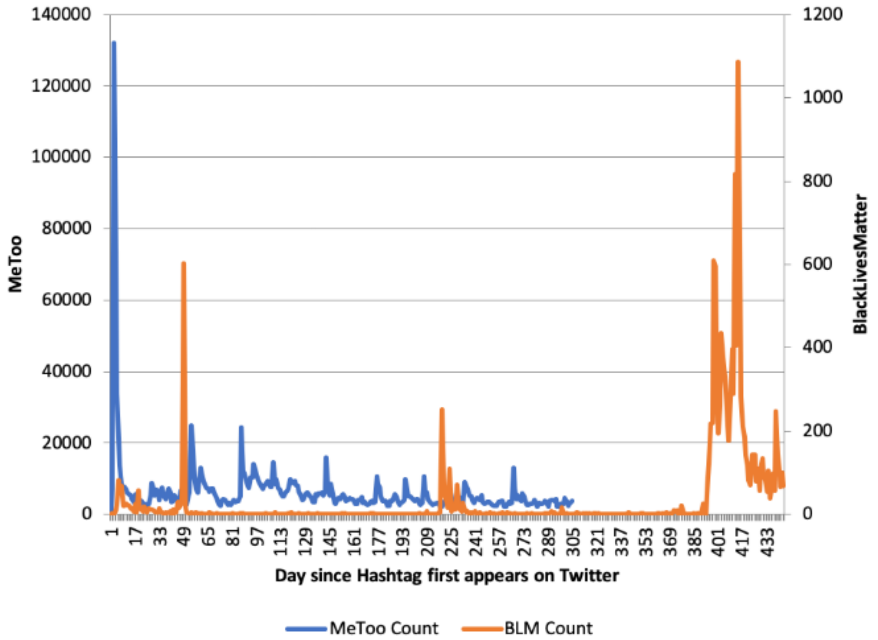
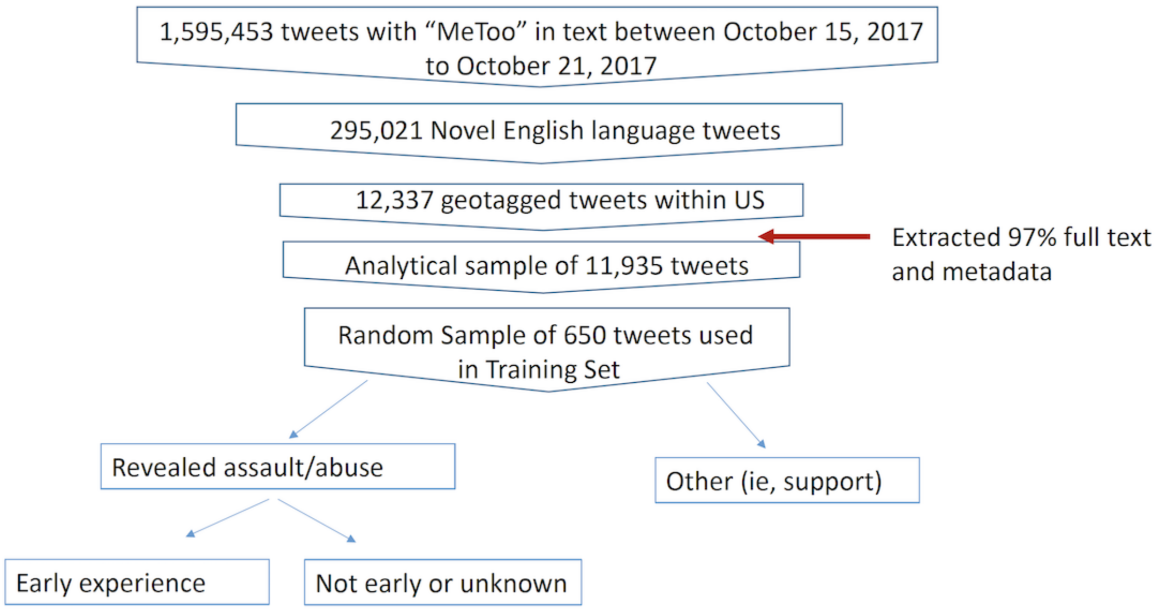


Figure 2. Data flow chart. LASSO: least absolute shrinkage and selection operator; SVM: support vector machine.



Analyses

We first present evidence that the sample of novel English #MeToo tweets in the United States, which we selected for analysis, was comparable in terms of time trends to all novel English #MeToo tweets. We graphed the number of #MeToo tweets per hour starting on October 14, to show the time trends comparing all #MeToo tweets, novel English #MeToo tweets, and US-based geotagged novel English language #MeToo tweets. We also report quantitative measures of correlation on the number of tweets in each of these categories by hour (the Pearson correlation and coefficient of determination).

Annotation Process

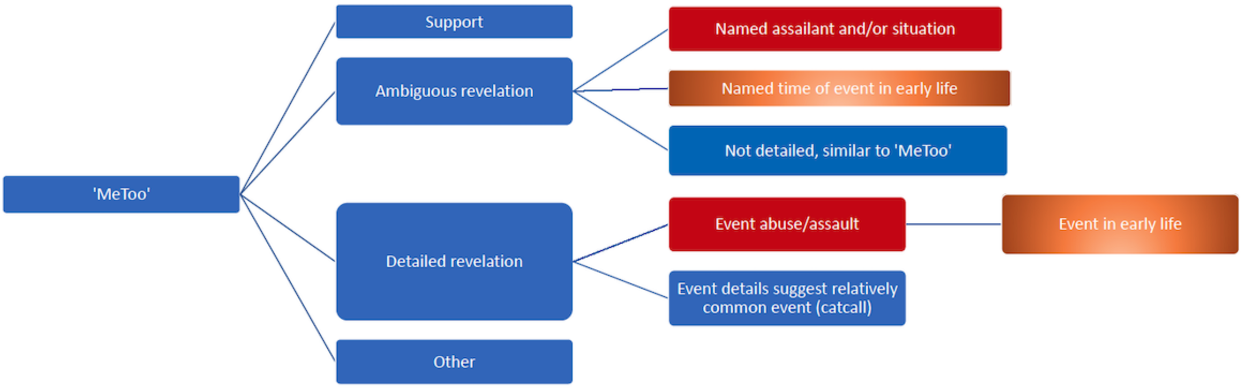
Before analyzing the data, both authors reviewed approximately 2000 tweets from the first week of the movement. Based on this initial review, we chose to focus the analysis on first-person revelations of sexual assault and abuse and childhood experience of sexual assault and abuse. The tweets reviewed revealed several categories of comments with #MeToo hashtags. They included (1) tweets with #MeToo that were support statements (ie, "How many women not in the spotlight have #MeToo #Notaceleb #StillAStar. Proud of them! As Bosch says- 'everybody counts or nobody counts'"), (2) statements with ambiguous revelations of events (ie, “MeToo” alone and “Kept it buried down for many years and didn't even really realize how it impacted me.”), (3) statements with detailed revelations of events (ie, “I was sexually assaulted by a family member when I was 8 the trauma *[sic]* it causes never ends, remember that before you ruin someone's life”), and (4) others (ie, either negative comments, unrelated content, or riding off the hashtag; “Not watched any games this weekend @NFL @nflcommish @nflnetwork #metoo didnt buy nfl merchandise this year Got refund for #NFLSundayTicket”).

Based on this initial review, we chose to focus on the second and third categories and created an annotation rubric to categorize first-person revelations. For ambiguous statements (ie, category 2) where it was clear an event occurred, such as a revelation with an alleged assailant’s name or a situation but little details on the actual actions, we decided that this recollection likely reflects enduring trauma and should be categorized as a revealed abuse or assault. If the ambiguous events were experienced in childhood, it would be classified as an early life experience of abuse or assault (“I was a child too scared to speak. Till this day, it still haunts me”). If the statement was too ambiguous, did not make references to childhood, or provided little detail, we did not classify it as a case of abuse (ie, “Sometimes the stories we don't share are the ones that affect us and continue to scare us the most” and “Tried to tweet about my #metoo moment and deleted it bc it still feels like my fault. The voices you aren’t hearing in this are deafening”). For the third category—the clearest case—based on the purported event, we categorized the revelations as abuse or assault and early life experience based on the details provided. Figure 3 provides a schematic of our annotation process.

We randomly selected a subset of 650 tweets from the novel English US-based #MeToo tweets as the training set for the analysis. Based on our rubric, both authors classified the 650 randomly selected tweets. Each tweet was categorized along two main dimensions: if the tweet revealed details of an experience of sexual assault and abuse and if the details suggest that the event happened in early life. All other types of tweets (support, too ambiguous, or other) were categorized as not revealing an incident of abuse or assault. We used an approximate age cut off of 22 years (or references to college or earlier schooling) to distinguish and delineate early life experience. This was chosen to capture college experience under the category of early life. All other tweets were categorized as “other.” The concordance between the two authors’ categorizations was 94% for sexual abuse and assault and 98% for childhood experience. Most disagreements on the abuse/assault categorization were on cases with ambiguous revelations, which could be argued either way. Most disagreements on childhood experience were in cases where the revealed event was placed in the past. Many cases point to an event in the distant past, but it was hard to ascertain the age of the poster at the time of the revealed event, because we did not have the current age of the poster. Given these minor differences in interpretation, we decided to use one author’s (BC) categorization in the training set. To assess our model performance, we used the other author’s (SM) classification in the test set for calculating positive and negative predictive values.

Tweets were categorized along two dimensions: if they disclose an experience of sexual assault and abuse (red in Figure 3) and if the details suggest that the event happened in early life (orange in Figure 3). All other tweets were categorized as “other.” Multimedia Appendix 2 provides examples of the categorizations performed by the authors, which were guided by the categorizations in the “Rape Culture Pyramid” graphic version 1 created by Ranger Cervix and Jaime Chandra in mid-2016 [17].

Figure 3. Classification flow chart.



Prediction Methods

To provide exemplars of the disclosed tweets, we used simple supervised machine learning methods— least absolute shrinkage and selection operator (LASSO) regression models—on the training sample to find most predictive words for classification along both dimensions [18]. This allows us to present the most predictive words, which serve as an archetype of the tweets with such revelations.

LASSO regressions are a common tool for economists who are using text as data and performing simple computational linguistic analyses [19]. The LASSO, a penalized linear model with L1 penalization, serves as a shrinkage method to help perform feature/word selection to identify the most predictive words from a list of candidate words in a supervised learning environment. These methods are recommended because they are intuitive and interpretable. LASSO regression was chosen over ridge regression, another penalized linear models with L2 penalization, because it is more efficient for variable selection. Furthermore, LASSO was chosen over elastic net regression, which is sometimes preferred when there is substantial correlation between features/words, because we found limited correlation in the words in the training set data.

After removing stop words, we had 11,931 unique words in our training set from 650 tweets. We stemmed each word, examined the list for misspelling, and considered words (and stubs) that appear in at least five tweets to limit the sparsity of our sample, leading to 1186 unique words. From these words, the authors selected 109 words related to sexual abuse and assault. This filtering limited the sparsity of the data further. We then used LASSO regressions on a matrix of 650 tweets coded for sexual assault/abuse (or coded for early life experience) and 109 words, each of which was treated as a binary flag if it appeared in each tweet. For example, if the tweet included the word "Rape" or "Raped" in it, then the variable "Rape" was coded as 1.

The LASSO model minimizes an objective function , which is a constrained ordinary least squares model (OLS), to find the words with the strongest predictive power on tweets coded as having a revelation of sexual assault or abuse.

Equation 1: , where is the L1 Norm of the estimated β coefficients; this is the sum of the absolute value of all the β coefficients, and is the sum of squared deviation of the predicted outcome relative to the actual outcome for each observation. This algorithm is the same as an OLS with an added penalization for large estimated β coefficients. The form of the penalization with an L1 Norm drives the magnitude of many β coefficients to zero. This essentially makes the algorithm identify and select features/words that are most predictive [18].

The LASSO models were implemented in the statistical software R ([computer program] Version 3.5.0. Vienna, Austria: R Foundation for Statistical Computing; 2018) using the "glmnet" package. The results of the LASSO model are validated with ten-fold cross-validation. The cross-validation allows us to pick the model with the λ value having the lowest mean squared error or model variance. From this selected model, we obtained a list of the 35 most predictive words for tweets with revelations of sexual assault/abuse words and 34 most predictive words for tweets with revelations of early life experience of sexual assault/abuse words. These word lists were then categorized by the authors.

We then used the same training sample of tweets to train the support vector machine (SVM) models to classify the remaining sample of tweets (N=11,285) along the same two dimensions—experience of sexual assault/abuse and early life experience of sexual assault/abuse [18,20].

An SVM is a discriminative classifier that is used on training data to define a separating hyperplane in multidimensional space and then uses this hyperplane to categorize new data. In the training set, data points that are closer to a potential defining hyperplane (ie, support vectors) are given more weight. The goal of the underlying algorithm is to maximize the margin between the support vector data points and the separating hyperplane using a cost function [18].

In our case, we have a matrix of 11,931 unique words (dimensions) and 650 labeled outcome data. From this information, the algorithm defines a separating hyperplane. Based on this hyperplane, the remaining 11,285 tweets are categorized. To perform the SVM model operations, the package "RTextTools" was used in R software.

We assessed the quality of the SVM model's classification with four test sets, two for each classification category, to estimate positive and negative predictive values. To calculate the positive predictive value (PPV) and negative predictive value (NPV), we sampled 50 tweets four times as test sets. The first two samples were used to assess PPV and NPV for the model predictions of sexual assault and abuse. The second two samples were used to assess PPV and NPV for the model predictions of early life experience of sexual assault and abuse. Given that BC's classification was used in the training set, we used SM's classification on the test set to assess the model.

For each sample of the first two samples, one author (SM) assessed the content of the tweet for a revelation of an experience of sexual assault and abuse. We treated the human-assessed content as the gold standard and calculated the proportion of the time that the SVM algorithm's classification was the same as the human classification. PPV was calculated as the number of true positives (agreement between human and algorithm) divided by the number of positive cases found by the SVM algorithm. NPV was calculated as the number of true negatives divided by the number of negative cases found by the SVM algorithm. We repeat the process and manually categorize the content of the last two samples for revelations of early life experience of sexual assault or abuse to calculate PPV and NPV

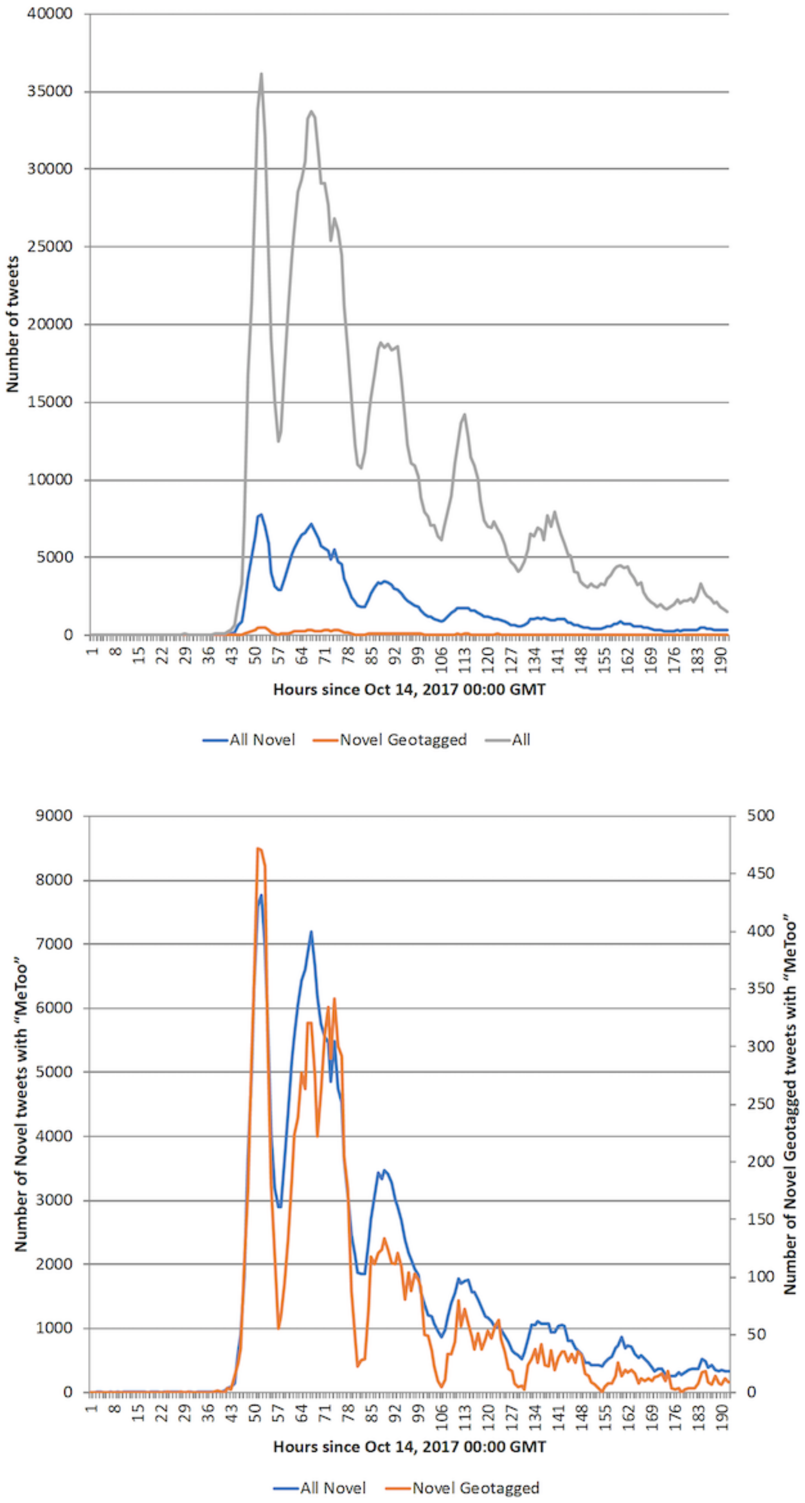
Based on these classifications, we used a previously vetted commercial service, Demographics Pro [21,22], to infer the demographic characteristics of the individuals who revealed an incident of abuse/assault and childhood experience (Multimedia Appendix 1 for description of their algorithm) . We then empirically derived estimates of the 25%-75% range of the number of followers from each poster who revealed an event. This allows us to conduct a crude calculation to capture a lower bound on the reach of the tweets with revelations of sexual assault/abuse in the first week of the movement. This estimate represents the number of potential Twitter users who would have been exposed to revelations of sexual assault/abuse posted by someone they follow.

Results

Figure 4 shows the time trends of #MeToo tweets, illustrating how the movement grew. We presented hourly counts of tweets from October 14, 2017, 00:00 GMT, the day before the movement began. We also compared all #MeToo tweets, the novel English #MeToo tweets, and the novel geotagged English #MeToo tweets—the main sample used in this study—to show that they follow a similar overall time trend in the first week. The Pearson correlation between the counts of hourly #MeToo tweets and novel geotagged English #MeToo tweets was 0.96, and the coefficient of determination was 0.92. The counts derived from the Premium API in Figure 4 vary by the date of query. These counts were extracted on June 12, 2018.

Author BC's classification of the 650 randomly selected tweets in the training set revealed that approximately 19% of these tweets included a first-person revelation of sexual assault/abuse. BC's classification was used in supervised machine learning methods with LASSO regressions to identify the most predictive words. Textbox 1 summarizes and organizes the most predictive words related to revelations of sexual abuse/assault. The most predictive words are organized into descriptions of time, persons, actions, body parts, and words indicating state and other. For example, the predictive actions verbs include "Rape," "Grope," "Grab," and "Shout" among others. Other highly predictive words relate to an intoxicated state ("Drunk" or "Drug"). Many predictive words placed the revealed event in the past, particularly in early life ("Age," "College," "First Time," and "Years Old").

Figure 4. Hourly counts of "MeToo" tweets. Top: Hourly counts of "MeToo" tweets by category (overall, novel English, and geotagged novel English). Bottom: Hourly counts of all novel English language tweets with the phrase "MeToo" and hourly counts of all novel English language geotagged tweets in the United States.



Textbox 1. Words contained in Tweets that consistently predict a revelation of sexual abuse or assault (most predictive words categorized).

Time/age

- Age
- College
- First time
- Fifteen
- Grade
- Hasnumber (Hasnumber is an overall indicator of whether there is a number in the text.)
- Kindergarten
- Old
- Year
- Years ago
- Years old

Person

- Boyfriend
- Coworker
- Date
- Father
- Man
- Rapist
- Police
- Stranger
- Teacher
- Uncle

Action

- Advantage
- Chase
- Grab
- Grope
- Rape
- Shout

Body Part

- Arm
- Boob
- Butt

State

- Drug
- Drunk

Other

- Daylight
- Door
- Interview

- Frat (This term can indicate a location, time/age, or an implied level of intoxication. The term frat can indicate a time (college is when most people go to frats), a place such as a frat house, and a level of intoxication because frats are often where alcohol is served.)

Of the 650 randomly selected tweets in the training set, the authors categorized 5% of the tweets as having indicated an early experience of sexual assault/abuse. Textbox 2 summarizes and organizes the most predictive words related to early experience of sexual abuse or assault. Although many of the predictive terms are the same, the list of persons, or possible assailants, listed for those who had an early life experience include terms like “Neighbor” and “Step-Father.” In addition, the states of being “Asleep” and “Afraid” are highly predictive. Together, these predictive words provide a picture of the types of tweets posted during the early phase of the movement. Note that predictive words are different from a word cloud that captures the frequency of a word. For example, “MeToo” was in every one of the tweets we examined but was not, and could not be, predictive of a disclosure of sexual assault/abuse, given our method. The word “Grope,” on the other hand, was not in every tweet and was a predictive word in tweets with revelations of assault/abuse.

Next, we use the SVM algorithm on the same 650 classified tweets in the training set. Of the remaining 11,285 tweets, the SVM algorithm categorized 1287 tweets (11.4%) as having a revelation of sexual assault/abuse and 657 tweets (5.8%) as having a revelation of early sexual assault/abuse (Multimedia Appendix 3). To assess the validity of the SVM classification, we calculate both the PPV and NPV in comparison with human categorized test sets. The SVM classification for sexual assault/abuse had a PPV of 87%, meaning that there is 87% concordance between the algorithm and our manual assessment of the tweets with revelation of sexual assault/abuse. The NPV was 83%, suggesting that the algorithm may have slightly underreported the number of tweets with revelations of sexual assault/abuse in comparison to the authors’ assessment. For early life experience classification, the PPV is 79% and the NPV is 95%. There was a slightly lower concordance between the algorithm and the authors’ assessment of tweets identified as revealing an early life experience of sexual assault/abuse, but the false negative rate was quite low. Together, these assessments of the algorithm suggest that our algorithm performed well and would be highly congruent with the human-generated classifications.

In Table 1, we present the percentage of daily tweets with revelations of sexual assault/abuse and early life experience of sexual assault/abuse based on the SVM classification. In the first 2 days of the movement (October 15-17), 11%-13% of tweets revealed an experience of sexual abuse or assault and over half of the tweets revealed that the experience of sexual assault/abuse occurred in early life. As the movement carried forward, the revelations of sexual assault/abuse decreased to about 6% of MeToo tweets in the last day of the first week, but the relative proportion with early life experience of these events increased to over 80%, suggesting that more traumatic events were being shared.

Based on the SVM classification, we then used the Demographics Pro prediction service to understand and compare the demographics of posters of sexual assault/abuse during the early #MeToo movement. We present these demographics in Table 2. Based on the results of the SVM model, we identified 1168 unique posters/Twitter handles that revealed an incident of sexual assault/abuse and 612 unique posters that revealed an early life experience of sexual assault/abuse. We shared these Twitter handles with Demographics Pro, who then gave us aggregated predictions on the distribution of the demographic characteristics of these posters. We found that 90% of these Twitter users who posted about their experience of sexual assault/abuse were women. This is congruent with national estimates that 90% of sexual assault victims are women [23]. We also found that white women were overrepresented in the early conversation in our Twitter data with regard to their population in the United States, their proportion as Twitter users, and national estimates of those who report experiencing sexual assault (data not shown). The age distribution also shows that older users of Twitter, aged 25-50 years, disproportionately revealed events.

Finally, from the SVM classification, we used the provided metadata on the number of followers of posters who revealed sexual assault/abuse in the first week of the #MeToo movement to better understand the reach of such revelations. Based on posters’ follower count, we took the 25%-75% range of followers from users with such experiences and calculated a lower bound range on the number of Twitter user who may have seen a first-person revelation. This simplistic calculation provided a range of 5,955,342 to 34,251,628 Twitter users for the reach, which we believe is a substantial lower bound (Multimedia Appendix 3).

We believe the reported reach is an underestimate, because we did not include replies or retweets and our algorithm has a higher false negative rate. We did not examine the network of people who reveal events, which could be highly overlapping and would mean that many users would have seen multiple tweets with revelations. Furthermore, the distribution of followers who posted about sexual assault does not seem to vary substantially from the distribution of followers for Twitter users overall (Multimedia Appendix 4). Finally, we only captured posts on Twitter. Many were also posting #MeToo posts on other social media platforms such as Facebook.

Textbox 2. Words contained in Tweets that consistently predict a revelation of an early experience of sexual abuse or assault (most predictive words categorized).

Time/age

- Age
- College
- First time
- Freshman
- Grade
- Hasnumber (Hasnumber is an overall indicator of whether there is a number in the text.)
- High school
- Kindergarten
- Old
- School
- Years ago
- Years old

Person

- Cop
- Coworker
- Date
- Doctor
- Father
- Male
- Neighbor
- Rapist
- Step father
- Teacher
- Uncle

Action

- Rape
- Rip
- Screw

Body part

- Arm
- Butt
- Pussy

State

- Afraid
- Asleep
- Drunk

Other

- Concert

Table 1. Counts and percent of #MeToo tweets with disclosures of sexual abuse/assault and early experience tweets by date.

Date	Total, n[a]	Abuse/assault, n (%)	Early experience, n (%)
10/15/17	371	43 (11.59)	25 (6.74)
10/16/17	5987	817 (13.65)	420 (7.02)
10/17/17	3174	336 (10.59)	142 (4.47)
10/18/17	1155	113 (9.78)	54 (4.68)
10/19/17	676	57 (8.43)	21 (3.11)
10/20/17	356	31 (8.71)	14 (3.93)
10/21/17	215	14 (6.51)	12 (5.58)

[a] Number of geotagged novel English language tweets in United States.

Table 2. Demographic characteristics of abuse/assault and early life experience samples among unique Twitter users.

Characteristic	US census, %[a]	Twitter overall, %[b]	Abuse/assault sample (N=1168), %[b,c]	Early experience sample (N=612), %[b,c]
Sex				
Male	49.2	45.8	10.6	9.2
Female	50.8	54.2	89.4	90.8
Age (years)				
≤19	25.4	25.02	15.2	13.1
20-24	6.70	45.33	25.5	24.1
25-29	7.10	16.10	20	20
30-34	6.70	7.16	17.8	22
35-39	6.60	2.40	8.1	7.3
40-49	12.5	3.25	9.7	9
50-59	13.3	0.49	2.6	3.7
≥60	21.7	0.25	1	0.8
Race/ethnicity				
White/Caucasian	60.7	78.7	90.7	89.8
Hispanic	18.1	7.6	6.2	6.1
African American	13.4	13.1	2.6	3.3
Asian	5.8	0.6	0.4	0.8
Native American/Pacific Islander	1.5	—[d]	—	—

[a] Age distribution based on 2017 American Community Survey 1-Year Estimates (July 1, 2017).

[b] Proportions provided by Demographics Pro on October 18, 2018.

[c] Based on our classification in the analytical sample of geotagged novel English language tweets in the United States.

[d] Not available.

Discussion

Principal Findings

We conduct the first quantitative text analysis of the content of the early conversation in the #MeToo movement on Twitter, which was the largest hashtag movement on Twitter in 2017 [24] and the largest public discourse on sexual violence. We use machine learning to provide exemplars of disclosures during the first week of the movement. The most predictive words create an archetype of the content of tweets with the revelation of sexual assault/abuse. They include words that we would expect like “grope” and “rape” as well as states like “drunk,” “asleep,” and “afraid.” Based on our models, 11% of novel tweets in the first week of the movement publicly revealed an experience of sexual assault/abuse and 6% revealed an early life experience of sexual assault/abuse. The initial women sharing were predominately white women aged 25-50 years. Tweets from older posters suggest that these experiences were lodged into these women’s memories and were not inconsequential passing events. Notably, African Americans were underrepresented relative to their Twitter presence in this early conversation. Moreover, given the connectivity possible on Twitter, we estimate that between 6 and 34 million Twitter users may have been exposed to at least one such detailed revelation.

Limitations

There are key limitations to this study that should be noted. First, based on financial considerations, we could not extract all novel English language posts during the first week of the #MeToo movement. Instead, we chose to focus on those geotagged within the United States. This restriction made the purchase of almost all the tweets that fit this category possible. To examine the representativeness of these tweets, we examined time patterns relative to all #MeToo posts and novel English language #MeToo posts (Figure 3). This analysis suggests that the timing of novel English language geotagged tweets was not different from novel English language tweets. However, if users who allow Twitter to geotag tweets systematically revealed sexual assault/abuse events at a different rate than those who do not allow Twitter to geotag their tweets, then our estimates may be inaccurate. However, we noted that users would have likely enabled geotagging on Twitter at some time prior to their MeToo tweet. Thus, allowing geotagging is not necessarily related to the MeToo tweet. Second, we include only English tweets on the topic of #MeToo even though there were similar movements in different languages (such as #YoTambien or #BalanceTonPorc), and we did note examine tweets outside the United States. For example UK-based or Canada-based English language #MeToo tweets were not included in our data. Again, there could be systematic differences that would affect our estimates, and thus, we acknowledge that our estimates reflect the conversation in the United States. Third, while we examined the time dynamics of many related or counter hashtag movements happening the same week, such as #HimThough, #Ibelieveyou, #Ihearyou, #Iwillnot, and #howIwillchange, we did not include these in our analyses because these hashtags were in response to the #MeToo posts. The content of tweets with these other hashtags were less likely to have revelations of sexual assault/abuse, but rather just voice support. Future analyses could examine who voiced support rather than who revealed events. Fourth, we did not conduct network analysis examining common retweets, which would have had much greater reach. Our focus was on novel revelations that were rarely retweeted relative to support statements, which were often retweeted. Fifth, we only examined the first week of the movement. This was a deliberate choice because the movement went in many directions afterward, with calls to stop posting traumatic events because they triggered women [25] or demands that men post instead of women having to post and relive their trauma. Finally, many tweets were ambiguous, revealing that something happened but without enough detail to determine if there was abuse or assault. Nonetheless, we used a consistent method based on our reading of thousands of tweets. The NPV and PPV of the trained machine learning estimates were good, based on our annotation. However, a different pair of annotators may have had slightly different estimates.

Conclusions

Despite the noted limitations, our results highlight the magnitude of the mass sharing of personal experiences of sexual assault/abuse, filled with narratives of early life experience, which enabled the spread of the #MeToo movement and had a broad reach across Twitter. Further, these tweets and the archetypes presented here provide rich details to augment existing statistics captured traditionally from survey data and small in-depth studies of sexual assault and abuse survivors. The archetype provides a picture of what the public saw. The candid and revealing statements reminded followers and the public alike of the magnitude sexual violence and how it is often first experienced in early life and hidden. Therefore, the descriptive narratives could be used in public health survey development to assess whether there has been a change or a deeper public understanding of the prevalence, early life experience, and enduring trauma of sexual assault and abuse.

The summary of the content of tweets presented in this study highlights the initial conversation and demographics of participants in the conversation in the early stages of the #MeToo movement. Although our results present a snapshot of the public discourse on sexual violence and the initial participants, future work could examine the content and directions of the national conversation, which has since taken many directions and varies by populations.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Details of the selection process.

[PDF File (Adobe PDF File), 61KB - jmir_v21i9e13837_app1.pdf]

Multimedia Appendix 2

All versus novel tweets: example table.

[PDF File (Adobe PDF File), 284KB - jmir_v21i9e13837_app2.pdf]

Multimedia Appendix 3

Estimates of number, proportion, and reach of MeToo tweets from Oct 14 to 21, 2017.

[PDF File (Adobe PDF File), 9KB - jmir_v21i9e13837_app3.pdf]

Multimedia Appendix 4
Comparison of Twitter followers for abuse/assault and early experience sample against all twitter users.
[PDF File (Adobe PDF File), 183KB - jmir_v21i9e13837_app4.pdf]

References

1. Cavazos-Rehg PA, Krauss M, Fisher SL, Salyer P, Grucza RA, Bierut LJ. Twitter chatter about marijuana. J Adolesc Health 2015 Feb;56(2):139-145 [FREE Full text] [doi: 10.1016/j.jadohealth.2014.10.270] [Medline: 25620299]
2. Eshleman R, Jha D, Singh R. Identifying individuals amenable to drug recovery interventions through computational analysis of addiction content in social media. : IEEE; 2017 Presented at: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); Nov 13-16, 2017; Kansas City, MO, USA URL: https://ieeexplore.ieee.org/document/8217766 [doi: 10.1109/BIBM.2017.8217766]
3. Cavazos-Rehg P, Krauss M, Grucza R, Bierut L. Characterizing the followers and tweets of a marijuana-focused Twitter handle. J Med Internet Res 2014 Jun;16(6):e157 [FREE Full text] [doi: 10.2196/jmir.3247] [Medline: 24974893]
4. Paul MJ, Dredze M. You Are What You Tweet: Analyzing Twitter for Public Health. 2011 Jul 17 Presented at: 5th Int AAAI Conf Weblogs Soc Media ICWSM; July 17, 2011; Barcelona, Spain.
5. Caputi TL, Nobles AL, Ayers JW. Internet Searches for Sexual Harassment and Assault, Reporting, and Training Since the #MeToo Movement. JAMA Intern Med 2019 Feb 01;179(2):258-259. [doi: 10.1001/jamainternmed.2018.5094] [Medline: 30575847]
6. Evans H, Cordova V, Sipole S. Twitter Style: An Analysis of How House Candidates Used Twitter in Their 2012 Campaigns. PS: Political Science and Politics 2014 Apr 14;47(2):454-462. [doi: 10.1017/S1049096514000389]
7. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
8. Aramaki E, Maskawa S, Morita M. Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter. USA: Association for Computational Linguistics; 2011 Jul 27 Presented at: 2011 Conference on Empirical Methods in Natural Language Processing; July 27, 2011; Edinburgh, Scotland, UK p. 1568-1576 URL: http://dl.acm.org/citation.cfm?id=2145432.2145600
9. Pater J, Haimson O, Andalibi N, Mynatt E. “Hunger Hurts but Starving Works”: Characterizing the Presentation of Eating Disorders Online. New York, NY, USA: ACM Press; 2016 Feb 27 Presented at: CSCW '16 Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; Feb 27, 2016; San Francisco, CA, USA p. 1185-1200. [doi: 10.1145/2818048.2820030]
10. De Choudhury M, Counts S, Horvitz E, Hoff A. Characterizing and Predicting Postpartum Depression from Shared Facebook Data. New York, NY, USA: ACM Press; 2014 Feb 15 Presented at: The 17th ACM Conference on Computer Supported Cooperative Work and Social Computing; Feb 15, 2014; Baltimore, Maryland, USA p. 628-638. [doi: 10.1145/2531602.2531675]
11. Freelon D, McIlwain C, Clark M. Beyond the Hashtags: #Ferguson, #Blacklivesmatter, and the Online Struggle for Offline Justice. In: SSRN Journal. Washington, DC: Center for Media & Social Impact; Feb 26, 2016.
12. Keller J, Mendes K, Ringrose J. Speaking ‘unspeakable things’: documenting digital feminist responses to rape culture. Journal of Gender Studies 2016 Jul 28;27(1):22-36. [doi: 10.1080/09589236.2016.1211511]
13. Mendes K, Ringrose J, Keller J. #MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. European Journal of Women's Studies 2018 Apr 29;25(2):236-246. [doi: 10.1177/1350506818765318]
14. Kuo R. Racial justice activist hashtags: Counterpublics and discourse circulation. New Media & Society 2016 Aug 12;20(2):495-514. [doi: 10.1177/1461444816663485]
15. Walsh K, Galea S, Koenen KC. Mechanisms Underlying Sexual Violence Exposure and Psychosocial Sequelae: A Theoretical and Empirical Review. Clin Psychol Sci Pract 2012 Oct 29;19(3):260-275. [doi: 10.1111/cpsp.12004]
16. Houry D, Smith S, Chen J, Basile K, Gilbert L, Merrick M, et al. Centers for Disease Control and Prevention. 2017 May 01. The National Intimate Partner and Sexual Violence Survey (NISVS) | 2010-2012 State Report URL: https://www.cdc.gov/violenceprevention/pdf/NISVS-StateReportBook.pdf [accessed 2019-02-25]
17. Rape Culture Pyramid. 11th Principle Consent URL: https://www.11thprincipleconsent.org/consent-propaganda/rape-culture-pyramid/ [accessed 2019-02-26] [WebCite Cache ID 76UDQ5U4p]
18. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. In: Springer. New York: Springer; 2013.
19. Gentzkow M, Kelly B, Taddy M. The National Bureau of Economic Research. Cambridge, MA; 2017 Mar. Text as Data URL: https://www.nber.org/papers/w23276
20. Kowalczyk A. SVM Tutorial. 2014 Nov 23. SVM Tutorial: How to classify text in R URL: https://www.svm-tutorial.com/2014/11/svm-classify-text-r/ [accessed 2018-08-01]
21. Demographics Pro. URL: https://www.demographicspro.com/ [accessed 2019-01-25]
22. Cavazos-Rehg PA, Krauss MJ, Sowles S, Connolly S, Rosas C, Bharadwaj M, et al. A content analysis of depression-related Tweets. Comput Human Behav 2016 Jan 01;54:351-357 [FREE Full text] [doi: 10.1016/j.chb.2015.08.023] [Medline: 26392678]

23. Rennison CM. U.S. Department of Justice. Washington, DC: U.S. Department of Justice; 2002 Aug 01. Rape and Sexual Assault: Reporting to Police and Medical Attention, 1992-2000 URL: https://www.bjs.gov/content/pub/pdf/rsarp00.pdf [accessed 2019-07-29]
24. Sini R. BBC. 2017 Dec 27. #MeToo, #TakeAKnee and #Covfefe: Hashtags that dominated in 2017 URL: https://www.bbc.com/news/world-42251490 [accessed 2019-02-25]
25. LaMotte S. CNN Health. 2017 Oct 19. For some, #MeToo sexual assault stories trigger trauma not empowerment URL: https://www.cnn.com/2017/10/19/health/me-too-sexual-assault-stories-trigger-trauma/index.html [accessed 2019-02-27] [WebCite Cache ID 76UD9grl9]

Abbreviations

API: application programming interface
LASSO: least absolute shrinkage and selection operator
NPV: negative predictive value
OLS: ordinary least squares model
PPV: positive predictive value
SVM: support vector machine

Edited by G Eysenbach; submitted 26.02.19; peer-reviewed by A Daughton, A Benis; comments to author 28.03.19; revised version received 28.06.19; accepted 07.07.19; published 03.09.19.

Please cite as:
Modrek S, Chakalov B
The #MeToo Movement in the United States: Text Analysis of Early Twitter Conversations
J Med Internet Res 2019;21(9):e13837
URL: https://www.jmir.org/2019/9/e13837/
doi:10.2196/13837
PMID:31482849

Original Paper

Disparities in Electronic Patient Portal Use in Prenatal Care: Retrospective Cohort Study

Erinma P Ukoha[1], MD, MPH; Joe Feinglass[2], PhD; Lynn M Yee[3], MD, MPH

[1]Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California San Francisco, San Francisco, CA, United States

[2]Division of General Internal Medicine and Geriatrics, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

[3]Division of Maternal Fetal Medicine, Department of Obstetrics and Gynecology, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

Corresponding Author:
Erinma P Ukoha, MD, MPH
Department of Obstetrics, Gynecology, and Reproductive Sciences
University of California San Francisco
550 16th Street
San Francisco, CA, 94143
United States
Phone: 1 415 476 5192
Email: erinma.ukoha@ucsf.edu

Abstract

Background: Electronic patient portals are websites that provide individuals access to their personal health records and allow them to engage through a secure Web-based platform. These portals are becoming increasingly popular in contemporary health care systems. Patient portal use has been found to be beneficial in multiple specialties, especially in the management of chronic disease. However, disparities have been identified in portal use in which racial and ethnic minorities and individuals with lower socioeconomic status have been shown to be less likely to enroll and use patient portals than non-Hispanic white persons and individuals with higher socioeconomic status. Electronic patient portal use by childbearing women has not been well studied, and data on portal use during pregnancy are limited.

Objective: This study aimed to quantify the use of an electronic patient portal during pregnancy and examine whether disparities related to patients' demographics or clinical characteristics exist.

Methods: This was a retrospective cohort study of women who received prenatal care at an academic medical center from 2014 to 2016. Clinical records were reviewed for portal use and patient data. Patients were considered enrolled in the portal if they had an account at the time of delivery, and enrollees were compared with nonenrollees. Enrollees were further categorized based on the number of secure messages sent during pregnancy as active (≥1) or inactive (0) users. Bivariable chi-square and multivariable Poisson regression models were used to calculate the incidence rate ratio of portal enrollment and, if enrolled, of active use based on patients' characteristics.

Results: Of the 3450 women eligible for inclusion, 2530 (73.33%) enrolled in the portal. Of these enrollees, 72.09% (1824/2530) were active users. There was no difference in portal enrollment by maternal race and ethnicity on multivariable models. Women with public insurance (adjusted incidence rate ratio; aIRR 0.60, 95% CI 0.49-0.84), late enrollment in prenatal care (aIRR 0.78, 95% CI 0.69-0.89 for second trimester and aIRR 0.50, 95% CI 0.39-0.64 for third trimester), and high-risk pregnancies (aIRR 0.82, 95% CI 0.75-0.89) were significantly less likely to enroll. Conversely, nulliparity (aIRR 1.10, 95% CI 1.02-1.20) and having more than 8 prescription medications at prenatal care initiation (aIRR 1.19, 95% CI 1.06-1.32) were associated with greater likelihood of enrollment. Among portal enrollees, the only factor significantly associated with active portal use (ie, secure messaging) was nulliparity (aIRR 1.11, 95% CI 1.01-1.23).

Conclusions: Among an obstetric population, multiple clinical and socioeconomic factors were associated with electronic portal enrollment, but not subsequent active use. As portals become more integrated as tools to promote health, efforts should be made to ensure that already vulnerable populations are not further disadvantaged with regard to electronic-based care.

(J Med Internet Res 2019;21(9):e14445) doi:10.2196/14445

KEYWORDS

patient portals; socioeconomic factors; pregnancy; cohort studies

Introduction

Background

The use of technology to communicate between patients and providers has become routine in medicine. One such technological development is the use of Web-based patient portal services in the electronic health record, which has been adopted in large part because of the requirements of the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act. HITECH created the Electronic Health Record Incentive Programs and required meaningful use of programs such as patient portals [1]. Financial support for patient portals by the government has encouraged its widespread adoption [2].

A patient portal is a secure Web-based platform directly linked to an electronic medical record (EMR) that provides personal health information, encouraging participants to become more active in their health care [3]. Portals allow individuals to communicate with health care providers, access portions of their medical record, refill prescriptions, and schedule appointments [4]. Of the numerous studies that have explored the effect of patient portals on clinical care, many have shown portal use to be associated with positive patient-reported outcomes, including improved patient satisfaction and patient-provider communication [1,5-8]. In particular, secure messaging is a unique aspect within patient portals that may facilitate patient self-management, shared decision making, and patient satisfaction by allowing patients more opportunities to communicate with providers [9]. Secure messaging has been associated with favorable clinical outcomes, specifically in studies examining diabetes management, in which patients who engage in secure messaging have lower hemoglobin A_{1c} values [1,8,10-15].

Objectives

The use of patient portals has not been widely investigated in obstetrics [16,17]. Moreover, with the rapid expansion of patient portals, disparities in health outcomes and care that already exist may be perpetuated by disparities in technology access and electronic health literacy [4,18-23]. Given that studies outside of obstetrics suggest portals may be associated with favorable clinical outcomes, there is an impetus to better understand the patterns of and factors associated with portal use during pregnancy. Thus, this study aimed to estimate the use of an electronic patient portal during pregnancy and determine whether disparities in portal use exist.

Methods

Study Cohort

This study was a cohort study of all women who received antenatal care at a single large-volume academic medical center from January 1, 2014, to January 1, 2016. Women eligible for inclusion must have been older than 18 years and received prenatal care, as defined by at least three clinical encounters, with Northwestern Medical Group providers. The study was approved by the institutional review board of Northwestern University before initiation, and all data were obtained from the EMR used for clinical care.

Northwestern Memorial Hospital and Northwestern Medical Group use EpicCare EMR and MyChart, the associated commercial patient portal. Through Northwestern's MyChart, patients are able to view records, review laboratory and imaging results, message providers, schedule appointments, and request medication refills. A unique access code used for portal enrollment can be generated by the provider during a clinical encounter or patients can self-enroll online through their email without an access code. Patients then have to activate their MyChart account through the portal website before they are able to freely access and use portal functions. The portal can then be used via either Web-based interfaces or mobile apps. In this practice, as part of the standard clinic workflow, all patients are invited to initiate MyChart accounts at their prenatal visit through various avenues, for example, as part of their routine prenatal intake packet, in the after-visit summary from provider encounters, or through interactions with the front desk staff. Some women may already have had accounts from previous interactions within the health care system. However, beyond enrollment, there are no additional health system campaigns or initiatives designed to encourage particular aspects of portal use.

Available portal use data included whether participants had a portal account, when it was initiated, and whether they used the portal for secure messaging. Using these data, we classified participants on the basis of antenatal portal enrollment and portal use. Patients were considered to be enrolled in the patient portal if they had an account at the time of delivery, regardless of when the account was created. Patient portal enrollees were further subcategorized into active or inactive users based on their secure message use. Active users were those who had sent 1 or more secure messages during pregnancy, whereas inactive users sent 0 messages antenatally. Only patient portal use for communication with obstetric providers (physicians, nurse practitioners, certified nurse midwives, or nurses within the Department of Obstetrics and Gynecology) qualified for this analysis (ie, communication with nonprenatal care providers was not included).

Statistical Analysis

The primary analyses focused on portal enrollment. We compared sociodemographic characteristics of portal enrollees (cases) with those of nonenrollees (controls), including age, race, ethnicity, insurance type, primary language, and neighborhood income (for which lower income neighborhood was defined as a zip code area in which the median annual household income was less than US $40,000 based on the 2015 five-year American Community Survey census data). The following clinical characteristics were also measured: parity, obesity, high-risk pregnancy (defined as a woman with any of the following characteristics: pregestational diabetes, gestational diabetes, chronic hypertension, gestational hypertension or preeclampsia, multifetal gestation, or care with a maternal-fetal medicine physician), gestational age at initiation of care, and number of prescription medications documented in the medical record at initial prenatal encounter. Comparisons were also made based on whether, once enrolled, a woman demonstrated active versus inactive use.

We used bivariable chi-square analysis and calculated unadjusted incidence rate ratios of portal enrollment and, once enrolled, of active use based on aforementioned patients' characteristics. To calculate adjusted incidence rate ratios (aIRR), we used multivariable Poisson regression models, adjusting for potentially confounding variables with *P*<.10 in bivariable analyses and also taking into account the time in relation to when electronic portals first launched to control for potential time bias. Poisson regression models were used for the analysis as they can be a more accurate estimate for treatment effect than adjusted odds ratios when the incidence of the outcome of interest is more common [24,25]. Analyses were performed using IBM SPSS Statistics for Macintosh, Version 24.0 and Stata Statistical Software, Release 13. All analyses were 2-tailed, and *P*<.05 was used to define statistical significance.

Results

A total of 3450 women were eligible for inclusion. Of these women, 1553 (45.01%) were non-Hispanic white (NHW), 372 (10.78%) were non-Hispanic black (NHB), 394 (11.42%) were Hispanic, 270 (7.83%) were Asian, and 861 (24.96%) were either listed as other or unknown. Overall, 3029 (87.80%) had private insurance. By zip code of residence, 437 (12.67%) of women lived in a low-income residential area, and 98 (2.84%) had a primary language other than English. In addition, 1819 (52.72%) participants were nulliparous, and 2530 (73.33%) eligible women enrolled in the patient portal (Table 1).

In bivariable analysis, compared with NHW patients, NHB and Hispanic women were significantly less likely to enroll in the patient portal. Similarly, women who were younger, had public insurance, had a low household income, or did not speak English as a primary language all had lower frequencies of portal enrollment. On multivariable models accounting for potential confounders (Table 2), women with public insurance remained significantly less likely to enroll in the patient portal compared with women with private insurance (aIRR 0.60, 95% CI 0.49-0.84). Other demographic factors did not remain associated with portal enrollment after adjusting for potential confounders.

Regarding clinical characteristics, women with high-risk pregnancies were less likely to enroll in the patient portal—56.7% (522/920) of nonenrolled versus 34.11% (863/2530) of enrolled; *P*<.001, aIRR 0.82, 95% CI 0.75-0.89. Gestational age at initiation of care was also associated with patient portal enrollment, with those who began care in the second trimester and third trimester significantly less likely to enroll. In contrast, nulliparity (aIRR 1.10, 95% CI 1.02-1.20) and having more than 8 prescription medications at the first prenatal visit (aIRR 1.19, 95% CI 1.06-1.32) were both associated with a significantly increased likelihood of enrolling (Table 2).

Of the 2530 women who enrolled in the patient portal, 1824 (72.09%) were categorized as active users (Table 3). In contrast to the differences in socioeconomic and clinical factors seen in portal enrollment, these characteristics were not associated with active portal use among enrollees (Table 4). After adjusting for potential confounders, only nulliparity remained associated with an increased likelihood of engaging in secure messaging (aIRR 1.12, 95% CI 1.02-1.23).

Table 1. Sociodemographic and clinical characteristics by patient portal enrollment (N=3450).

Patients' characteristics	Patient portal, n (%)		*P* value
	Nonenrollee (n=920)	Enrollee (n=2530)	
Age (years)			<.001
18-29	242 (26.3)	407 (16.09)	
30-39	608 (66.1)	1927 (76.17)	
≥40	70 (7.6)	196 (7.75)	
Race/ethnicity			<.001
Non-Hispanic white	310 (33.7)	1243 (49.13)	
Non-Hispanic black	169 (18.4)	203 (8.02)	
Hispanic	147 (16.0)	247 (9.76)	
Asian	56 (6.1)	214 (8.46)	
Other/unknown	238 (25.9)	623 (24.62)	
Insurance[a]			<.001
Private	703 (76.4)	2326 (91.94)	
Medicaid/Medicare	188 (20.4)	105 (4.15)	
Self-pay/uninsured	29 (3.2)	99 (3.91)	
Low household income[b]	190 (20.7)	247 (9.76)	<.001
Non-English as primary language	48 (5.2)	50 (1.98)	<.001
High-risk pregnancy[c]	522 (56.7)	863 (34.11)	<.001
Nulliparous	392 (42.6)	1427 (56.40)	<.001
Obese (body mass index ≥30 kg/m^2)	296 (32.2)	473 (18.70)	<.001
Gestational age at initiation of care			<.001
0 to 13 0/7 weeks gestational age	555 (60.3)	2188 (86.48)	
13 1/7 to 26 6/7 weeks gestational age	231 (25.1)	273 (10.79)	
≥27 0/7 weeks gestational age	134 (14.6)	69 (2.72)	
Number of prescription medications at initial prenatal encounter			<.001
0-3 medications	539 (58.6)	1263 (49.92)	
4-7 medications	274 (29.8)	803 (31.74)	
≥8 medications	107 (11.6)	464 (18.34)	

[a] Insurance was determined at the patient's first prenatal visit.

[b] Low household income was defined as patients whose zip code corresponded to a residential area in which the median household income was less than US $40,000 based on the 2015 five-year American Community Survey census data.

[c] High-risk pregnancy was defined as a patient with any of the following characteristics: type I diabetes, type II diabetes, gestational diabetes, chronic hypertension, gestational hypertension, preeclampsia, multifetal gestation, or receiving care from the Maternal-Fetal Medicine division.

Table 2. Unadjusted and adjusted incidence rate ratio of electronic patient portal enrollment (N=3450).

Patients' characteristics	Incidence rate ratio (95% CI)	Adjusted incidence rate ratio (95% CI)[a]
Age (years)		
18-29	Reference	Reference
30-39	1.21 (1.09-1.35)	1.07 (0.96-1.20)
≥40	1.17 (0.99-1.39)	1.12 (0.94-1.33)
Race/ethnicity		
Non-Hispanic white	Reference	Reference
Non-Hispanic black	0.68 (0.59-0.79)	0.88 (0.75-1.03)
Hispanic	0.78 (9.68-0.90)	0.96 (0.83-1.10)
Asian	0.99 (0.86-1.14)	1.05 (0.91-1.22)
Other/unknown	0.90 (0.82-1.00)	0.99 (0.90-1.10)
Insurance[b]		
Private	Reference	Reference
Medicaid/Medicare	0.47 (0.38-0.57)	0.60 (0.49-0.84)
Self-pay/uninsured	1.01 (0.82-1.23)	1.0 (0.81-1.22)
Low household income[c]	0.75 (0.65-0.85)	0.91 (0.79-1.05)
Non-English as primary language	0.69 (0.52-0.91)	0.80 (0.60-1.06)
High-risk pregnancy[d]	0.77 (0.71-0.84)	0.82 (0.75-0.89)
Nulliparous	1.16 (1.07-1.25)	1.10 (1.02-1.20)
Obesity (body mass index ≥30 kg/m^2)	0.80 (0.73-0.89)	0.99 (0.89-1.10)
Gestational age at initiation of care		
0 to 13 0/7 weeks gestational age	Reference	Reference
13 1/7 to 26 6/7 weeks gestational age	0.68 (0.60-0.77)	0.78 (0.69-0.89)
≥27 0/7 weeks gestational age	0.43 (0.34-0.54)	0.50 (0.39-0.64)
Number of prescription medications at initial prenatal encounter		
0-3 medications	Reference	Reference
4-7 medications	1.05 (0.96-1.15)	1.06 (0.97-1.16)
≥8 medications	1.16 (1.04-1.29)	1.19 (1.06-1.32)

[a] Multivariable models were adjusted for variables with *P*<.10 in bivariable analyses: age, race/ethnicity, insurance, household income, primary language, pregnancy risk, parity, obesity, gestational age at initiation of care, number of prescription medications at initial prenatal encounter, and time.

[b] Insurance was determined at the patient's first prenatal visit.

[c] Low household income was defined as patients whose zip code corresponded to a residential area in which the median household income was less than US $40,000 based on the 2015 five-year American Community Survey census data.

[d] High-risk pregnancy was defined as a patient with any of the following characteristics: type I diabetes, type II diabetes, gestational diabetes, chronic hypertension, gestational hypertension, preeclampsia, multifetal gestation, or receiving care from the Maternal-Fetal Medicine division.

Table 3. Sociodemographic and clinical characteristics by secure messaging use among patient portal enrollees (N=2530).

Patients' characteristics	Inactive user: sent 0 messages (n=706), n (%)	Active user: sent ≥1 message (n=1824), n (%)	*P* value
Age (years)			.007
18-29	121 (17.1)	286 (15.68)	
30-39	513 (72.7)	1414 (77.52)	
≥40	72 (10.2)	124 (6.80)	
Race/ethnicity			.04
Non-Hispanic white	331 (46.9)	912 (50.00)	
Non-Hispanic black	70 (9.9)	133 (7.29)	
Hispanic	66 (9.3)	181 (9.92)	
Asian	49 (6.9)	165 (9.05)	
Other/unknown	190 (26.9)	433 (23.74)	
Insurance[a]			.02
Private	653 (92.5)	1673 (91.72)	
Medicaid/Medicare	36 (5.1)	69 (3.78)	
Self-pay/uninsured	17 (2.4)	82 (4.49)	
Low household income[b]	81 (11.5)	166 (9.10)	.71
Non-English as primary language	15 (2.1)	35 (1.91)	.74
High-risk pregnancy[c]	279 (39.5)	584 (32.01)	<.001
Nulliparous	344 (48.7)	1083 (59.37)	<.001
Obesity (body mass index ≥30 kg/m^2)	146 (20.7)	327 (17.92)	.11
Gestational age at initiation of care			<.001
0 to 13 0/7 weeks gestational age	577 (81.7)	1611 (88.32)	
13 1/7 to 26 6/7 weeks gestational age	96 (13.6)	177 (9.70)	
≥27 0/7 weeks gestational age	33 (4.7)	36 (1.97)	
Number of prescription medications at initial prenatal encounter			.28
0-3 medications	340 (48.2)	923 (50.60)	
4-7 medications	223 (31.6)	580 (31.79)	
≥8 medications	143 (20.2)	321 (17.59)	

[a] Insurance was determined at the patient's first prenatal visit.

[b] Low household income was defined as patients whose zip code corresponded to a residential area in which the median household income was less than US $40,000 based on the 2015 five-year American Community Survey census data.

[c] High-risk pregnancy was defined as a patient with any of the following characteristics: type I diabetes, type II diabetes, gestational diabetes, chronic hypertension, gestational hypertension, preeclampsia, multifetal gestation, or receiving care from the Maternal-Fetal Medicine division.

Table 4. Unadjusted and adjusted incidence rate ratio of secure messaging use among electronic patient portal enrollees (N=2530).

Patients’ characteristics	Incidence rate ratio (95% CI)	Adjusted incidence rate ratio (95% CI)[a]
Age (years)		
18-29	Reference	Reference
30-39	1.04 (0.92-1.19)	1.06 (0.93-1.21)
≥40	0.89 (0.72-1.10)	0.94 (0.76-1.17)
Race/ethnicity		
Non-Hispanic white	Reference	Reference
Non-Hispanic black	0.90 (0.75-1.07)	0.95 (0.78-1.14)
Hispanic	0.99 (0.85-1.16)	1.02 (0.87-1.21)
Asian	1.04 (0.88-1.23)	1.04 (0.88-1.23)
Other/unknown	0.94 (0.85-1.16)	0.97 (0.87-1.10)
Insurance[b]		
Private	Reference	Reference
Medicaid/Medicare	0.90 (0.71-1.15)	0.99 (0.77-1.28)
Self-pay/uninsured	1.14 (0.91-1.42)	1.09 (0.87-1.37)
Low household income[c]	0.94 (0.80-1.10)	0.99 (0.84-1.17)
Non-English as primary language	0.99 (0.71-1.38)	1.01 (0.72-1.41)
High-risk pregnancy[d]	0.92 (0.83-1.01)	0.97 (0.87-1.07)
Nulliparous	1.12 (1.02-1.23)	1.12 (1.02-1.23)
Obesity (body mass index ≥30 kg/m^2)	0.96 (0.86-1.09)	1.02 (0.90-1.15)
Gestational age at initiation of care		
0 to 13 0/7 weeks gestational age	Reference	Reference
13 1/7 to 26 6/7 weeks gestational age	0.89 (0.76-1.03)	0.93 (0.79-1.09)
≥27 0/7 weeks gestational age	0.76 (0.56-1.05)	0.80 (0.58-1.10)
Number of prescription medications at initial prenatal encounter		
0-3 medications	Reference	Reference
4-7 medications	0.99 (0.89-1.10)	1.01 (0.91-1.12)
≥8 medications	0.94 (0.83-1.07)	0.97 (0.85-1.11)

[a] Multivariable models were adjusted for variables with *P*<.10 in bivariable analyses: age, race/ethnicity, insurance, household income, primary language, pregnancy risk, parity, obesity, gestational age at initiation of care, number of prescription medications at initial prenatal encounter, and time.

[b] Insurance was determined at the patient’s first prenatal visit.

[c] Low household income was defined as patients whose zip code corresponded to a residential area in which the median household income was less than US $40,000 based on the 2015 five-year American Community Survey census data.

[d] High-risk pregnancy was defined as a patient with any of the following characteristics: type I diabetes, type II diabetes, gestational diabetes, chronic hypertension, gestational hypertension, preeclampsia, multifetal gestation, or receiving care from the Maternal-Fetal Medicine division.

Discussion

Principal Findings

In this study of a large and diverse obstetric population, we found significant socioeconomic and clinical disparities in patient portal enrollment among pregnant women. Specifically, women who were publicly insured, medically higher risk, and late to initiate prenatal care were less likely to enroll in the electronic patient portal, whereas women who were nulliparous and taking more medications were more likely to enroll in the electronic patient portal. However, once enrolled, we found few differences in patient characteristics based on active use within the portal.

Comparison of Results With Previous Studies

The disparities identified in this obstetric population are consistent with previous studies examining portal use in a nonpregnant population [4,18-23]. These findings are supported by a case-control study in Boston that demonstrated individuals who registered for the electronic portal were more likely to be NHW, less likely to have Medicare or Medicaid insurance, and were younger and healthier compared with nonenrollees [19]. Furthermore, studies have shown that nonwhite and Hispanic persons, those with the lowest incomes, and publicly insured and uninsured persons were less likely to activate and subsequently use their electronic patient portal account after registration [18,21,22]. Conversely, other studies have also suggested that race and ethnicity are independent risk markers for portal use. For example, a cohort study of more than 1700 individuals at Kaiser Permanente Georgia demonstrated that compared with NHW individuals, NHB individuals were less likely to register for the patient portal. These differences by race were not accounted for by differences in education, income, or internet access, although greater education and internet access were independently associated with portal registration [4]. In our study, once potential confounders were controlled for, we identified no statistically significant difference in portal use by race. This, however, may be because of incomplete race and ethnicity data. Such disparities may, in fact, exist in a better identified population cohort.

In addition, we noted differences in portal enrollment based on clinical characteristics. Notably, women with complicated pregnancies were significantly less likely to be enrolled in the patient portal. Previous studies have come to conflicting conclusions regarding patient clinical status and portal use [18-20]. It has been suggested that the *worried well* may use the portal more, which was consistent with our findings that women classified as having high-risk pregnancies complicated by diabetes, hypertensive disorders, multifetal gestations, or under the care of the Maternal-Fetal Medicine division were less likely to be enrolled in the patient portal. Ideally, given the potential benefits of portal use, enrollment should be targeted at more complicated patients who may benefit from enhanced provider-patient communication and additional support [6]. In contrast to other findings, patients with more prescription medications at the initiation of prenatal care were in fact more likely to enroll in the patient portal. If medications are properly reconciled in accordance with electronic health record use standards, it may be that these women with more medical problems, and thus more prescription medications, are more engaged in patient portals. However, it is possible that the number of medications may actually be proxy for integration into the medical system as opposed to health status and medical morbidity. In addition, women who initiated prenatal care at a later gestational age were significantly less likely to have patient portal accounts. Factors that are associated with late presentation to prenatal care—including nonwhite ethnic group, immigrant status, lower education level, lower socioeconomic status, nontraditional perception of the value of prenatal care, uninsured, unemployed, and poor reproductive health knowledge—also may play a role in whether patients are offered the electronic patient portal, their attitude toward the patient portal, or enrollment patterns [26].

Despite disparities in portal enrollment, we identified very few differences in secure messaging use. After adjusting for potential confounders, only nulliparity remained associated with an increased likelihood of engaging in secure messaging. This finding can be accounted for in that women with their first pregnancy may have more questions in general, independent from health risk. Thus, once the initial barrier to portal enrollment is overcome, patients are overall equally likely to engage and message their providers, a finding that has also been noted in studies of nonpregnant populations [23]. This finding regarding portal use underscores the importance of identifying and overcoming potential barriers that dissuade individuals from enrolling in the patient portal.

Clinical Implications

Electronic patient portal interfaces require several steps before achieving access, and potential barriers exist at each level from being offered registration to active use [23]. A 2016 national survey study demonstrated that compared with their white and non-Hispanic counterparts, respectively, black and Hispanic individuals were significantly less likely to be offered electronic portal access by their health care provider. Moreover, individuals who are older, have poor health, and are poorly educated were also significantly less likely to be offered access by their provider and to engage in the patient portal [27]. Yet, an overwhelming majority of individuals considered online access to their personal health information important with no difference noted by race or economic status [27]. These findings highlight the role health care providers and their inherent bias may play in adoption of electronic patient portals, a critical step required to realize any benefit associated with patient portals.

Patients' characteristics that may affect enrollment and use of an electronic portal include, but are not limited to, electronic/computer literacy, health literacy and numeracy, perceived benefit of portal use, patient preferences regarding provider communication, and trust in the health care system and electronic mediums [18,23]. In particular, health and computer literacy have both been shown to be associated with portal enrollment and long-term portal use [28-31]. High perceived health literacy along with internet access at home, high self-rated ability when using the internet, and high overall online ability have all been associated with increased likelihood of portal use—all factors that may vary among different patient populations [28,29]. Patients' attitudes regarding the patient portal and electronic communications may also have a varying impact on its use depending on the specific population. For example, 2 qualitative studies with focus group interviews have shown that black and Hispanic patients may have negative attitudes toward the portal, were dubious regarding electronic communications and their potential benefit, and were fearful that the portal would diminish existing relationships with providers [30,31]. These studies emphasize the importance of identifying potential barriers that may dissuade individuals from engaging in the patient portal.

Strengths and Limitations

Strengths of this study are that it included a large population of pregnant women and applied few exclusion criteria. However, given the retrospective nature of this study, any associations cannot be assumed to be causal, and there is potential for unmeasured confounding. For example, we did not have direct measures for health literacy, internet access, education level, income level, or self-care behaviors, all of which may play a role in portal engagement. In addition, at our institution, patients can enroll in the portal on their own accord without directly receiving an access code from their provider, and the exact steps leading to any given enrollment along with the temporality of enrollment are unknown. Therefore, because there is no standardized process for enrolling patients for portal access, the degree to which provider bias and discriminatory offering patterns may be related to the disparities seen in enrollment is unclear. We have identified subsets of patients less likely to enroll; thus, studies evaluating the logistics of enrollment and systems issues related to it can shed light on how to potentially increase portal use. Furthermore, although this sample was large and diverse, the patients were nonetheless receiving care at a large academic tertiary care setting, and therefore, findings may not be fully generalizable to other contexts.

Research Implications

Future work must identify reasons for portal nonuse, develop indicators of successful outcomes of portal use, and implement potential systems-based or provider-based interventions to increase portal enrollment [32]. Efforts to expand portal enrollment in populations of greatest need may be key to improving health communication in these populations. In addition, given that secure messaging has previously been associated with positive patient outcomes in primary care settings, more studies focusing on this particular aspect are needed [1,8,10-15]. Qualitative content analysis of patient-provider electronic communication can provide information on how secure messaging within portal use may be related to perinatal, maternal, and neonatal outcomes.

Conclusions

In summary, we identified socioeconomic and clinical disparities within portal enrollment and use during prenatal care. Disparities that were significant in regard to patient enrollment did not exist when examining subsequent secure messaging use. Thus, once initial barriers to portal enrollment were overcome, patient portal use was similar among most groups. As electronic patient portals become more integrated as tools to promote health, it is important to understand the patterns of use and the potential impact in pregnancy, especially as it relates to perinatal outcomes in already disadvantaged groups [32].

Acknowledgments

Research reported in this publication was supported, in part, by the *National Institutes of Health's National Center for Advancing Translational Sciences*, Grant Number UL1TR001422, via the *Northwestern University Clinical and Translational Sciences Institute Enterprise Data Warehouse Pilot Program.* In addition, LMY is supported by the *National Institute of Child Health and Human Development* K12 HD050121-11. 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. Kuo A, Dang S. Secure messaging in electronic health records and its impact on diabetes clinical outcomes: a systematic review. Telemed J E Health 2016 Sep 1;22(9):769-777. [doi: 10.1089/tmj.2015.0207] [Medline: 27027337]
2. Wright A, Feblowitz J, Samal L, McCoy AB, Sittig DF. The medicare electronic health record incentive program: provider performance on core and menu measures. Health Serv Res 2014 Feb;49(1 Pt 2):325-346 [FREE Full text] [doi: 10.1111/1475-6773.12134] [Medline: 24359554]
3. Irizarry TA, DeVito DA, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
4. Roblin DW, Houston TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. J Am Med Inform Assoc 2009;16(5):683-689 [FREE Full text] [doi: 10.1197/jamia.M3169] [Medline: 19567790]
5. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015 Feb 10;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
6. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res 2012 Nov 26;14(6):e162 [FREE Full text] [doi: 10.2196/jmir.2238] [Medline: 23183044]
7. Goldzweig CL, Orshansky G, Paige NM, Towfigh AA, Haggstrom DA, Miake-Lye I, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. Ann Intern Med 2013 Nov 19;159(10):677-687. [doi: 10.7326/0003-4819-159-10-201311190-00006] [Medline: 24247673]
8. Osborn CY, Mayberry LS, Mulvaney SA, Hess R. Patient web portals to improve diabetes outcomes: a systematic review. Curr Diab Rep 2010 Dec;10(6):422-435 [FREE Full text] [doi: 10.1007/s11892-010-0151-1] [Medline: 20890688]
9. Schickedanz A, Huang D, Lopez A, Cheung E, Lyles CR, Bodenheimer T, et al. Access, interest, and attitudes toward electronic communication for health care among patients in the medical safety net. J Gen Intern Med 2013 Jul;28(7):914-920 [FREE Full text] [doi: 10.1007/s11606-012-2329-5] [Medline: 23423453]

10. Wade-Vuturo AE, Mayberry LS, Osborn CY. Secure messaging and diabetes management: experiences and perspectives of patient portal users. J Am Med Inform Assoc 2013 May 1;20(3):519-525 [FREE Full text] [doi: 10.1136/amiajnl-2012-001253] [Medline: 23242764]
11. Harris LT, Haneuse SJ, Martin DP, Ralston JD. Diabetes quality of care and outpatient utilization associated with electronic patient-provider messaging: a cross-sectional analysis. Diabetes Care 2009 Jul;32(7):1182-1187 [FREE Full text] [doi: 10.2337/dc08-1771] [Medline: 19366959]
12. Harris LT, Koepsell TD, Haneuse SJ, Martin DP, Ralston JD. Glycemic control associated with secure patient-provider messaging within a shared electronic medical record: a longitudinal analysis. Diabetes Care 2013 Sep;36(9):2726-2733 [FREE Full text] [doi: 10.2337/dc12-2003] [Medline: 23628618]
13. Goldzweig CL, Towfigh AA, Paige NM, Orshansky G, Haggstrom DA, Beroes JM, et al. Systematic Review: Secure Messaging Between Providers and Patients, and Patients' Access to Their Own Medical Record. Washington, DC: Department of Veterans Affairs; 2012.
14. Shimada SL, Allison JJ, Rrosen AK, Feng H, Houston TK. Sustained use of patient portal features and improvements in diabetes physiological measures. J Med Internet Res 2016 Jul 1;18(7):e179. [doi: 10.2196/jmir.5663] [Medline: 27369696]
15. Petullo B, Noble B, Dungan KM. Effect of electronic messaging on glucose control and hospital admissions among patients with diabetes. Diabetes Technol Ther 2016 Sep;18(9):555-560. [Medline: 27398824]
16. Shaw E, Howard M, Chan D, Waters H, Kaczorowski J, Price D, et al. Access to web-based personalized antenatal health records for pregnant women: a randomized controlled trial. J Obstet Gynaecol Can 2008 Jan;30(1):38-43. [doi: 10.1016/S1701-2163(16)32711-6] [Medline: 18198066]
17. Wieland D, Gibeau A, Dewey C, Roshto M, Frankel H. Patient portal readiness among postpartum patients in a safety net setting. Appl Clin Inform 2017 Jul 5;8(3):698-709 [FREE Full text] [doi: 10.4338/ACI-2016-12-RA-0204] [Medline: 28678891]
18. Ancker JS, Barrón Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, et al. Use of an electronic patient portal among disadvantaged populations. J Gen Intern Med 2011 Oct;26(10):1117-1123 [FREE Full text] [doi: 10.1007/s11606-011-1749-y] [Medline: 21647748]
19. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95 [FREE Full text] [doi: 10.1197/jamia.M1833] [Medline: 16221943]
20. Carrell D, Ralston JD. Variation in adoption rates of a patient web portal with a shared medical record by age, gender, and morbidity level. AMIA Annu Symp Proc 2006;2006:871 [FREE Full text] [Medline: 17238491]
21. Miller H, Vandenbosch B, Ivanov D, Black P. Determinants of personal health record use: a large population study at Cleveland clinic. J Healthc Inf Manag 2007;21(3):44-48. [Medline: 19195293]
22. Wallace LS, Angier H, Huguet N, Gaudino JA, Krist A, Dearing M, et al. Patterns of electronic portal use among vulnerable patients in a nationwide practice-based research network: from the OCHIN practice-based research network (PBRN). J Am Board Fam Med 2016;29(5):592-603 [FREE Full text] [doi: 10.3122/jabfm.2016.05.160046] [Medline: 27613792]
23. Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW. Disparities in enrollment and use of an electronic patient portal. J Gen Intern Med 2011 Oct;26(10):1112-1116 [FREE Full text] [doi: 10.1007/s11606-011-1728-3] [Medline: 21538166]
24. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. J Am Med Assoc 1998 Nov 18;280(19):1690-1691. [doi: 10.1001/jama.280.19.1690] [Medline: 9832001]
25. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004 Apr 1;159(7):702-706. [doi: 10.1093/aje/kwh090] [Medline: 15033648]
26. Smith A, Bassett-Novoa E. Late Presentation to Prenatal Care. Am Fam Physician 2015 Sep 1;92(5):391-397 [FREE Full text] [Medline: 26371574]
27. Peacock S, Reddy A, Leveille SG, Walker J, Payne TH, Oster NV, et al. Patient portals and personal health information online: perception, access, and use by US adults. J Am Med Inform Assoc 2017 Apr 1;24(e1):e173-e177. [doi: 10.1093/jamia/ocw095] [Medline: 27413120]
28. Noblin AM, Wan TT, Fottler M. The impact of health literacy on a patient's decision to adopt a personal health record. Perspect Health Inf Manag 2012;9:1-13 [FREE Full text] [Medline: 23209454]
29. Woods SS, Forsberg CW, Schwartz EC, Nazi KM, Hibbard JH, Houston TK, et al. The association of patient factors, digital access, and online behavior on sustained patient portal use: a prospective cohort of enrolled users. J Med Internet Res 2017 Oct 17;19(10):e345. [doi: 10.2196/jmir.7895] [Medline: 29042345]
30. Goel MS, Brown TL, Williams A, Cooper AJ, Hasnain-Wynia R, Baker DW. Patient reported barriers to enrolling in a patient portal. J Am Med Inform Assoc 2011 Dec;18(Suppl 1):i8-12 [FREE Full text] [doi: 10.1136/amiajnl-2011-000473] [Medline: 22071530]
31. Lyles CR, Allen JY, Poole D, Tieu L, Kanter MH, Garrido T. 'I want to keep the personal relationship with my doctor': understanding barriers to portal use among African Americans and Latinos. J Med Internet Res 2016 Oct 3;18(10):e263. [doi: 10.2196/jmir.5910] [Medline: 27697748]
32. Ukoha EP, Yee LM. Use of electronic patient portals in pregnancy: an overview. J Midwifery Womens Health 2018 May;63(3):335-339. [doi: 10.1111/jmwh.12761] [Medline: 29758124]

Abbreviations

aIRR: adjusted incidence rate ratio
EMR: electronic medical record
HITECH: Health Information Technology for Economic and Clinical Health
NHB: non-Hispanic black
NHW: non-Hispanic white

Edited by G Eysenbach; submitted 18.04.19; peer-reviewed by B Chaudhry, M Reynolds, B Crotty; comments to author 10.07.19; revised version received 23.07.19; accepted 23.07.19; published 23.09.19.

Please cite as:
Ukoha EP, Feinglass J, Yee LM
Disparities in Electronic Patient Portal Use in Prenatal Care: Retrospective Cohort Study
J Med Internet Res 2019;21(9):e14445
URL: https://www.jmir.org/2019/9/e14445
doi:10.2196/14445
PMID:31586367

Original Paper

Interaction Patterns of Men Who Have Sex With Men on a Geosocial Networking Mobile App in Seven United States Metropolitan Areas: Observational Study

Nadia N Abuelezam[1], ScD; Yakir A Reshef[2], PhD; David Novak[3], MSW; Yonatan Hagai Grad[4,5], MD, PhD; George R Seage III[6], MPH, ScD; Kenneth Mayer[7], MD; Marc Lipsitch[6,8], DPhil

[1]William F Connell School of Nursing, Boston College, Chestnut Hill, MA, United States
[2]Department of Computer Science, Harvard University, Cambridge, MA, United States
[3]DSN Consulting, LLC, Boston, MA, United States
[4]Department of Immunology and Infectious Disease, Harvard TH Chan School of Public Health, Boston, MA, United States
[5]Division of Infectious Disease, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
[6]Department of Epidemiology, Harvard TH School of Public Health, Boston, MA, United States
[7]The Fenway Institute, Boston, MA, United States
[8]Center for Communicable Disease Dynamics, Boston, MA, United States

Corresponding Author:
Nadia N Abuelezam, ScD
William F Connell School of Nursing
Boston College
140 Commonwealth Ave
Chestnut Hill, MA, 02467
United States
Phone: 1 617 552 2312
Email: nadia.abuelezam@bc.edu

Abstract

Background: The structure of the sexual networks and partnership characteristics of young black men who have sex with men (MSM) may be contributing to their high risk of contracting HIV in the United States. Assortative mixing, which refers to the tendency of individuals to have partners from one’s own group, has been proposed as a potential explanation for disparities.

Objective: The objective of this study was to identify the age- and race-related search patterns of users of a diverse geosocial networking mobile app in seven metropolitan areas in the United States to understand the disparities in sexually transmitted infection and HIV risk in MSM communities.

Methods: Data were collected on user behavior between November 2015 and May 2016. Data pertaining to behavior on the app were collected for men who had searched for partners with at least one search parameter narrowed from defaults or used the app to send at least one private chat message and used the app at least once during the study period. Newman assortativity coefficient (R) was calculated from the study data to understand assortativity patterns of men by race. Pearson correlation coefficient was used to assess assortativity patterns by age. Heat maps were used to visualize the relationship between searcher’s and candidate’s characteristics by age band, race, or age band and race.

Results: From November 2015 through May 2016, there were 2,989,737 searches in all seven metropolitan areas among 122,417 searchers. Assortativity by age was important for looking at the profiles of candidates with correlation coefficients ranging from 0.284 (Birmingham) to 0.523 (San Francisco). Men tended to look at the profiles of candidates that matched their race in a highly assortative manner with R ranging from 0.310 (Birmingham) to 0.566 (Los Angeles). For the initiation of chats, race appeared to be slightly assortative for some groups with R ranging from 0.023 (Birmingham) to 0.305 (Los Angeles). Asian searchers were most assortative in initiating chats with Asian candidates in Boston, Los Angeles, New York, and San Francisco. In Birmingham and Tampa, searchers from all races tended to initiate chats with black candidates.

Conclusions: Our results indicate that the age preferences of MSM are relatively consistent across cities, that is, younger MSM are more likely to be chatted with and have their profiles viewed compared with older MSM, but the patterns of racial mixing are more variable. Although some generalizations can be made regarding Web-based behaviors across all cities, city-specific usage patterns and trends should be analyzed to create targeted and localized interventions that may make the most difference in the lives of MSM in these areas.

(J Med Internet Res 2019;21(9):e13766) doi:10.2196/13766

KEYWORDS

men who have sex with men; sexual behavior; race factors; population dynamics

Introduction

Background

Sexually transmitted infection (STI) and HIV transmission risk remain high among men who have sex with men (MSM) in the United States. MSM account for 68.2% of all primary and secondary syphilis cases in the United States in 2017 [1]. They account for 38.5% of isolates testing positive for *Neisseria gonorrhoeae* in the United States in 2017, up from 3.9% in 1989 [1]. Specific subpopulations at increased risk of infection include young black MSM and Latino MSM (MSM of color). Compared with their proportions in the US population, MSM of color have significantly higher burden of STIs than white MSM [1]. Rates of reported chlamydia cases were highest among black men (907.3/100,000), Native Hawaiian or Pacific Islander (376.6/100,000), and Asian men (402.6/100,000) in the United States when compared with white men (137.1/100,000) [1]. Gonorrhea rates are highest among black (660.7/100,000) and Asian (238.0/100,000) men in the United States [1].

Young black MSM are at highest risk for HIV and STI transmission in the United States despite reporting similar frequency of HIV risk behaviors as other groups [2,3]. Some believe that the structure of their sexual networks and partnership characteristics may be contributing to this high risk of HIV [4]. Assortative mixing (also known as homophily), which refers to the tendency of individuals to have partners from their own group, where a group can be defined by age, race, height, weight, or other characteristic factors involved in choosing a sexual partner, has been proposed as a potential explanation for disparities [5]. When assortative mixing by race is high, it can amplify STI and HIV prevalence differentials and may explain disparities in incidence among racial groups of MSM in the United States [4]. In a Web-based study of an ethnically diverse group of MSM, black MSM were more than 11 times as likely to have black partners than other racial group. In other studies of MSM in a number of US cities, black men were observed to have more black partners than any other racial group [6-9] and, in some cases, the majority of black men exclusively reported sex with black partners [10]. Nonblack MSM reported lesser preference for black partners in 1 study of MSM in San Francisco [11].

Assortative and disassortative mixing by age have also been found to be important to HIV transmission in particular contexts and are disparate by race [12], with more black MSM reporting sexual partners in a different age group (±5 years) than white men [13]. Black MSM in a number of studies were found to have a large number of partners that were in an older age group [9,10,13]. Evidence shows that having older sexual partners was associated with higher sexual risk behaviors in black MSM [12]. Odds of unprotected sex increased dramatically for young black MSM with partners who were 3 years older (a predictor only present for black MSM in this study) [4]. As black MSM have the highest HIV prevalence in the United States, this may potentially drive HIV transmission to young black MSM [12]. Studies to understand the relationship between sexual network composition and structure and risk for STIs are being conducted in these communities, but there remains a great deal of stigma associated with homosexuality and HIV that makes recruitment for studies difficult [14,15].

Understanding men’s preferences using Web-based apps may lead to better hypotheses and understanding of behaviors and risks associated with STI transmission in the real world. Increasingly, MSM use various forms of technology and the internet to locate sexual partners outside of physical local venues [16]. With the advancement of mobile technologies, geosocial networking (GSN) mobile apps have arisen that utilize global positioning system (GPS) to allow app users to find and chat with partners in their immediate city or community. Users post pictures and descriptions about themselves and are able to search for partners and chat with potential connections [17]. GSN apps allow men to *favorite* potential partners and identify when they may be in close proximity throughout their day [18]. Men’s use of these networks reflects the fact that they have control over who they interact with and respond to, and there is a relative ease with which they find partners [19]. For MSM of color, who have very few venues for socializing with other MSM because of stigma and taboo within their communities, Web-based platforms and apps play an important role in finding partners and providing a social atmosphere [20]. About one-third of all MSM have reported ever using a GSN app and 85% of those using apps engage with them daily [21,22].

Previous studies to estimate assortativity by age and race have focused on collecting egocentric and self-reported data from MSM through systemically or conveniently selected samples. Although important, these estimates are potentially limited by social desirability and recall biases of respondents [23-26], especially with regard to racial preferences. In addition, research has found that recruiting and retaining MSM of color in research using Web-based platforms has been difficult [26]. Studies that are able to observe behavior, without interfering with it, are ideal as they avoid the potential for bias and social desirability among the MSM being observed. Observing behavior on GSN apps provides a unique and novel perspective to collecting more robust objective data on sexual preferences and behavior.

Objective

Our study was designed to objectively identify the age- and race-related search patterns of users of a diverse GSN mobile app in seven major metropolitan areas in the United States to better understand the disparities in STI and HIV risk in MSM communities.

Methods

Study Population

Data were collected on user behavior while using a diverse MSM GSN mobile app on an Apple or Android device between November 2015 and May 2016. Data entered by the user upon initiation of an account (including age, race, height, weight, and partner preferences) were collected where available. When signing up for a profile, individuals had the option of self-identifying their race as one of the following categories: Asian, Black, Caucasian, Latino, Middle Eastern, Mixed, Pacific Islander, or Other. We coded individuals into the following racial categories: Asian, black, white (indicated Caucasian), Latino, and other (indicated Middle Eastern, Mixed, Pacific Islander, or Other). Every time a user logged into the app and searched for a partner, information was collected on the parameters of the conducted search, including the following: GPS location of the searcher; the list of potential candidates resulting from the search; whether or not the user looked at the details of a candidate's profile, favorited a candidate in the search list, or initiated a chat with a candidate in the search list; and the provided details of the candidate (including age, race, height, and weight). The data collected were composed of de-identified user identities for the searcher and the candidates resulting from the search.

Behavior on the app was collected for men who had searched for partners with at least 1 search parameter narrowed from defaults or used the app to send at least 1 private chat message to another individual and used the app at least once during the study period (November 2015 through May 2016). Users were categorized into 7 major metropolitan areas if their GPS coordinates fell within the metropolitan census tracts for the cities of Birmingham (Alabama), Boston (Massachusetts), Los Angeles (California), New York City (New York), San Francisco (California), Tampa (Florida), and Washington DC.

We had 2 main behavioral observations of interest: acquiring profile details and initiating a chat. If an individual clicked on a candidate's name and looked at the profile details of a candidate in their search list, we considered this to be a situation in which details were acquired. If an individual initiated a chat with a candidate on their search list by sending a message to a candidate from the search list, we considered this to be a situation in which a chat was initiated (independent of chat length or duration).

Human Subjects

An Institutional Review Board (Harvard TH Chan School of Public Health) and an independent ethics committee (Western Institutional Review Board) approved the study protocol and amendments for this study.

Assortativity

For each city, a mixing matrix (e) was created by age and race, respectively: the i th to j th entry of e was the proportion of all searcher–candidate pairs for which the searcher fell in category i and the candidate fell in category j. The matrix e assessed the proportional cross-tabulation of age (or race) for searcher and candidate in each city, so that the sum of the entries in the matrix is 1. These mixing matrices were analyzed to better understand the patterns of activity on the GSN app by age and race.

A total of 2 quantities were calculated from the study data to understand assortativity patterns of men using the GSN app by age and race. Newman assortativity coefficient, R, a parameter used to assess the extent to which a population exhibits assortative, neutral, or disassortative sexual mixing patterns, was calculated from mixing matrices using the following equation:

$$R = \mathrm{Tr}\, e - \| e^2 \| / 1 - \| e^2 \|$$

where R is the assortativity coefficient and e is the mixing matrix whose elements are e_{ij}. Tr e is the trace of the mixing matrix (the sum of its diagonal elements), and $\| e^2 \|$ is the sum of the squared values of the elements in the mixing matrix.

An R value of 1 represents perfect assortativity in which people mix only with others having the same characteristics. An R value of 0 represents random mixing and an R value of −1 represents perfect disassortative mixing. We calculated Newman assortativity coefficient for racial categories and age categorized into 5-year increments up to the age of 40 years. As this measure compares mixing within the same group with mixing between groups, it is sensitive to the choice of band size for continuous measures such as age, with smaller age bands yielding lower values of assortativity. We, therefore, also calculated Pearson correlation coefficient to understand assortativity by age taken continuously to avoid this sensitivity and decrease it to outliers. Separate assortativity coefficients were calculated for each city of interest and each classification of the interaction (looking at candidate's profile details or initiating a chat). When making reference to a specific group, such as men of Asian race/ethnicity, we define assortativeness for that group as a higher proportion of Asian candidates among the candidates of Asian searchers (the diagonal entry on the heat map) than among the candidates of all searchers (the corresponding marginal entry on the heat map).

On the basis of previous study, assortativity coefficients between 0.15 and 0.25 are considered minimally assortative, between 0.26 and 0.34 are considered moderately assortative, and 0.35 or larger are considered assortative [27,28].

Heat maps were used to help visualize the relationship between characteristics of the searcher and the candidate: age band, race, or age band and race. All values for heat maps were normalized by numbers of searchers in a category, thus the total value of a column (searcher) adds up to 100%. Besides (and above) each heat map, a column representing the distribution of candidates (and searchers) in the age, race, and age and race categories is shown. Heat maps were created for each city of interest and each classification of the interaction (looking at candidate's profile details or initiating a chat).

Results

Descriptive Statistics

From November 2015 through May 2016 there were 2,989,737 searches in all seven major metropolitan areas among 122,417 searchers. The median number of searches per searcher in all cities was 3 to 4, but there were outliers in each city. In New York and Washington DC, for example, some searchers had over 15,000 searches in the study period. All searches resulted in 752,832 unique candidates. The age and race profiles of the searchers and candidates for each metropolitan area are outlined in Table 1. Notably, the majority of searchers in all metropolitan areas were aged between 21 and 29 years, with 5.59% (230/4114, Birmingham) to 9.38% (1014/10,929, San Francisco) of searchers in each city aged >40 years. The age distribution of searchers across metropolitan areas was relatively consistent.

The racial composition of the searchers varied by city. The majority of searchers self-identified as black or other, respectively, in Birmingham (77.72%, 3188/4102 and 13.99%, 574/4102) and Washington DC (64.96%. 17,699/27,245 and 20.67%, 5632/27,245) (Table 1). Boston had the largest proportion of white searchers (20.98%, 923/4400) and the second highest proportion of Asian searchers (22.52%, 991/4400) behind San Francisco (36.02%, 3902/10,833). Los Angeles had the highest proportion of searchers self-identifying as Latino (15.25%, 2003/13,132) and a high proportion self-identifying as other (22.96%, 3015/13,132) behind New York (24.98%, 10,490/41,991) and Tampa (23.44%, 1389/5925).

Table 1. Characteristics of men who have sex with men that searched, or were candidates themselves resulting from a search, for partners on a diverse social networking app focusing on such men in seven US metropolitan cities, using the app between November 2015 and May 2016.

Category	Birmingham	Boston	Los Angeles	New York	San Francisco	Tampa	Washington DC
Instances (n)	N=72,793	N=90,837	N=265,376	N=1,413,803	N=219,768	N=79,597	N=847,563
Get details	14,363	12,842	40,583	183,085	30,488	18,901	132,581
Chat	4978	3309	10,756	51,166	6759	6175	40,429
Race, n (%)							
Searchers	N=4102	N=4400	N=13,132	N=41,991	N=10,833	N=5925	N=27,245
White	211 (5.14)	923 (20.98)	1147 (8.73)	2978 (7.09)	1573 (14.52)	670 (11.31)	1867 (6.85)
Black	3188 (77.72)	1088 (24.73)	4534 (34.53)	18,113 (43.14)	2022 (18.67)	3109 (52.47)	17,699 (64.96)
Latino	35 (0.85)	473 (10.75)	2003 (15.25)	5751 (13.70)	919 (8.48)	533 (9.00)	934 (3.43)
Asian	94 (2.29)	991 (22.52)	2433 (18.53)	4659 (11.10)	3902 (36.02)	224 (3.78)	1113 (4.09)
Other	574 (13.99)	925 (21.02)	3015 (22.96)	10,490 (24.98)	2417 (22.31)	1389 (23.44)	5632 (20.67)
Candidates	N=15,633	N=23,843	N=50,682	N=125,551	N=48,222	N=18,049	N=75,867
White	555 (3.55)	3986 (16.72)	5974 (11.79)	13,239 (10.54)	7016 (14.55)	2069 (11.46)	6518 (8.59)
Black	11,522 (73.70)	4517 (18.94)	13,641 (26.91)	46,512 (37.05)	7612 (15.79)	9455 (52.39)	41,175 (54.27)
Latino	294 (1.88)	2048 (8.59)	6036 (11.91)	11,379 (9.06)	3467 (7.19)	1680 (9.31)	3671 (4.84)
Asian	226 (1.45)	9648 (40.46)	15,275 (30.14)	29,621 (23.59)	22,333 (46.31)	673 (3.73)	9168 (12.08)
Other	3036 (19.42)	3644 (15.28)	9756 (19.25)	24,800 (19.75)	7794 (16.16)	4172 (23.11)	15,335 (20.21)
Age (years), n (%)							
Searchers	N=4114	N=4408	N=13,183	N=42,305	N=10,929	N=5937	N=27,546
18-20	646 (15.70)	553 (12.55)	1119 (8.49)	4371 (10.33)	1235 (11.30)	815 (13.73)	2641 (9.59)
21-24	1152 (28.00)	1205 (27.34)	3344 (25.37)	10,375 (24.52)	2627 (24.04)	1586 (26.71)	6017 (21.84)
25-29	1215 (29.53)	1364 (30.94)	4686 (35.55)	14,700 (34.75)	3291 (30.11)	1906 (32.10)	9219 (33.47)
30-34	597 (14.51)	607 (13.77)	2077 (15.76)	6617 (15.64)	1812 (16.58)	760 (12.80)	4725 (17.15)
35-39	274 (6.66)	315 (7.15)	977 (7.41)	3359 (7.94)	950 (8.69)	402 (6.77)	2469 (8.96)
40 and older	230 (5.59)	364 (8.26)	980 (7.43)	2883 (6.81)	1014 (9.38)	468 (7.88)	2475 (8.98)
Candidates	N=15,814	N=24,114	N=51,124	N=126,942	N=48,770	N=18,151	N=76,908
18-20	2094 (13.24)	2861 (11.86)	4574 (8.95)	12197 (9.61)	4278 (8.77)	2264 (12.47)	7467 (9.71))
21-24	4704 (29.75)	6783 (28.13)	12,593 (24.63)	30,259 (23.84)	11,421 (23.42)	4940 (27.22)	17,987 (23.39)
25-29	5134 (32.46)	7959 (33.01)	17,749 (34.72)	43,800 (34.50)	16,202 (33.22)	6047 (33.31)	26,530 (34.50)
30-34	2173 (13.74)	3421 (14.19)	8223 (16.08)	20,725 (16.33)	8483 (17.39)	2463 (13.57)	12,644 (16.44)
35-39	915 (5.79)	1658 (6.88)	4247 (8.31)	10,548 (8.31)	4441 (9.11)	1223 (6.38)	6416 (8.34)
40 and older	794 (5.02)	1432 (5.38)	3738 (7.31)	9413 (7.42)	3945 (8.09)	1214 (6.69)	5864 (7.62)

Statistics for Age

Assortativity by age, as measured by Pearson correlation coefficient, ranged from 0.117 (Tampa) to 0.210 (Los Angeles) across all search activities.

Details

Assortativity by age was important for looking at the profile details of candidates, with correlation coefficients ranging from 0.284 (Birmingham) to 0.523 (San Francisco) (Figure 1). Across all cities, searchers looked at the profiles of candidates that were from their own age category (suggesting assortative preference) and men aged between 21 and 29 years in all areas (seen by the horizontal striping across the heat maps for these 2 age groups) (Figures 2 and 3). Mean absolute age differences between the searcher and the candidate ranged from 4.33 (Boston, SD 7.2) to 5.65 (Tampa, SD 8.1) when details were examined (results not shown).

Figure 1. Assortativity results for race (top) and age (bottom) for each of the seven metropolitan areas.

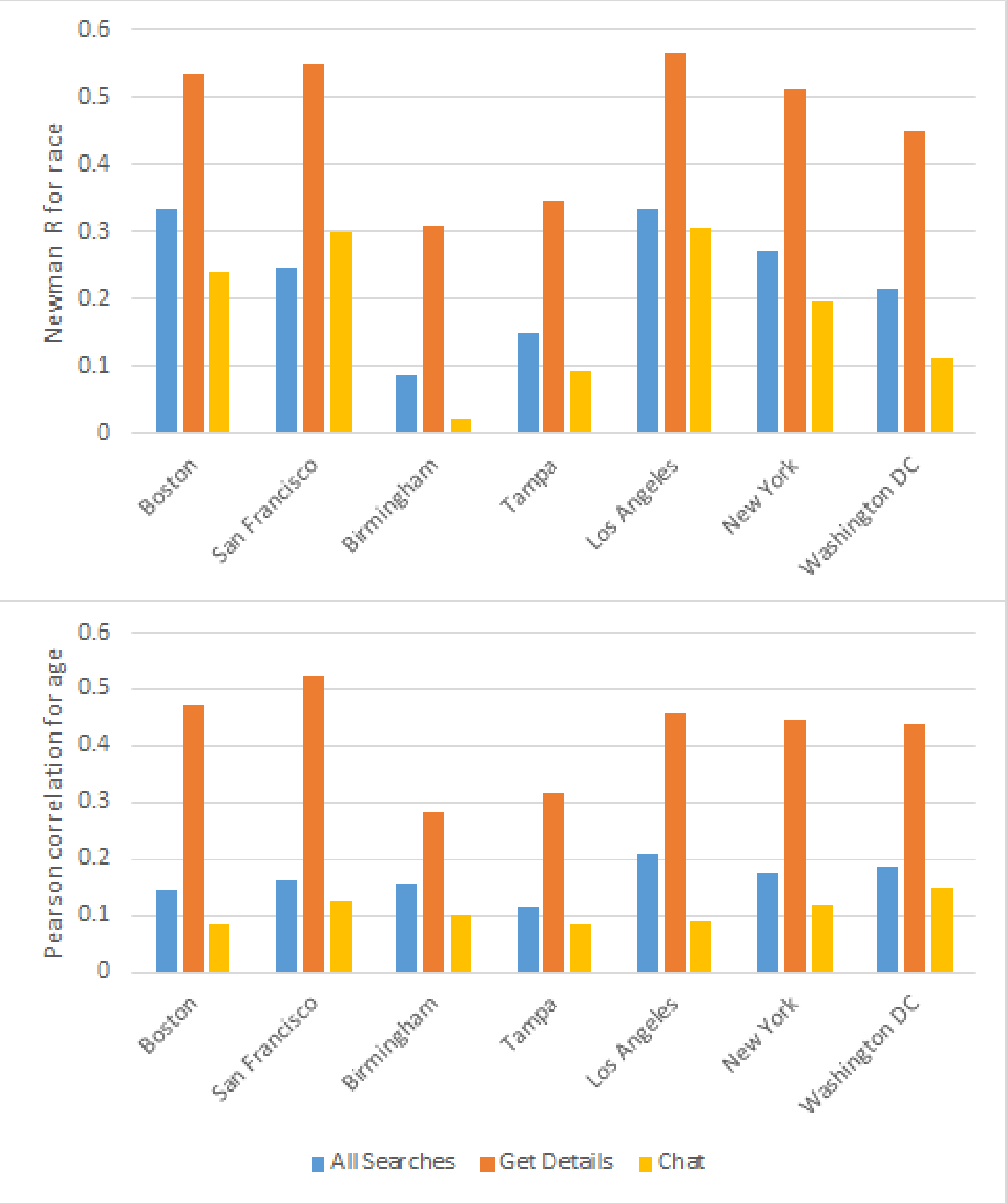


Figure 2. Assortativity by age (years) for details that are acquired by searchers on candidates in Birmingham, Los Angeles, Boston, and New York. Heat maps show the relationship between searcher's age band and the candidate's age band. All values are normalized by numbers of searchers in each age category (the total value of a column adds up to 100%). Besides each heat map, a column representing the distribution of candidates in the age band and above each heat map, a row representing the distribution of searchers is shown.

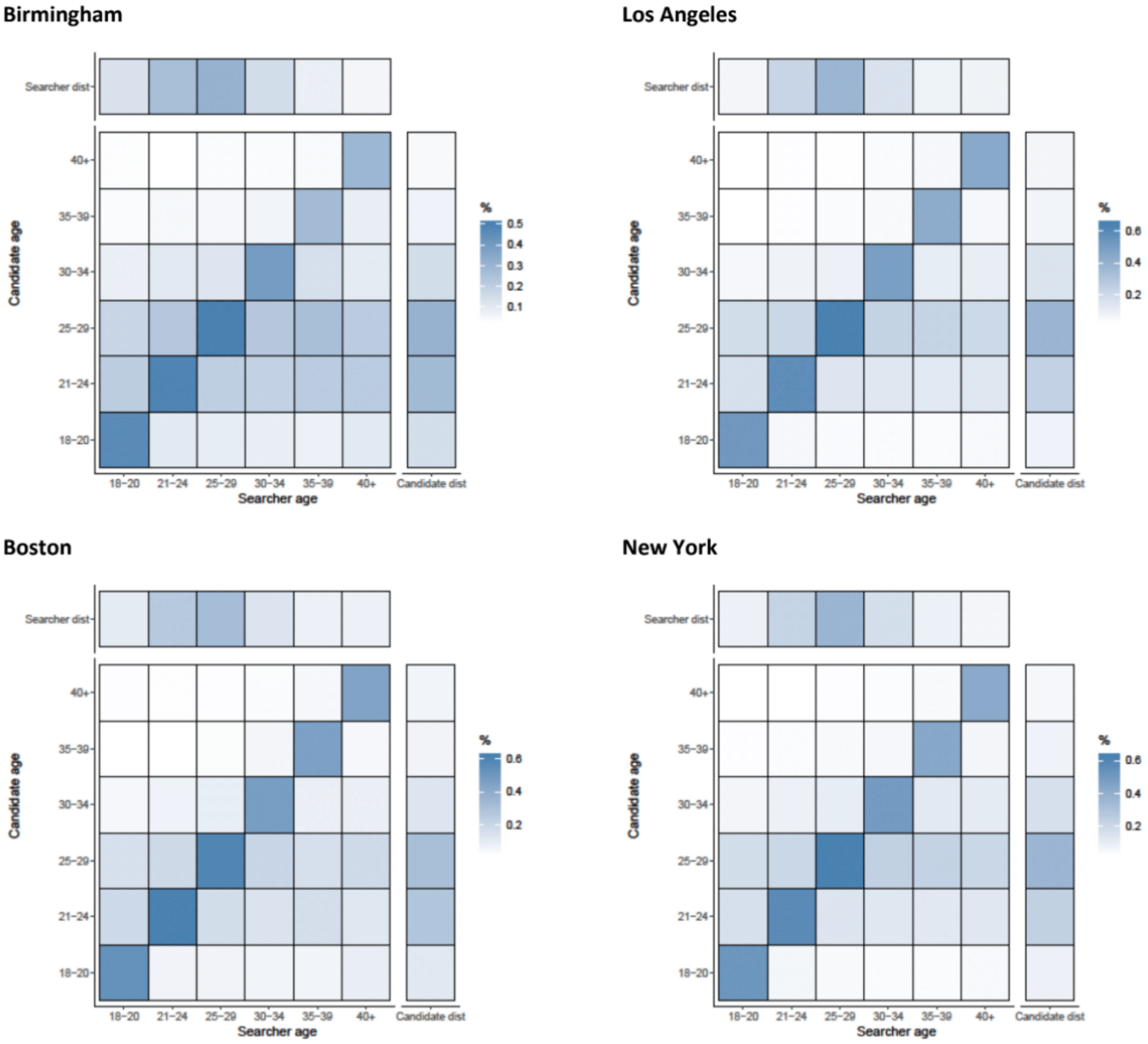
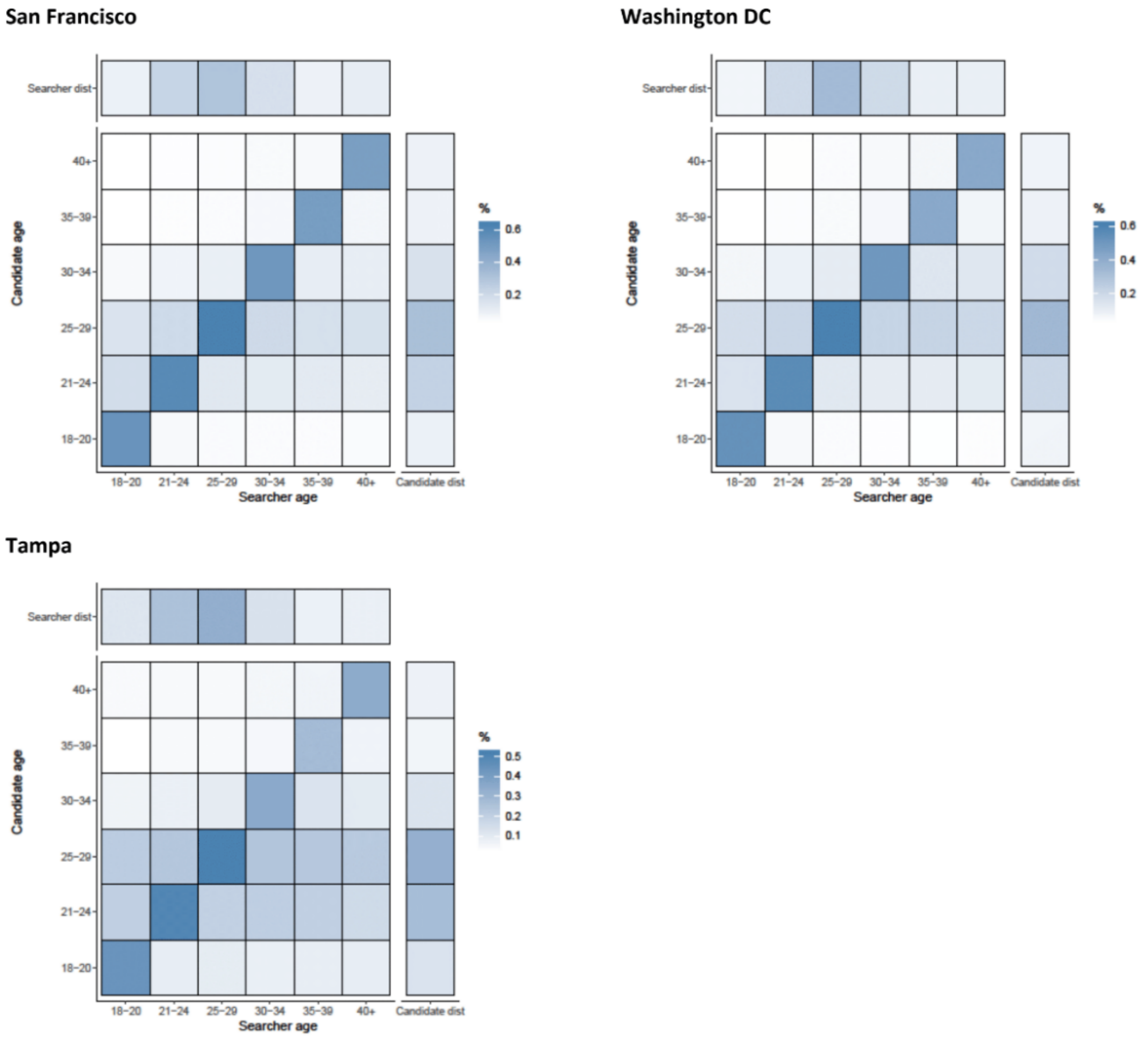


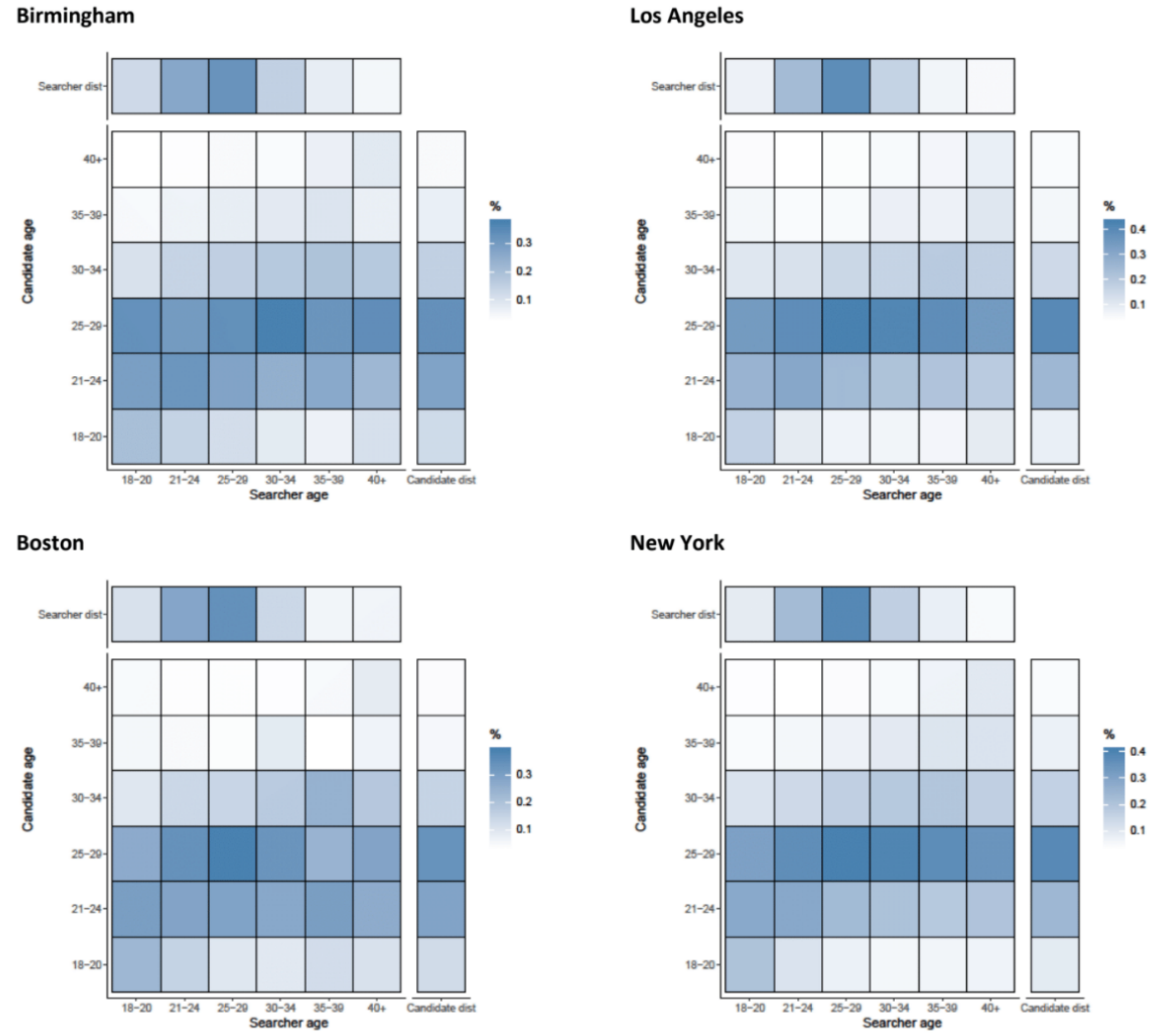
Figure 3. Assortativity by age (years) for details that are acquired by searchers on candidates in San Francisco, Washington DC, and Tampa. Heat maps show the relationship between searcher's age band and the candidate's age band. All values are normalized by numbers of searchers in each age category (the total value of a column adds up to 100%). Besides each heat map, a column representing the distribution of candidates in the age band and above each heat map, a row representing the distribution of searchers is shown.



Chats

When initiating chats with candidates, searchers were not highly selective by age, with correlation coefficients ranging from 0.085 (Boston) to 0.148 (Washington DC) (Figure 1). Younger men were more often chosen for chats in all cities, with 25 to 29-year-olds being more often chosen on chats than older men. Searchers from all age categories initiated chats with young men (aged 21-29 years), as can be seen by the horizontal stripes in the chat initiated heat maps (Figures 4 and 5). There was some age assortativity for men aged 25 to 29 years initiating chats with other men aged 25 to 29 years, as can be seen by the darker shaded box on the diagonal for these age groups in most cities, compared with the box for this age at the right margin, which gives the frequency with which this age group was chatted with by searchers of all ages. Chats were initiated with men from younger age categories most commonly in Boston (36.78%, 1217/3309) and least commonly in New York (34.72%, 17,763/51,166; results not shown). Chats were initiated with men from older age categories most commonly in San Francisco (34.39%, 2325/6759) and least commonly in Los Angeles (31.47%, 3385/10,756; results not shown). Mean absolute age differences between searcher and candidates ranged from 6.49 (New York, SD 7.9) and 7.09 (San Francisco, SD 8.1) when chats were initiated (results not shown).

Figure 4. Assortativity by age (years) for chats initiated by searchers with candidates in Birmingham, Los Angeles, Boston, and New York. Heat maps show the relationship between searcher's age band and the candidate's age band. All values are normalized by numbers of searchers in each age category (the total value of a column adds up to 100%). Besides each heat map, a column representing the distribution of candidates in the age band and above each heat map, a row representing the distribution of searchers is shown.



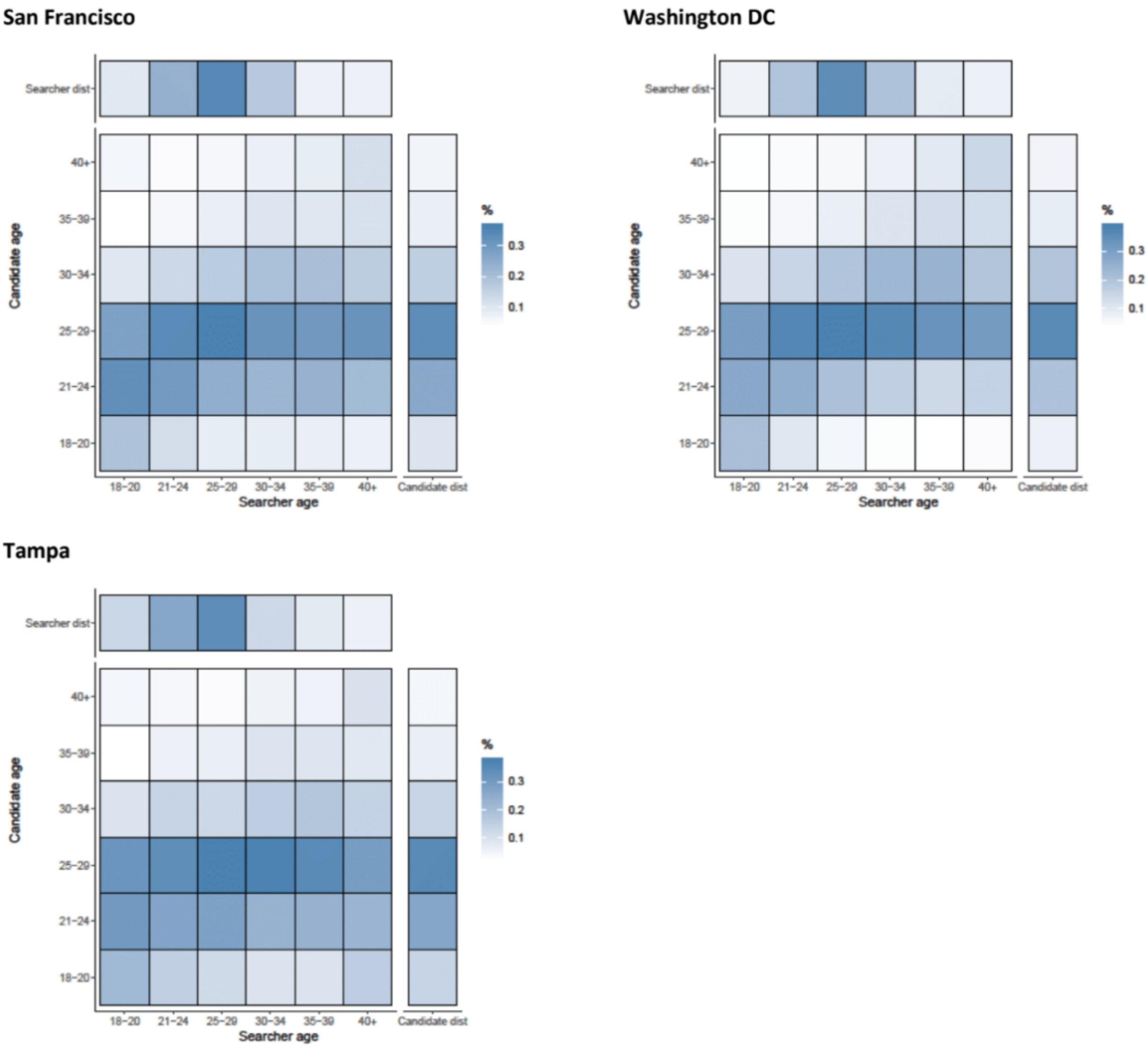
Birmingham

Los Angeles

Boston

New York

Figure 5. Assortativity by age (years) for chats initiated by searchers with candidates in San Francisco, Washington DC, and Tampa. Heat maps show the relationship between searcher's age band and the candidate's age band. All values are normalized by numbers of searchers in each age category (the total value of a column adds up to 100%). Besides each heat map, a column representing the distribution of candidates in the age band and above each heat map, a row representing the distribution of searchers is shown.



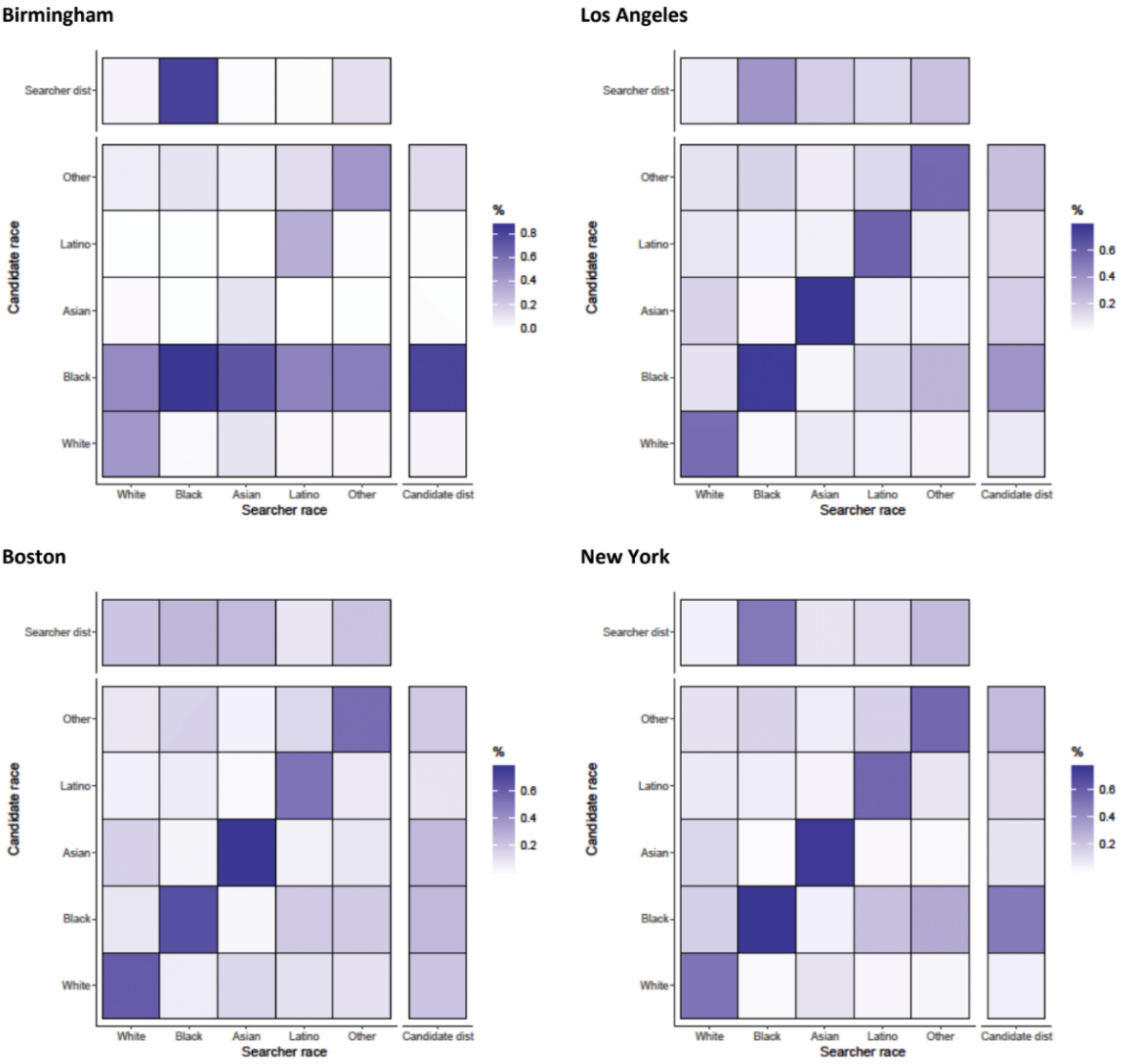
Newman Assortativity for Race

Across all searches in all of the metropolitan areas examined, there was evidence for moderate assortativity by race with Newman assortativity coefficient R>0 for all cities. Across all search activities, R was highest in Los Angeles (0.33) and Boston (0.335) and lowest in Birmingham (0.088) and Tampa (0.149) (Figure 1).

Details

Men tended to look at the details of candidates that matched their race in a highly assortative manner with Newman R coefficients ranging from 0.310 (Birmingham) to 0.566 (Los Angeles) (Figure 1). In Birmingham and Tampa, all searchers (regardless of race) viewed black candidates' profiles, potentially a byproduct of the proportion of candidates who were black in these cities (Figures 6 and 7). The majority of candidates in Birmingham, New York, Tampa, and Washington DC were black, whereas the majority of searchers in San Francisco were Asian (Figure 4). Black searchers looked at details for black candidates most commonly in Washington DC (83.20%, 110,307/132,581) and Birmingham (85.60%, 12,295/14,363) and least commonly in Boston (66.70%, 8566/12,842; results not shown).

Figure 6. Assortativity by race for details that are acquired by searchers on candidates in Birmingham, Los Angeles, Boston, and New York. Heat maps show the relationship between searcher's race group and the candidate's race group. All values are normalized by numbers of searchers in each race category (the total value of a column adds up to 100%). Besides each heat map, a column representing the distribution of candidates in the race group and above each heat map, a row representing the distribution of searchers is shown.



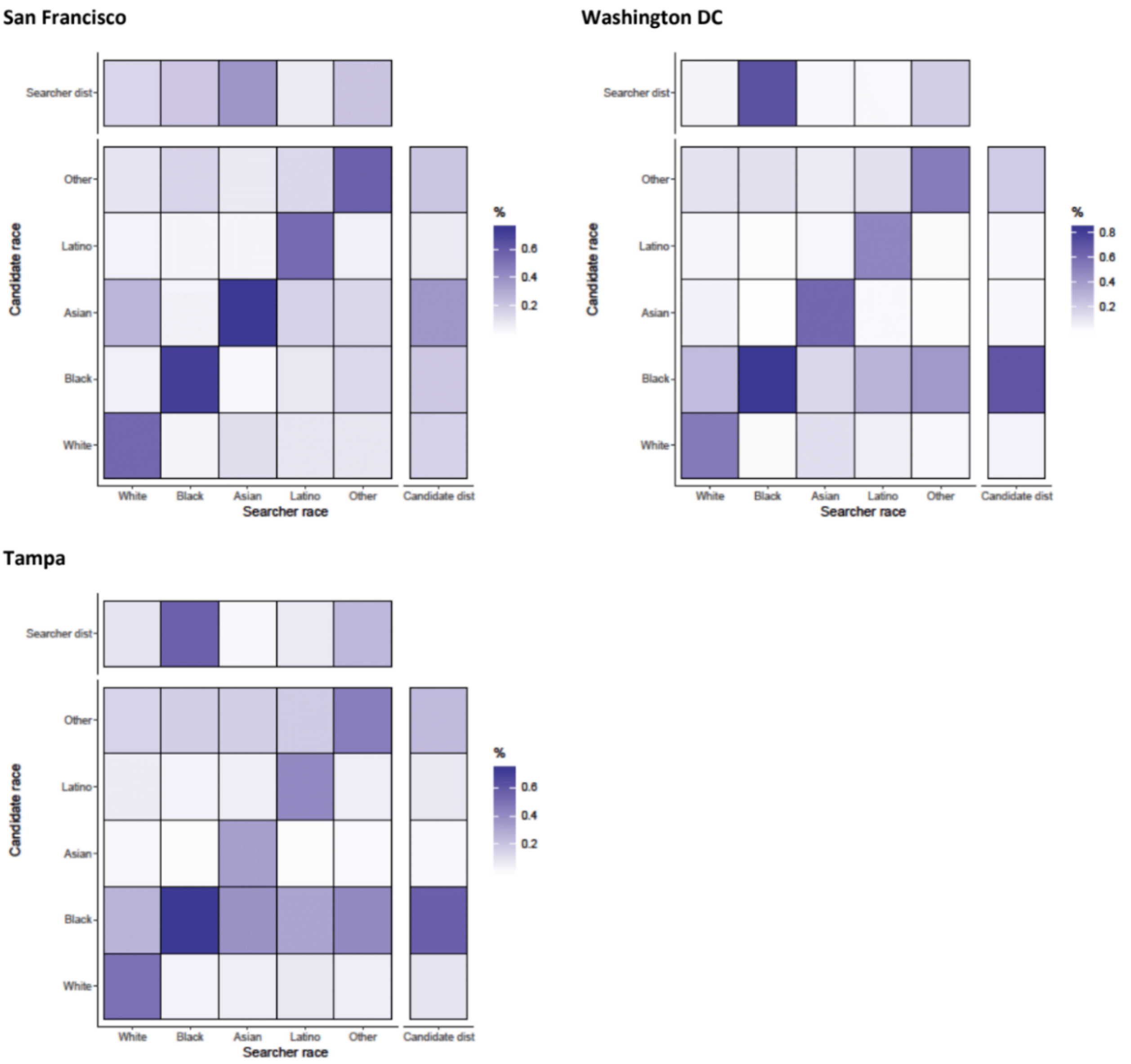
Birmingham

Los Angeles

Boston

New York

Figure 7. Assortativity by race for details that are acquired by searchers on candidates in San Francisco, Washington DC, and Tampa. Heat maps show the relationship between searcher's race group and the candidate's race group. All values are normalized by numbers of searchers in each race category (the total value of a column adds up to 100%). Besides each heat map, a column representing the distribution of candidates in the race group and above each heat map, a row representing the distribution of searchers is shown.



Chats

For the initiation of chats, race appeared to be assortative for some groups with R ranging from 0.023 (Birmingham) to 0.305 (Los Angeles) (Figure 1). Asian searchers were most assortative in initiating chats with Asian candidates in Boston, Los Angeles, New York, and San Francisco (Figures 8 and 9). In Birmingham and Tampa, searchers from all races tended to initiate chats with black candidates, as evidenced by the dark horizontal line in the heat maps for black candidate race. This reduced overall assortativity for chats in these cities. All cities showed strong assortativity among black searchers and candidates, with black searchers chatting with black candidates most commonly in Birmingham (82.50%, 4107/4978) and least commonly in Boston (54.82%, 1814/3309). Searchers with an "other" self-classified race tended to chat with black candidates in Los Angeles, New York, and Washington DC white searchers initiated chats with Asian candidates in Boston and San Francisco but not in other cities. There was slight evidence that Latino searchers were assortative in initiating chats with other Latinos in Boston, Los Angeles, and New York.

Figure 8. Assortativity by race for chats that are initiated by searchers on candidates in Birmingham, Los Angeles, Boston, and New York. Heat maps show the relationship between searcher's race group and the candidate's race group. All values are normalized by numbers of searchers in each race category (the total value of a column adds up to 100%). Besides each heat map, a column representing the distribution of candidates in the race group and above each heat map, a row representing the distribution of searchers is shown.

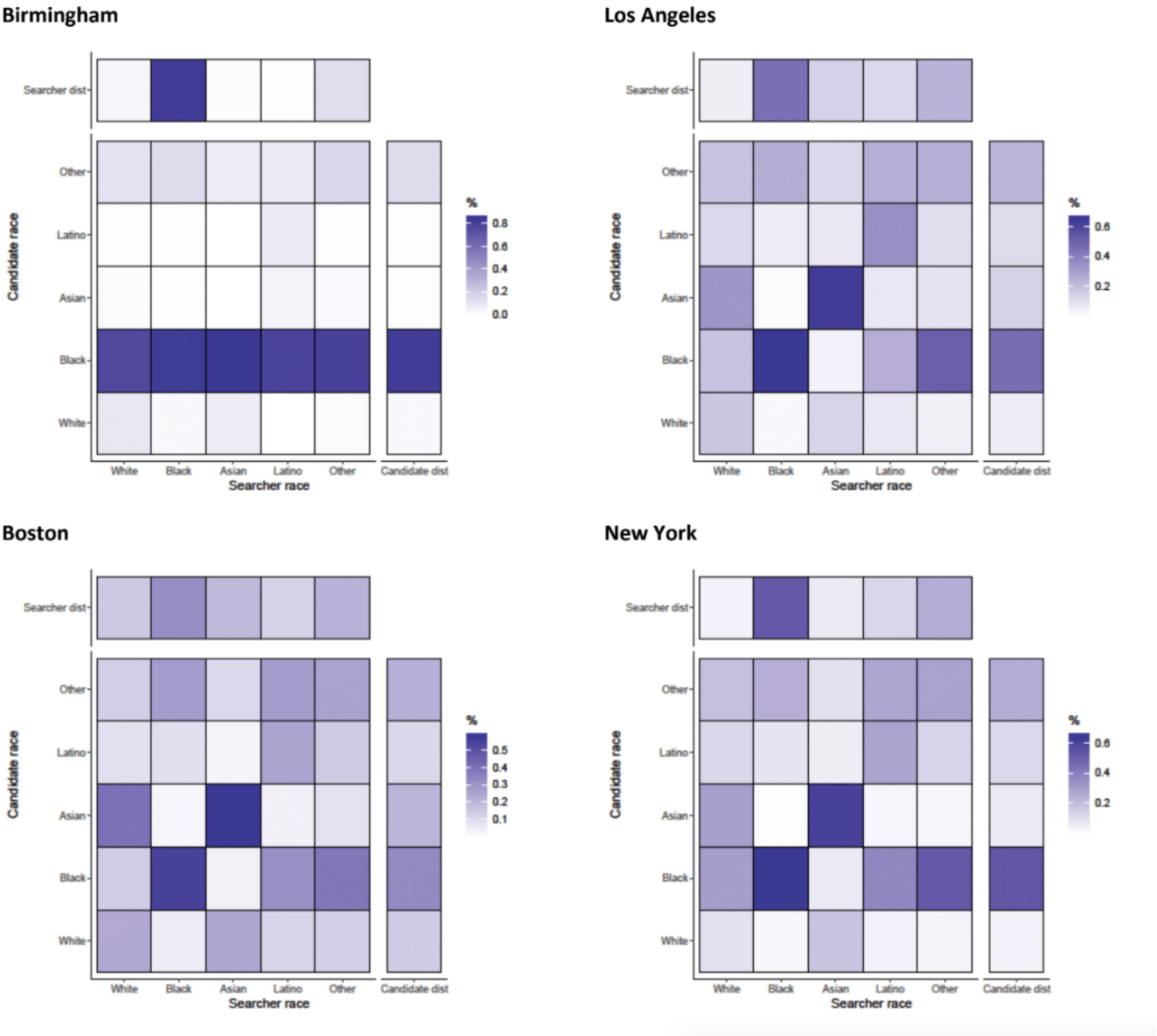
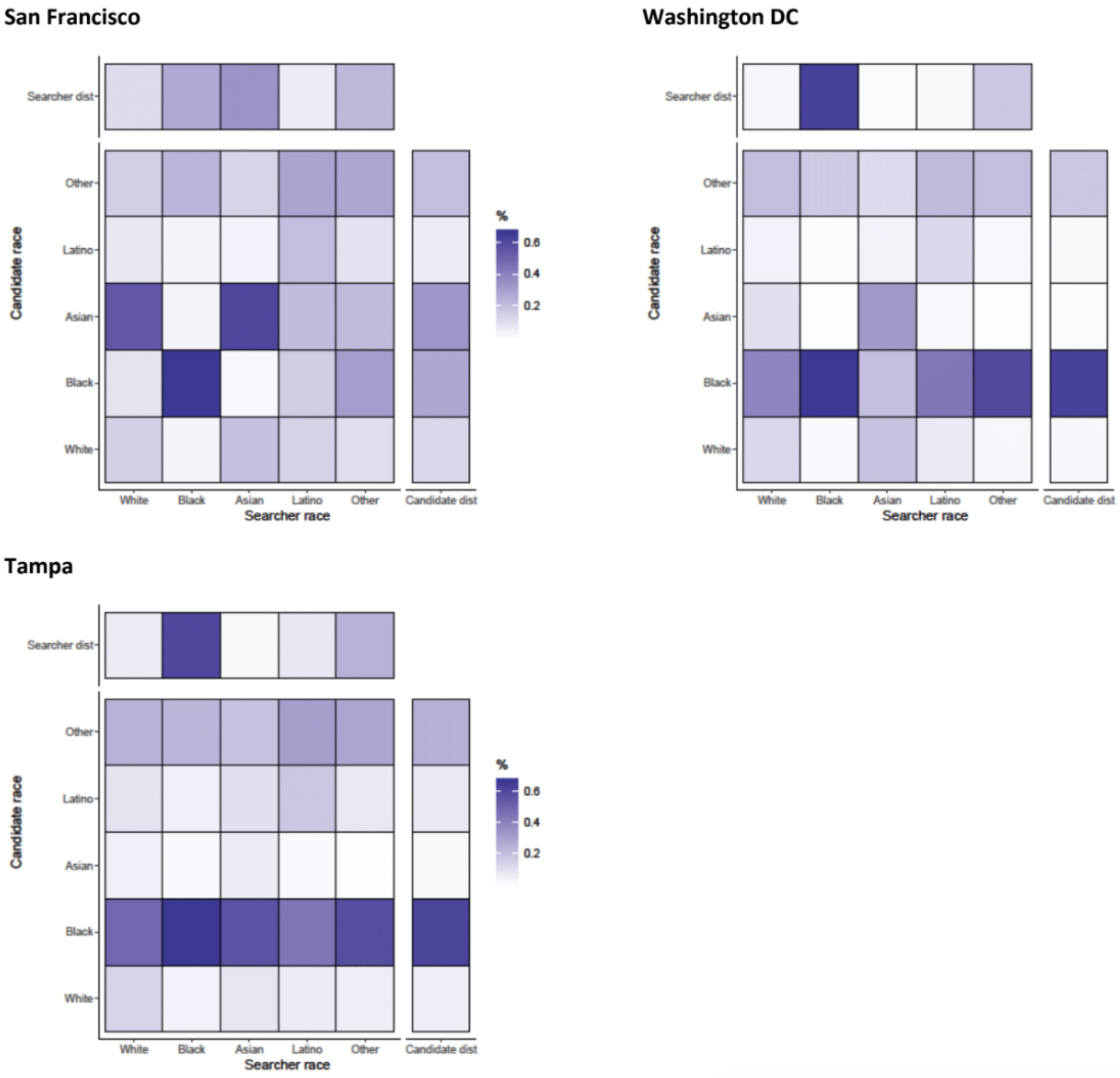


Figure 9. Assortativity by race for chats that are initiated by searchers on candidates in San Francisco, Washington DC, and Tampa. Heat maps show the relationship between searcher's race group and the candidate's race group. All values are normalized by numbers of searchers in each race category (the total value of a column adds up to 100%). Besides each heat map, a column representing the distribution of candidates in the race group and above each heat map, a row representing the distribution of searchers is shown.



Race and Age

When examining heat plots of both age and race together, we see that across searchers of various age/race combinations, 25 to 29-year-old black men are the most highly chatted with candidates, with similar trends for black men aged 20 to 24 and 30 to 34 years (Figures 10 and 11). In New York, black and Asian searchers are highly assortative by age when initiating chats but also more likely to chat with 25 to 29-year-old black and Asian candidates, respectively (Figure 11). Washington DC was the only city to show evidence of older black candidates (aged >35 years) being actively chatted with across all age and race groups of the searchers.

Figure 10. Assortativity by age (years) and race for chats initiated in Washington DC. Heat maps show the relationship between searcher's age and race group and the candidate's age and race group. All values are normalized by numbers of searchers in each age and race category (the total value of a column adds up to 100).

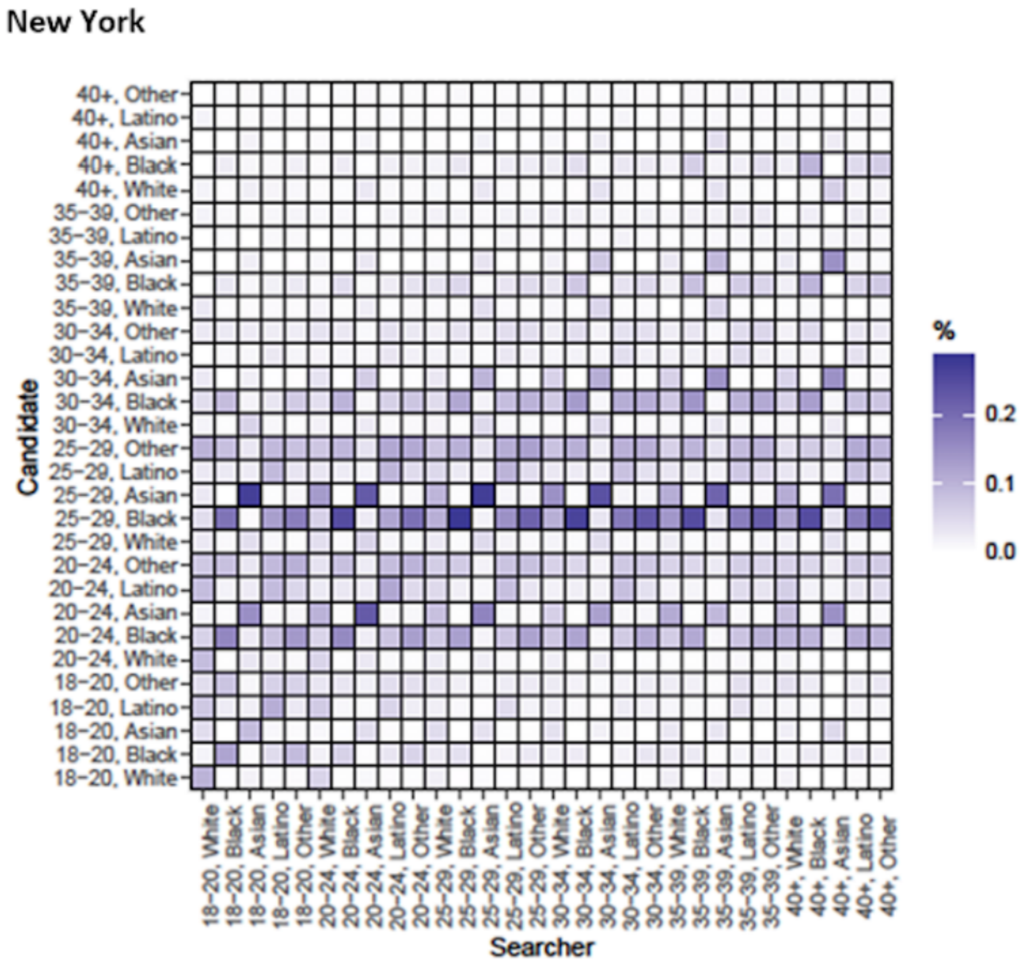
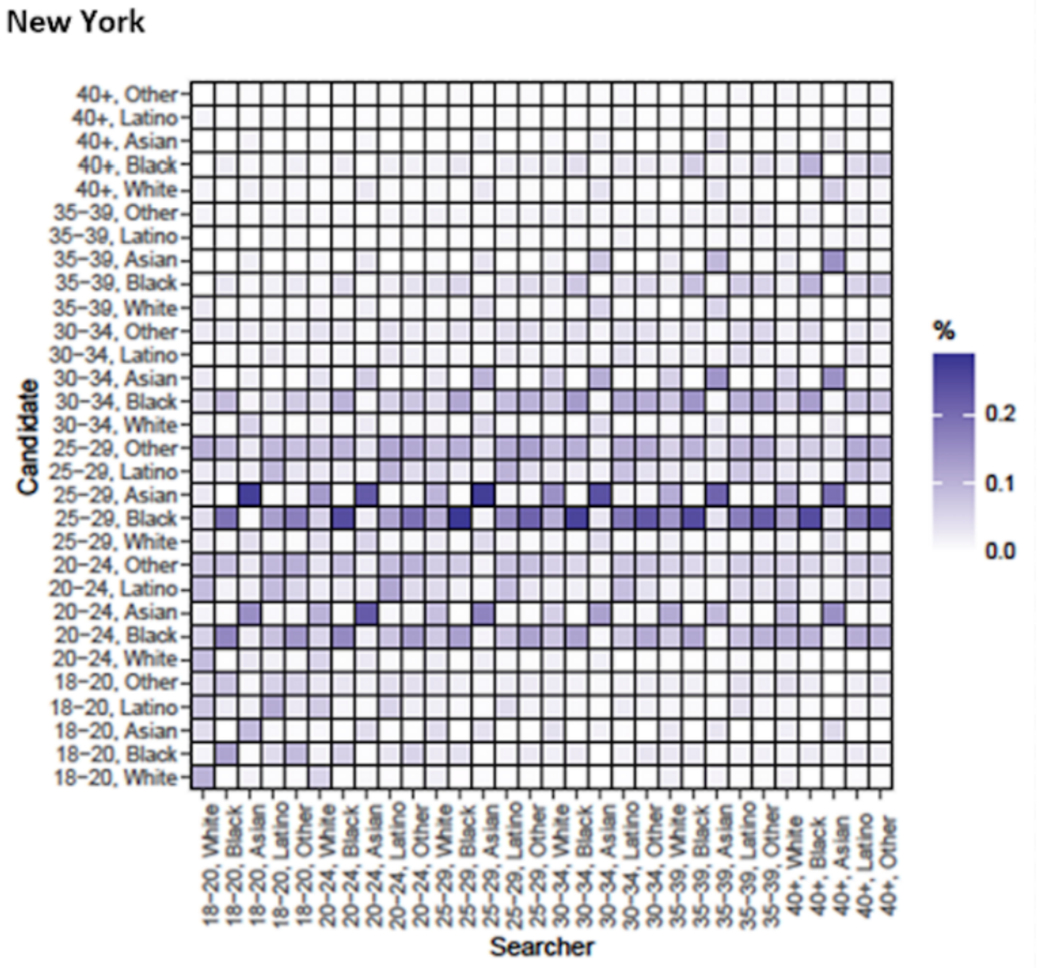


Figure 11. Assortativity by age (years) and race for chats initiated in New York. Heat maps show the relationship between searcher's age and race group and the candidate's age and race group. All values are normalized by numbers of searchers in each age and race category (the total value of a column adds up to 100).



Discussion

Principal Findings

Our results indicate that the age preferences of MSM are relatively consistent across cities, that is, younger MSM profiles are more likely to be viewed and chatted with compared with older MSM, but the patterns of racial mixing are more variable. Specifically, we see that men tend to look at profiles and access details for other men on the platform from similar age, race, and age/race subgroups, but men initiate chats with men aged between 20 and 29 years most often, independent of searcher age. Assortativity patterns with regard to age were similar to other studies done with MSM across the country. A study by Tieu et al [10] in 6 US cities found that assortativity for age was estimated to be 0.20, within the range of age assortativity coefficients calculated in our study. We also see that Asian and black men tend to initiate chats with other Asian and black men in most cities. Cities such as Birmingham and Tampa may have different racial mixing patterns, with young black candidates being chatted with more often than other subgroups of the population (independent of searcher race).

Although we cannot, with certainty, generalize the interaction patterns of men in Web-based settings to real-world settings, Web-based partnership selection behavior may inform our understanding of real-world patterns in STI spread. Our observation of behavior on a GSN app, without relying on the self-reporting of behavior, is a major strength of our analysis. Similar to previous studies, our results confirm that black men are the most assortative by race with regard to viewing profiles and chatting with candidates. Specifically, our results lend evidence to patterns that suggest that black MSM are highly assortative in their preferences, consistent with propagation of HIV and other STIs on their sexual networks [7,9,29]. Previous studies found that black MSM often had exclusively black partners, had a majority of black partners, or were more likely to select a black partner [4,10]. We found that black searchers were highly assortative in both viewing profiles and chatting across all cities, supporting previous evidence to suggest that black MSM might have less diverse networks than other racial peers [4]. We also found evidence for high assortativity among Asian MSM in Boston, Los Angeles, New York, and San Francisco, suggesting that this may be a group needing targeted intervention. In cities like Tampa and Birmingham, where the overall proportion of black individuals is high, black candidates are being chatted with by all racial groups in these areas.

Viewing someone's profile is considered a latent interaction and has been found to be extremely common among users of Web-based social networks and is often not reciprocated [30].

Little to no research has been conducted to understand latent interactions on GSN apps targeted at MSM, often because of a lack of available data. A searcher may be more inclined to look at profile details or have a latent interaction with someone in whom they are interested but who may not be perceived as a socially acceptable partner (thus not initiating a chat) [31]. Although this pressure would be expected to matter less in a Web-based setting, because of a lack of observers, these pressures may still exist and should be investigated. It would be presumed that searchers who initiated a chat with a candidate had a reason to do so (ie, desire for interaction, whether on the Web or in person). The fact that more age-disparate chats were initiated suggests a longing for younger partners and friends. Evidence suggests that many of the decisions made in Web-based settings are made purely on the basis of the photograph available on the profile [32]. Although the photograph may indicate phenotypic clues to a candidate's race and age, we were not able to analyze the content of the photos in relation to the race and age declared of the profile.

Limitations and Strengths

Caution should be used when employing and interpreting assortativity coefficients as they may mask subgroup behaviors that are important to sexual behavior and STI spread. Specifically, assortativity coefficients are less useful in areas where 1 or 2 groups make up the majority of the population (as in Tampa and Birmingham, in our examples). Therefore, comparing assortativity coefficients across metropolitan areas may only be useful together with information on the distribution of race or age in these areas. In addition, we are only reporting on the behavior of users on a single mobile app. Different mobile apps aimed at connecting MSM have different cultures and composition of users. The observations made from this mobile app may not be generalizable to other groups and apps. An added limitation is that we were not able to examine the behavior of men who did not specify a race or an age in their user profile. These men may have had a reason not to display an age or name (perhaps because they phenotypically belonged to a particular race or age group). The mixing observed is, therefore, limited to those individuals who specified their age or racial group. Finally, assortativity and age preferences are reported here at the population level, potentially obscuring interindividual variability. For example, an individual with 3 to 4 searches (the median) may have different assortativity patterns than an individual with 15,000 searches. These differences in patterns were not explored in this analysis and will be explored in future analyses.

Despite these limitations, this analysis has a number of strengths that provide a unique contribution to the field. Specifically, the results of our analysis include a number of different geographic regions within the United States. Our data also include a large number of MSM of color, and specifically black men, which allows for an objective examination of their behavior on the app. The data we are using are not egocentric or sociometric network data but rather observations on Web-based behavior. Our analysis is, therefore, not likely to be affected by social desirability bias, distortion, self-reporting biases, or recall bias similar to many other analyses examining assortative partnering behaviors among MSM.

Conclusions

One-size-fits-all interventions aiming to target MSM of color may not work in all contexts or among all minority subgroups. Specific geographic and community-level interventions need to be tailored to complement the needs of each individual population. Although some generalizations can be made regarding Web-based behaviors across all cities, city-specific usage patterns and trends should be analyzed to create targeted and localized interventions that may make the most difference in the lives of MSM in these areas.

Acknowledgments

This study was supported by Grant Number U54GM088558 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

Authors' Contributions

All authors (NNA, YAR, DN, YHG, GRS, KM, and ML) were involved in the conception and the design of the analysis. NNA and ML were responsible for performing the analysis. All authors were responsible for the interpretation of the study results and the writing of the paper.

Conflicts of Interest

DN is a former employee of the company that owns the GSN app discussed in this paper.

References

1. Centers for Disease Control and Prevention. 2017. Sexually Transmitted Disease Surveillance 2017 URL: https://www.cdc.gov/std/stats17/default.htm [accessed 2019-07-31]
2. Millett GA, Peterson JL, Wolitski RJ, Stall R. Greater risk for HIV infection of black men who have sex with men: a critical literature review. Am J Public Health 2006 Jun;96(6):1007-1019. [doi: 10.2105/AJPH.2005.066720] [Medline: 16670223]

3. Millett GA, Peterson JL, Flores SA, Hart TA, Jeffries WL, Wilson PA, et al. Comparisons of disparities and risks of HIV infection in black and other men who have sex with men in Canada, UK, and USA: a meta-analysis. Lancet 2012 Jul 28;380(9839):341-348. [doi: 10.1016/S0140-6736(12)60899-X] [Medline: 22819656]
4. Newcomb ME, Mustanski B. Racial differences in same-race partnering and the effects of sexual partnership characteristics on HIV risk in MSM: a prospective sexual diary study. J Acquir Immune Defic Syndr 2013 Mar 1;62(3):329-333 [FREE Full text] [doi: 10.1097/QAI.0b013e31827e5f8c] [Medline: 23187943]
5. Newman ME. Mixing patterns in networks. Phys Rev E 2003 Feb 27;67(2 Pt 2):026126. [doi: 10.1103/physreve.67.026126]
6. Tieu HV, Xu G, Bonner S, Spikes P, Egan JE, Goodman K, et al. Sexual partner characteristics, serodiscordant/serostatus unknown unprotected anal intercourse and disclosure among human immunodeficiency virus-infected and uninfected black men who have sex with men in New York city. Sex Transm Dis 2011 Jun;38(6):548-554. [doi: 10.1097/OLQ.0b013e318203e2d7] [Medline: 21217419]
7. Tieu HV, Murrill C, Xu G, Koblin BA. Sexual partnering and HIV risk among black men who have sex with men: New York city. J Urban Health 2010 Jan;87(1):113-121. [doi: 10.1007/s11524-009-9416-x] [Medline: 19949990]
8. Tieu HV, Nandi V, Hoover DR, Lucy D, Stewart K, Frye V, NYC M2M Study Team. Do sexual networks of men who have sex with men in New York city differ by race/ethnicity? AIDS Patient Care STDS 2016 Jan;30(1):39-47 [FREE Full text] [doi: 10.1089/apc.2015.0237] [Medline: 26745143]
9. Mimiaga MJ, Reisner SL, Cranston K, Isenberg D, Bright D, Daffin G, et al. Sexual mixing patterns and partner characteristics of black MSM in Massachusetts at increased risk for HIV infection and transmission. J Urban Health 2009 Jul;86(4):602-623 [FREE Full text] [doi: 10.1007/s11524-009-9363-6] [Medline: 19466554]
10. Tieu HV, Liu TY, Hussen S, Connor M, Wang L, Buchbinder S, HPTN 061. Sexual networks and HIV risk among black men who have sex with men in 6 US Cities. PLoS One 2015;10(8):e0134085 [FREE Full text] [doi: 10.1371/journal.pone.0134085] [Medline: 26241742]
11. Raymond HF, McFarland W. Racial mixing and HIV risk among men who have sex with men. AIDS Behav 2009 Aug;13(4):630-637 [FREE Full text] [doi: 10.1007/s10461-009-9574-6] [Medline: 19479369]
12. Mustanski B, Birkett M, Kuhns LM, Latkin CA, Muth SQ. The role of geographic and network factors in racial disparities in HIV among young men who have sex with men: an egocentric network study. AIDS Behav 2015 Jun;19(6):1037-1047 [FREE Full text] [doi: 10.1007/s10461-014-0955-0] [Medline: 25430501]
13. Bingham TA, Harawa NT, Johnson DF, Secura GM, MacKellar DA, Valleroy LA. The effect of partner characteristics on HIV infection among African American men who have sex with men in the young men's survey, Los Angeles, 1999-2000. AIDS Educ Prev 2003 Feb;15(1 Suppl A):39-52. [doi: 10.1521/aeap.15.1.5.39.23613] [Medline: 12630598]
14. Eaton LA, Earnshaw VA, Maksut JL, Thorson KR, Watson RJ, Bauermeister JA. Experiences of stigma and health care engagement among black MSM newly diagnosed with HIV/STI. J Behav Med 2018 Aug;41(4):458-466. [doi: 10.1007/s10865-018-9922-y] [Medline: 29626312]
15. Sang JM, Matthews DD, Meanley SP, Eaton LA, Stall RD. Assessing HIV stigma on prevention strategies for black men who have sex with men in the United States. AIDS Behav 2018 Dec;22(12):3879-3886. [doi: 10.1007/s10461-018-2181-7] [Medline: 29860555]
16. Grosskopf NA, LeVasseur MT, Glaser DB. Use of the internet and mobile-based 'apps' for sex-seeking among men who have sex with men in New York city. Am J Mens Health 2014 Nov;8(6):510-520. [doi: 10.1177/1557988314527311] [Medline: 24658284]
17. Phillips G, Magnus M, Kuo I, Rawls A, Peterson J, Jia Y, et al. Use of geosocial networking (GSN) mobile phone applications to find men for sex by men who have sex with men (MSM) in Washington, DC. AIDS Behav 2014 Sep;18(9):1630-1637. [doi: 10.1007/s10461-014-0760-9] [Medline: 24682866]
18. Gudelunas D. There’s an app for that: the uses and gratifications of online social networks for gay men. Sex Cult 2012 Jan 14;16(4):347-365 [FREE Full text] [doi: 10.1007/s12119-012-9127-4]
19. Miller B. 'They’re the modern-day gay bar': exploring the uses and gratifications of social networks for men who have sex with men. Comput Human Behav 2015 Oct;51:476-482. [doi: 10.1016/j.chb.2015.05.023]
20. Duncan DT, Park SH, Hambrick HR, Dangerfield II DT, Goedel WC, Brewer R, et al. Characterizing geosocial-networking app use among young black men who have sex with men: a multi-city cross-sectional survey in the southern United States. JMIR Mhealth Uhealth 2018 Jun 14;6(6):e10316 [FREE Full text] [doi: 10.2196/10316] [Medline: 29903702]
21. Zou H, Fan S. Characteristics of men who have sex with men who use smartphone geosocial networking applications and implications for HIV interventions: a systematic review and meta-analysis. Arch Sex Behav 2017 May;46(4):885-894. [doi: 10.1007/s10508-016-0709-3] [Medline: 27043837]
22. Muessig KE, Pike EC, Fowler B, LeGrand S, Parsons JT, Bull SS, et al. Putting prevention in their pockets: developing mobile phone-based HIV interventions for black men who have sex with men. AIDS Patient Care STDS 2013 Apr;27(4):211-222 [FREE Full text] [doi: 10.1089/apc.2012.0404] [Medline: 23565925]
23. Hamilton DT, Morris M. Consistency of self-reported sexual behavior in surveys. Arch Sex Behav 2010 Aug;39(4):842-860 [FREE Full text] [doi: 10.1007/s10508-009-9505-7] [Medline: 19588240]
24. Fenton KA, Johnson AM, McManus S, Erens B. Measuring sexual behaviour: methodological challenges in survey research. Sex Transm Infect 2001 Apr;77(2):84-92 [FREE Full text] [doi: 10.1136/sti.77.2.84] [Medline: 11287683]

25. Schroder KE, Carey MP, Vanable PA. Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. Ann Behav Med 2003 Oct;26(2):104-123 [FREE Full text] [doi: 10.1207/S15324796ABM2602_03] [Medline: 14534028]
26. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. J Med Internet Res 2011 May 13;13(2):e38 [FREE Full text] [doi: 10.2196/jmir.1797] [Medline: 21571632]
27. Doherty IA, Schoenbach VJ, Adimora AA. Sexual mixing patterns and heterosexual HIV transmission among African Americans in the southeastern United States. J Acquir Immune Defic Syndr 2009 Sep 1;52(1):114-120 [FREE Full text] [doi: 10.1097/QAI.0b013e3181ab5e10] [Medline: 19506485]
28. Schneider JA, Cornwell B, Ostrow D, Michaels S, Schumm P, Laumann EO, et al. Network mixing and network influences most linked to HIV infection and risk behavior in the HIV epidemic among black men who have sex with men. Am J Public Health 2013 Jan;103(1):e28-e36 [FREE Full text] [doi: 10.2105/AJPH.2012.301003] [Medline: 23153147]
29. Clerkin EM, Newcomb ME, Mustanski B. Unpacking the racial disparity in HIV rates: the effect of race on risky sexual behavior among black young men who have sex with men (YMSM). J Behav Med 2011 Aug;34(4):237-243. [doi: 10.1007/s10865-010-9306-4] [Medline: 21107898]
30. Jiang J, Wilson C, Wang X, Sha W, Huang P, Dai Y, et al. Understanding latent interactions in online social networks. ACM Trans Web 2013 Oct 01;7(4):1-39. [doi: 10.1145/2517040]
31. Brown G, Maycock B, Burns S. Your picture is your bait: use and meaning of cyberspace among gay men. J Sex Res 2005 Feb;42(1):63-73. [doi: 10.1080/00224490509552258] [Medline: 15795806]
32. Miller B. 'Dude, where's your face?' self-presentation, self-description, and partner preferences on a social networking application for men who have sex with men: a content analysis. Sex Cult 2015 Apr 10;19(4):637-658. [doi: 10.1007/s12119-015-9283-4]

Abbreviations

GPS: global positioning system
GSN: geosocial networking
MSM: men who have sex with men
STI: sexually transmitted infection

Edited by G Eysenbach; submitted 22.02.19; peer-reviewed by J Opoku, P Janulis; comments to author 25.04.19; revised version received 08.07.19; accepted 19.07.19; published 12.09.19.

Please cite as:
Abuelezam NN, Reshef YA, Novak D, Grad YH, Seage III GR, Mayer K, Lipsitch M
Interaction Patterns of Men Who Have Sex With Men on a Geosocial Networking Mobile App in Seven United States Metropolitan Areas: Observational Study
J Med Internet Res 2019;21(9):e13766
URL: https://www.jmir.org/2019/9/e13766/
doi:10.2196/13766
PMID:31516124

Original Paper

A Theoretical Twist on the Transparency of Open Notes: Qualitative Analysis of Health Care Professionals' Free-Text Answers

Gudbjörg Erlingsdóttir[1], PhD; Lena Petersson[1], MSc; Karin Jonnergård[2], PhD

[1]Department of Design Sciences, Lund University, Lund, Sweden

[2]Department of Business Administration, Lund University, Lund, Sweden

Corresponding Author:
Lena Petersson, MSc
Department of Design Sciences
Lund University
Box 118
Lund
Sweden
Phone: 46 46 222 05 33
Email: Lena.Petersson@design.lth.se

Abstract

Background: The New Public Management movement strove for transparency so that policy makers and citizens could gain insight into the work and performance of health care. As the use of the electronic health record (EHR) started to diffuse, a foundation was laid for enhanced transparency within and between health care organizations. Now we appear to be experiencing a new kind of transparency in the health care sector. Many health care providers offer their patients online access to their EHRs (here referred to as Open Notes). The Open Notes system enables and strives for transparency between the health care organization and the patient. Hence, this study investigates health care professional (HCP) perceptions of Open Notes and deepens the understanding of the transparency that Open Notes implies.

Objective: Based on two survey studies of HCP perceptions of Open Notes, this paper aims to deepen the academic writing on the type of transparency that is connected to Open Notes.

Methods: HCPs in adult psychiatry in Region Skåne, Sweden, were surveyed before and after implementation of Open Notes. The empirical material presented consists of 1554 free-text answers from two Web surveys. A qualitative content analysis was performed.

Results: The theoretically informed analysis pivots around the following factors connected to transparency: effectiveness; trust; accountability; autonomy and control; confidentiality, privacy, and anonymity; fairness; and legitimacy. The results show that free-text answers can be sorted under these factors as trade-offs with transparency. According to HCPs, trade-offs affect their work, their relationship with patients, and not least, their work tool, the EHR. However, since many HCPs also state that they have not met many patients, and in some cases none, who have read their EHRs, these effects seem to be more connected to the possibility (or threat) of transparency than to the actual effectuated transparency.

Conclusions: The implementation (or reform) of Open Notes is policy driven while demanding real-time transparency on behalf of citizens/patients and not the authorities, which makes this particular form of transparency quite unique and interesting. We have chosen to call it governed individual real-time transparency. The effects of Open Notes may vary between different medical specialties relative to their sensitivity to both total and real-time transparency. When HCPs react by changing their ways of writing notes, Open Notes can affect the efficiency of the work of HCPs and the service itself in a negative manner. HCP reactions are aimed primarily at protecting patients and their relatives as well as their own relationship with the patients and secondly at protecting themselves. Thus, governed individual real-time transparency that provides full transparency of an actual practice in health care may have the intended positive effects but can also result in negative trade-offs between transparency and efficiency of the actual practice. This may imply that full transparency is not always most desirable but that other options can be considered on a scale between none and full transparency.

(J Med Internet Res 2019;21(9):e14347) doi:10.2196/14347

KEYWORDS

electronic health record; eHealth; transparency; mental health; Open Notes; psychiatry; health care professionals; Web survey; free-text answers

Introduction

Background

One trend in the New Public Management (NPM) movement was an increased demand for transparency [1,2]. The first wave of transparency was connected to performance evaluations and the use of visualizing technologies as a means of regulation, monitoring, and accountability [3-5]. In other words, the transparency implied that policy makers and citizens would gain insight into the work and performance of health care. Stressing the positive effects of transparency, voices were heard advocating for extending its areas of application [6]. As use of the electronic health record (EHR) started to diffuse in Sweden about twenty years ago [7] and about ten years ago internationally [8], a foundation was laid for enhanced transparency within and between health care organizations as well as in relation to their environment [9]. Currently, we appear to be on the next step of this development, as many health care organizations provide patients with online access to their EHRs (here referred to as Open Notes). In other words, transparency increases between the health care organization and the patient. This also implies that a number of different types of transparency are at play simultaneously in the health care systems regarding both what is made visible and to whom.

In the literature, transparency related to visualizing technologies [10,11] and, to a certain degree, the EHR [12,13] has been investigated empirically and discussed in theoretical terms. When it comes to Open Notes, however, most of the research is still at an early stage and either descriptive [14-17] or normative [18,19], with few exceptions [20,21]. Professionals and researchers in the field have interpreted Open Notes as a means of enhancing the transparency of health care professionals' (HCPs) work and their work tool [13,22-26]. Research on transparency rising from Open Notes reform has been published [27-34], but few attempts have been made to theorize the transparency based on empirical material. Previous research in the area has found that perceptions of Open Notes differ between HCPs and patients, HCPs are more skeptical than patients, and the benefits for patients depend on their condition and ability to take advantage of the technology [35].

Based on two survey studies of HCP perceptions of Open Notes, the aim of this article is to deepen the academic writing on the type of transparency connected to Open Notes. First, because of the gap in the literature indicated above. Second, because the transparency that the Open Notes system permits distinguishes itself from the earlier systems connected to accountability or organizational internal translucence. While Open Notes is advocated to be a part of patient empowerment, meant to engage patients in their health care, the ones that become transparent—the HCPs—are not supposed to be affected. This is different from other types of transparency. In addition, Open Notes has a restricted transparency. In the Open Notes system, the only ones who can view the notes are the patients. These features make the transparency of Open Notes unique and important to study in its own right. The third reason is because the early research has indicated that the noneffect of Open Notes on EHRs as a work tool does not agree with the experience of the HCPs [15]. This gives practical relevance for studying the phenomenon.

We chose to examine perceptions of HCPs when the Open Notes service was implemented in psychiatric care in Sweden. This is because, unlike other parts of health care, psychiatric care has a large degree of vulnerable patients. It is thus likely that some concerns are more salient in the psychiatric context than in other parts of health care where they occur less often because the patients there are considered sufficiently capable of benefitting from the Open Notes service. Previous research has shown that professionals aggregate their concerns before Open Notes is implemented [35]. Thus, we sent out two surveys: one preimplementation and one post. This afforded us the opportunity to compare HCP perceptions before and after implementation and thereby separate belief from experience. In this paper, we analyze the answers to open-ended questions in the two surveys in order to capture the perceptions and deepen the understanding of the transparency that Open Notes implies.

Transparency in Health Care

Previously, transparency in health care has primarily been mentioned in connection with the NPM wave, where one of the main aims was to shift from what Ouchi [36] calls clan control together with bureaucracies to market control. This in turn led to a demand for visible or transparent processes in health care and quality assurances in communication between the producer and the market. However, production processes in health care have, by tradition, mainly been black boxed and controlled and peer reviewed by the medical professions [37,38]. Thus, the shift to NPM led to reactions and resistance from the professionals [39] because they interpreted the government's aim for enhanced transparency as being in conflict with their autonomy. According to Levay and Waks [11], efforts to enhance the transparency of professional practice will not only make the practice more visible but will also make it more accessible for external audit and control, which may explain the professionals' negative reaction to NPM.

The most common academic use of the term or concept transparency in organizational theory comes from the body of literature that views it in terms of its relation to or consequence of governance [40], as in the NPM case. However, the transparency connected to Open Notes does not totally fit into the governance discourse because it does not result in generally open or accessible information for all citizens or for government agencies. In addition, the request for transparency is not aimed at the health care organizations, as such, but at the use of a technology (the Open Notes system) that allows patients to read their health records. The citizen/patient, in turn, can only access his or her own medical information. This can be seen as a transparency on one-to-one basis, that is, between the HCP who writes the note and the patient who reads it. Still, the reform is a consequence of electronic health (eHealth) policies in Sweden where access to more information is expected to empower patients/citizens and increase their participation when it comes to their own health. In this respect, the enhanced transparency can be regarded as a result of governing to some degree. According to Blomgren and Sundén [41], the concept of transparency in this way is often associated with the ideals of democracy, accountability, fairness, informed citizenship, or patient rights.

Transparency has several different definitions and meanings [42]. We have chosen to use Florini's [43], which states that transparency "refers to the degree to which information is available to outsiders that enables them to have informed voice in decisions and/or to assess the decisions made by insiders." The definition fits the Open Notes system, the aim of which is to make information available to patients that will enable them to have an informed voice and assess the decisions made by the insiders (the HCPs). Hansen and Flyverbom [44] point out that in the digital age, humans and technologies (materials) together produce transparency through mediating technologies, which also can be recognized in our Open Notes case. Flyverbom et al [45], in turn, state that mediating technologies and the transparency they enable can allow observational control. The authors further propose "that the relationship between power and transparency is best understood in terms of both 'observational control' and 'regularizing control.'" The production of transparency through mediated technologies stands in contrast to the traditional secrecy that Ball [42] states has existed around health records. Thus, the secrecy is not primarily aimed to protect the professionals who write the notes but to protect the patient about whom the notes are written. The Open Notes system has the same aim, in one sense, as the notes are kept secret from all but the patients about whom they are written. At the same time, HCP notes become more transparent than before. This means that secrecy is intact regarding patients but not professionals.

To understand the relationships between different actors involved in the production and use of the Open Notes transparency and the type of transparency it leads to, we were inspired by Heald [46]. Heald constructs an anatomy of transparency considering (1) its direction (upward, downward, inward, and outward), (2) the variety of transparency formulated in three dichotomies (event/process, retrospect/real time, and nominal/effective), and (3) the habitat of transparency.

Heald [46] points out that process transparency may cause efficiency losses both directly (providing the information) and indirectly (more expensive working practices are adopted) and that, in real-time transparency, the "accounting window is always open and surveillance is continuous." This in turn makes it difficult for the organization (and its members) to focus entirely on its productive activities [46]. Yet another consequence of enhanced transparency is when professionals become more self-monitoring in their everyday work, which in turn can affect them in their core identity [11]. According to Heald [47], there may also be tradeoffs with transparency including "effectiveness; trust; accountability; autonomy and control; confidentiality, privacy, and anonymity; fairness; and legitimacy." Heald also states that full transparency may not be equal to the most beneficial transparency.

Methods

Setting

HCPs in adult psychiatry in Region Skåne were surveyed before and after the implementation of the eHealth service. Adult psychiatry is a part of the Division of Psychiatric Care in Region Skåne in southern Sweden. In 2017, there were more than 575,000 appointments, of which almost one-fifth were with a doctor in the Division, with more than 56,000 unique patients. Patients in adult psychiatry were offered online access to their Open Notes in October 2015, and patients in forensic psychiatry and parents of patients in child and youth psychiatry in Region Skåne were offered the service in February 2019.

Empirical Materials

The empirical material in this study consists of free-text responses from two Web surveys. The demographics of the respondents, quantitative findings from these surveys, and results from two of the general open-ended questions in the postimplementation survey on how the service influenced patient groups with different diagnoses have been published elsewhere with detailed information about the technical prerequisites of the service, the settings, and the administration of the surveys [14,15]. Thus, the results of these two open-ended questions are not included here. Table 1 presents a summary of information about the two Web surveys.

Table 1. Overview of the preimplementation and postimplementation Web surveys.

Aspects	Preimplementation (baseline) survey	Postimplementation survey
Number of questions	34 fixed-choice questions and 3 open-ended questions	44 fixed-choice questions and 20 open-ended questions
Timeframe in which to answer	September 18 to October 2, 2015	March 16 to April 22, 2017
Population	3017 health care professionals	2521 health care professionals
Professional groups included in the population	assistant nurses, doctors, medical secretaries, nurses, occupational therapists, physical therapists, psychologists, and social workers	assistant nurses, doctors, medical secretaries, nurses, occupational therapists, physical therapists, psychologists, social workers, and unit managers
Response rate	28.86% (871/3017)	27.73% (699/2521)
Number of free-text answers	388 free-text answers	1166 free-text answers

The postimplementation survey was based on the baseline survey to permit comparisons between the expectations before implementation and after. Furthermore, both the baseline and postimplementation surveys were based on the surveys developed by the Open Notes Project in the United States [17,48-50]. In both cases, the original surveys were translated and adjusted to fit the Swedish context.

There can be different kinds of open-ended questions in surveys. The most common is the general question where respondents are asked to elaborate on the overall topic of the survey. Another is an expansion question that follows a fixed-choice question [51]. There were 3 open-ended questions in the baseline survey. They were all general in nature and did not relate to any specific fixed-choice question. In the postimplementation survey, there were 20 open-ended questions (the results of 18 of which are included here); 12 were expansions of specific fixed-choice questions in which the respondents were asked to elaborate on the answer given. The remaining 6 open-ended questions were general and did not relate to any specific fixed-choice questions. Both Web surveys were designed so that the respondents could choose not to answer all the questions.

Web surveys make it easy to offer respondents an opportunity to tell the researchers what is on their minds with respect to the survey subject; however, it is still uncommon that free-text answers are systematically coded and analyzed [52]. Free-text answers from open-ended questions in surveys can be described as elicited documents in which the research participants produce the data by answering [53]. One advantage of using this method to gather data for qualitative analysis that Charmaz [53] identifies is that the respondents are able to report things that they, for different reasons, would not have told an interviewer in person. Thus, these elicited texts can reveal sensitive information, and a person can choose to write as much or as little about the subject as they wish. This method for gathering data is appropriate when the participants have a stake in the topic addressed, when they have experience in the area, and when they have the writing skills to express themselves and their views on the topic [53].

Analysis

The empirical material was analyzed by means of qualitative content analysis. The material was coded by hand in several stages by all three authors. The units of analysis in this study were free-text answers from the baseline and postimplementation surveys. The units of analysis in a qualitative content analysis are identifiable with a low degree of interpretation [54]. When starting the analysis, authors focused on the manifest content of the free-text answers. Manifest content analysis is appropriate when researcher interest is directed at the visible and obvious meaning of the text [54]. Each free-text answer was a meaning unit and there was no need for condensation; the meaning units were statements from the survey respondents with concentrated content in every sentence. Thus, the three authors began to code and then sort out the coded manifest content into categories. A category is on a descriptive level and answers the question “What?” [54]. The research process in this study can be described as abductive, with an initial inductive approach during the coding and categorization of the empirical material from the baseline survey. In the next step, these inductively created categories were used in the analysis of the empirical material from the postimplementation survey. Then a deductive approach was applied during the creation of the final categories that are presented in the results section. An abductive approach is used in a qualitative content analysis when the analysis moves between inductive and deductive approaches during different stages of the research process [55]. The quotes of the respondents used in this paper were translated from Swedish to English by a native English speaking, professional proofreader. They have been slightly edited to improve readability.

Results

Overview

The results are structured according to the seven factors described by Heald [47] that can result in trade-offs and synergies with transparency: effectiveness; trust; accountability; autonomy and control; confidentiality, privacy, and anonymity; fairness; and legitimacy. According to Heald, the trade-offs between transparency and these factors are generally positive while there are some conditions where the tradeoffs may become negative. The anticipated effects are briefly described here in connection with each factor. To illustrate the difference between how the professionals anticipated the effects of the Open Notes service before implementation (baseline survey) and how they described their experience one and a half years after (postimplementation survey), we present quotes from the professionals, sorted under before implementation and after implementation. A summarizing analysis is presented at the end of each factor section.

Effectiveness

The relationship between transparency and effectiveness is mostly seen as positive since public scrutiny is presumed to raise performance. However, transparency that reveals the operational process in an organization may “affect behavior in an unanticipated way,” according to Heald [47]. As shown in the theoretic section of the paper, Open Notes concerns the operational process of the HCPs.

Before Implementation

HCPs mainly expressed concerns about how they thought Open Notes would negatively affect the efficiency of their work. They felt that more time would be spent on documentation because they anticipated they would have to change their way of expressing themselves and be less candid in their writing.

> *More valuable work time is likely to be spent on record keeping because I will have to be more careful in how I formulate what I write.*

HCPs anticipated that Open Notes would affect the appointment and that the treatment would pivot on the notes in the EHR instead of the patient’s health condition and treatment.

> *I see great danger in that the pact with the patient will be broken rather than strengthened, and that much of the work after that will be focused on the notes rather than on the rehabilitation/recovery from the psychiatric diagnosis.*

HCPs felt that when patients were able to read their EHRs online, the contents would change and the health record would not be as efficient a work tool as before for the HCPs themselves.

> *I'm worried that we won't be able to use the health record as the tool it is meant to be.*

But there were also comments about the positive effects that could enhance the efficiency.

> *We will gladly go through the notes with the patient. It is valuable for both the patient and coworkers. You can explain and come to a consensus.*

After Implementation

One of the most common comments in the postimplementation survey was that the HCPs had not met any (or very few) patients who had read their notes. This was interpreted as lack of efficiency of the service itself and for the patients. This differed from the baseline survey where this had not occurred as an anticipated effect of the Open Notes service.

> *Very few patients make use of their notes = very little effect on the patients.*
>
> *Never met a patient who said they read their notes.*
>
> *I've only heard from a couple of patients that they have read them.*

In contrast, several comments confirmed experiences supporting loss-of-efficiency concerns found in the baseline survey. HCPs put more effort and time into writing notes in the EHR.

> *I weigh my words. How does one describe the manic patient? And the basis of a certificate of illness? Or the young patient with a suspected onset of a psychosis disease?*
>
> *Adds more time to composing each dictation and for all patients.*
>
> *Try to use expressions that are easier to understand which can take a little more time!*

HCPs described how appointments were affected. There was a difference in expression from the baseline study, however, as the quotes were more illustrative, including the patients in the examples.

> *When it happens that a patient reads his notes, and, for example, becomes deeply depressed, this is a problem that I have to spend quite a lot of time on sorting it all out. This doesn't happen often, but when it does, it takes time away from the treatment.*

HCPs also confirmed that the EHR as a work tool had been affected

> *The notes become watered down and fall short in their function as a tool for the profession.*
>
> *The notes becomes a poorer quality work document for me as a doctor. I can no longer write down all the tentative diagnoses and hypotheses. Need to simplify the language. Every week.*

HCPs reported their stress levels were raised, affecting their work environment.

> *For my own part, there has been an increase in stress to sign off on the notes entries that I'm about 2 months behind on.*

Also, they reported that Open Notes affected the patients' feelings and perceptions of their illnesses.

> *Unfortunately, the patients can get upset when they read their notes, even though they are written respectfully and according to all the existing rules of the art. One patient, for example, suddenly felt worse when she read that she has a chronic risk for suicide. She felt that it meant she was a hopeless case.*

Even though most comments were negative, there were some positive ones about how the service could be used to make communication in the care situation more effective.

> *I recommend the notes to the patients as a memory help instead of giving them slips of paper about what we have agreed on. In some cases we agree that I don't need to send the lab test results as a regular mail letter, but that the patient can instead read about them in her or his notes when I have commented on/assessed the results.*
>
> *I encourage the patient to read often to make sure that I have understood things right.*

Effectiveness Analysis

There were three main concerns that the HCPs expressed both before and after the implementation. First, the loss of effectiveness when entering notes into the EHR because they became less candid in the way they wrote. Second, that the appointment was less efficient because it evolved to a larger degree around the notes in the EHR. Third, that the health records became a less efficient work tool for themselves and in communication with other professionals because they were watered down and thus less informative. This makes it clear that the transparency in Open Notes reveals the operational process of the professionals (in Heald's vocabulary [46]), affects the behavior of HCPs and patients, and impacts the effectiveness of the professionals' work and the EHR as a work tool. The difference between the baseline and postimplementation surveys is that in the latter, HCPs also described the perceived effects on the patients from reading the EHRs. That, in turn, could affect the care process and make it less effective.

Trust

Transparency was viewed as being positive in building trust in monetary issues, among others. One of Heald's examples of a transparency trade-off is that transparency can undermine professional credibility by exposing professional errors (or perceptions thereof) in health care [47].

Before Implementation

A main concern of the HCPs was how misunderstandings or dissatisfaction with the content of the EHR would affect the patients' trust and relationship to caregivers in general and the HCP in particular.

> *I am most afraid that a misunderstanding or an unpleasant feeling when reading can, in the blink of*

> *an eye, destroy a relationship between the patient and myself that we have built up over several years with some difficult patients. That would be a shame!*

HCPs also suspected this might lead to patients not seeking care even though they need it.

> *A great risk that some patients will become distrustful of health care and not seek care again when they really need it.*

But there were also comments that the service can assist in building mutual trust with the patient.

> *Psychiatry needs transparency and needs to learn that the patient should be part of the care plan and has the right to information, just as with all other care. This can contribute to the improvement of methods for the staff, and a greater understanding of the patients' need for transparency and the possibility to influence the care and better cooperation between the staff and patient.*

After Implementation

Comments in the postimplementation survey more or less confirmed the concerns that the HCPs had in the baseline survey: patients could react negatively to the content of the EHR notes and this could lead to mistrust toward the caregiver. However, there were more comments about patients' inability to understand the notes.

> *The hardest part is that there are many who do not understand what they read; [they] misunderstand and can feel offended because of that and the way it is written. Many have a need to be understood and be seen. When they then read something that they take to be the opposite, they lose trust.*

Still, some comments indicate that Open Notes could also be used in a positive way to establish trust.

> *The best part is to be able to point out that we have a collaboration without a hidden agenda from health care.*

Trust Analysis

Comments expressing thoughts (baseline) and experiences (post) of trust or distrust showed that HCPs were concerned about how the patients would interpret the content of their EHRs. In the postimplementation survey, they even expressed doubts about patients' ability to understand the notes. According to Heald [46], for transparency to be effective, the information that it makes accessible must be understood and used by the recipient. Thus, on the one hand, if the patients do not comprehend the information to the degree that the initiators and implementers of the Open Notes service anticipated, the transparency will not be as effective as planned. The misinterpreted or misunderstood information leads to mistrust between patient and HCPs. On the other hand, if patients do understand and can use the information in communication with their caregiver, trust can be enhanced.

Accountability

In the public sector, transparency is supposed to have trade-offs with political accountability but also with the distribution of power and resources. Accountability, however, is also about public service providers demonstrating their ability to generate outputs and "process values (such as due process, equity, participation, and deliberation)" [47].

Before Implementation

Comments related to accountability mainly evolved around how the HCPs would become accountable to the readers of the notes (the patients) and/or themselves as their notes became more visible.

> *I hope that the staff correct inaccuracies if the patient points them out. It is important for the patient that he or she feels that the personnel have understood them correctly. It is about time that the patients gain insight and in so doing, knowledge of the care around them.*
>
> *How should I write the notes so that they meet all the requirements for the stakeholders? It is a work tool and is intended to ensure that relevant information is transferred to those who are involved in care. It must be able to hold up under a Health and Social Care Inspectorate examination. It has to meet my own professional requirements for accuracy and relevant content. And the patients have to be able to be read what I write without causing them harm! I have concerns that the notes will end up being vague and watered down. As a work tool, the notes would be considerably degraded, and that is nevertheless their foremost function! This worries me a lot!*

Positive effects were expected.

> *The positive thing is that the staff will become more aware to always document correctly, clearly, and respectfully, and that is very important.*

After Implementation

Comments clearly showed HCP concerns had been proven right. They thought about how they were perceived by the readers of their notes in the EHR and how they became more accountable for the writing of their notes.

> *The hardest part is the concern of being misunderstood when writing the notes; that my documentation becomes an obstacle in treatment and that I can no longer use it as a work tool.*

Accountability Analysis

It was obvious that HCPs, in both baseline and postimplementation surveys, were aware errors could exist in the EHR notes and patients could discover them. This was perceived as both positive, when patients could correct the errors, and negative, since the notes then could be perceived as output that was less valuable or accountable. HCPs also stated they had to be accountable to different stakeholders and it was difficult to know whom to please the most, since the different stakeholders (patients, oneself as a professional, other professionals) may have different needs and/or standards for what is accountable. HCPs also had to follow the laws and regulations surrounding the EHR.

Autonomy and Control

In Heald's [47] description, trade-off between transparency and autonomy and control is depicted on an organizational level. Transparency is viewed as an external force that blurs the boundaries of the organization. Since the Open Notes service, however, operates on a one-to-one basis between HCP and patient, we chose to interpret autonomy and control on the individual level—that is, on the HCP level in relation to the external control that transparency toward the patient implies as well as the professionals' control over their work tool and autonomy over how to write their notes.

Before Implementation

Some comments in the baseline study indicated that HCPs were worried about losing control over their work tool (ie, the EHR) when the patients were given online access to it.

> *I'm worried that we won't be able to use the notes as the work tool it is meant to be.*

After Implementation

Comments in the postimplementation survey are in line with HCP concerns in the baseline survey about loss of control over their work tool, autonomy over what they can write in their notes, and control over when the information/notes in the EHR should be made visible and to whom.

> *I wish there was a function where we who do the documentation could decide when a note becomes visible in the Open Notes; that we ourselves had to approve it.*
>
> *Make sure that an inpatient on a unit cannot block his/her notes from the staff so that the only way medicine dispensation can be registered, for example, is by a staff member overriding the block each time.*

There were concerns about the effect this had on their work process.

> *It has become more difficult to write your assessment, as our patients tend to be quick to take offense. That's why the notes are censured.*

But HCPs also mentioned that the service can give patients more control over their care (or even a false feeling of control).

> *Increase transparency and hopefully increase the ability to influence one's care plan.*
>
> *Create a false sense of control and overview of self [referring to the patient].*

Autonomy and Control Analysis

In both surveys, HCPs expressed their worries/experiences of loss of autonomy in how to write notes and control over their work tool. In the postimplementation survey, comments also expressed HCP frustration about not being able to control what information is visible for the patient and when it becomes visible. HCPs thus opposed their inability to control and alter the technical features of the Open Notes service. This also included the fact that patients could block their notes from the HCPs, which of course would complicate the work of the HCPs. On the positive side, there was hope that patients could use the EHRs to more actively participate in and gain better control over their own health/treatment. However, HCPs did not have much experience in their patients making use of this opportunity. It appears that HCP feelings of loss of autonomy and control were not matched by the expected gain in autonomy and control of the patients.

Confidentiality, Privacy, and Anonymity

Even this set of factors needs to be reinterpreted from the government level, described by Heald [46], to fit the Open Notes. Here, confidentiality is interpreted as being related to the content of the notes since they are confidential as is all information about the patient. The privacy aspect is very closely related to confidentiality, but here the patients can both be forced to and/or break their privacy themselves. Anonymity, however, relates mostly to HCPs and third parties and whether they can be anonymous toward the patient.

Before Implementation

When it comes to confidentiality, privacy, and anonymity, HCPs had concerns about patients that could be forced by their relatives or partners to share their EHRs. This would be a type of violation of the patient's privacy but would also impair the confidentiality between the HCP and the patient.

> *The patients who live with threats in close relationships will have a hard time taking the risk of telling how things really are; won't dare to name or say who is abusing them because that person can pressure them into giving up the password to their notes.*

Other actors, such as an employer, might also insist on reading the information.

> *Great risk for privacy violations when notes not only can be read by the patient, but also spread within and outside of the health care system to people who can misuse the information. For example, an employer who knows his/her employee has been admitted asks to read the patient's notes.*

Professionals' own needs for anonymity toward the patient were also considered.

> *I think there can be an increase in threats and abuse of staff when the name of the staff member who has written the note is included.*

After Implementation

Similarities and differences exist between baseline and postimplementation survey comments. There were still comments about the violation of patients' privacy.

> *Those of my patients who said they read their notes are not positive about being able to. They fear that unauthorized people will be able to read about them. One of the patients talks about it at every visit.*
>
> *There still is a little but difficult risk group of those who live with partners in a relationship where violence occurs.*

Concerns about HCP anonymity were reinforced.

> *The level of threat has grown worse and personal safety has deteriorated. I think it would have been better if it just said, for example, Nurse Marta in the part that patients can read. If you have an unusual last name, the patient can easily find your home address and that doesn't feel safe. Patients with addictions sometimes place impossible demands that cannot be met and even have contacts [on the outside] that they can activate to threaten or harm the staff.*

A new area that surfaced in the comments concerned third parties, such as partners or relatives, who share information about the patients that should be kept confidential from the patient.

> *I'm less candid about the information that relatives provide. That has sometimes been a big problem because it is important information that otherwise falls out of the system.*

Another new area was information about actions taken that needed to be kept confidential from the patient.

> *Patients in outpatient care who are acutely deteriorating and where we in health care, for example, send a notification of concern to social services or in some cases write an institutional psychiatric care certificate that calls for police assistance, where the patients can immediately read this in the notes at home and know that the police are on their way. That poses a lot of problems for us in outpatient care.*

Confidentiality, Privacy, and Anonymity Analysis

HCPs were concerned about the loss of patient privacy in terms of their relatives in both surveys. This would also result in loss of confidentiality between HCPs and patients since HCPs cannot guarantee the confidentiality of information given by the patients after the implementation of the Open Notes service. HCPs also expressed fear for the loss of anonymity for their own sake, since their full names would be exposed in the note. There is a difference between the baseline and postimplementation surveys, however, in that HCPs expressed concerns in the latter that third parties could be harmed since the patients' relatives could not be kept anonymous in Open Notes. They also feared that a patient might gain access to information about actions taken toward them by the police, information that should be kept confidential from the patient but was important for HCPs. HCPs thus expressed both a need to protect patients' privacy and confidentiality and the anonymity of third parties and confidentiality about information that ought to be withheld from the patient.

Fairness

Heald [47] points out that in the political debate "fairness is often taken to mean less inequality" but that fairness can also "be conceptualized in terms of rights, deserts, or needs." In the latter, transparency can invoke envy but in the former, it can stimulate actions to be taken against inequality.

Before Implementation

When it comes to fairness, HCPs mainly feared that patients would not have the same access to the service or no access at all.

> *Unfortunately, I think that those of our patients who have a computer (few patients) and succeed in getting in and looking (even fewer) will perhaps be more upset after reading their notes.*

Or that if they did read their EHRs, they would not be able to understand them.

> *Lack of knowledge and ability to understand the content in the notes can result in increased worry and unnecessary anxiety. Is it ethically right to leave the patient on his or her own to try and interpret the meaning of the notes?*

After Implementation

Postimplementation survey comments indicated concern that few patients have the competence and/or resources to access their EHRs online but also that these abilities differed between patient groups (with different diagnoses).

> *In the spirit of equality. Several patients don't have access to or the ability to read Open Notes.*

At the same time, HCPs stated that patients in adult psychiatry, just like those in nonpsychiatric care, should have access to their EHRs through the internet. It is a question of fairness.

> *Psychiatry is equated with somatic disorders and access to notes is part of removing the stigma.*

Fairness Analysis

One of the arguments for implementing Open Notes in psychiatry was that adult psychiatry patients should have access to their EHRs to achieve equality between psychiatric and nonpsychiatric patients. This was also seen as positive in the HCP comments. Still, the comments in the baseline survey displayed a distrust from HCPs in the patients' ability to understand and make use of the information in their EHRs. In the postimplementation survey, HCPs, in addition to the above, pointed out that many of the psychiatric patients did not have the material resources to access their EHRs because they may lack computers, the necessary identification to log in to the service, and/or the cognitive ability to understand and manage the information it provides. There were thus both positive and negative connotations related to equality or fairness compared to patients in nonpsychiatric care and even between patients with different diagnoses within psychiatric care.

Legitimacy

According to Heald [47], transparency can legitimize certain institutions, organizations, or their actions. Here we interpret legitimacy as that of the implementation of the Open Notes service and the service itself in the eyes of HCPs.

Before Implementation

Before implementation, comments were about the service itself and its implementation and were positive and negative.

> *Poor basis for the decision. Why should we do this? How does it make health care more user controlled/patient safe? Who are the notes for? Should the notes be a work tool between professionals where working hypotheses can be written with the knowledge that those who read the notes have been trained in the language that is used and the content in general? Or should the notes be a diary for the patients?*
>
> *I think it is very good and will use it as an opportunity.*

After Implementation

Postimplementation comments were similar to those in the baseline survey. They focused on the implementation and the service itself, why it would not suit the adult psychiatry practice, and its effects.

> *A system that was introduced without at all taking into account what the employees in psychiatry think and that only harms the patient and makes it harder to do a good job.*

Legitimacy Analysis

The legitimacy of Open Notes was questioned by HCPs on the grounds that the implementation was not sufficiently prepared, other measures were more necessary, and the enhanced transparency negatively affected the professionals' work tool (the EHR). For some, however, the service also was viewed as offering possibilities.

Discussion

Principal Findings

As we have shown, HCP comments in the free-text answers in both the baseline and postimplementation surveys can easily be sorted under Heald's different factors as trade-offs with transparency. This is not surprising, since transparency toward the patient is the core value in the Open Notes service. The aim of the service is to enhance patients' empowerment over and participation in their own health care by giving them access to their medical information. In the terms of Blomgren and Sundén [41], the concept of transparency is associated with a number of ideals: democracy, accountability, fairness, informed citizenship, and patient rights. According to the definition of transparency by Florini [43], it increases patient access to information to enable the patient to be informed about decisions made by HCPs and take an informed part in them. While we agree with the definitions above, we would like to contribute to specifying the type of direct, process transparency that the material reveals because it differs from most of the examples of transparency illustrated in theory. The fact that implementation (or reform) of Open Notes is policy driven while demanding real-time transparency on behalf of the citizens/patients and not the authorities makes this particular form of transparency quite unique and interesting. We have chosen to call it *governed individual real-time transparency*.

Since patients have been able to obtain paper copies of their health records on request for decades, the Open Notes service is viewed by policy makers and implementers as a simplifier that has made EHRs much more accessible to patients in both time (whenever) and space (from wherever). The service was not intended to affect the work of HCPs to any degree. However, comments in the free-text answers show that the transparency of Open Notes is a much more complex matter than anticipated by the initiators. The fact that EHRs can be accessed by patients as soon as the HCPs push the enter button changes the type of transparency from transparency in retrospect to transparency in real time, according to Heald's anatomy of transparency [46]. This implies that the "accountability window is always open and surveillance is continuous," which does not give the HCP any time to focus on writing the notes without considering that they can be read at any time. This also implies that the transparency that Open Notes allows is what Flyverbom et al [45] describe as observational control that reallocates power between HCPs and patients.

The above described complexity of the transparency of Open Notes may be the main reason for the results from the open-ended questions in the two surveys, which reveal that HCPs experience trade-offs in all seven factors that Heald [47] addresses in his discussion about how enhanced transparency may affect practice. According to HCPs, trade-offs affect their work, their relationship with patients, and not least, their work tool, the EHR. This is in line with Hansen and Flyverbom's [44] thesis that transparency may add to the problem instead of solving it. However, since many of HCPs state that they have not met many, or in some cases any, patients who have read their EHRs, these effects seem to be more connected to the possibility (or threat) of transparency than the actual effectuated transparency. The fact that patients are able to read their EHRs is key to the HCP reactions rather than whether the patients actually do so. The professionals' response to Open Notes by changing the way they write their notes is thus a preventive action to a certain extent.

The professionals' claim that they mainly change their way of writing entries in the EHR system by becoming less candid or leaving out information could in Ball's [42] terms be interpreted as a way of keeping some of the secrecy intact around EHRs. The HCP comments also confirm Ball's findings that secrecy toward the patients is primarily aimed at protecting the patients and their relatives. Many HCPs claim that they produce a less complete account of their work and the patients' medical information with the Open Notes system than they did without it. This means that in some sense, the effects of Open Notes in psychiatry may be that it offers more direct transparency into the EHR but an EHR with less information than before the service. In the terms of Flyverbom et al [45], this can be interpreted as a self-regulatory response to the observational control of Open Notes transparency. HCPs also point out that an EHR that withholds information is less useful to them and their colleagues in their daily work. This confirms the hypothesis by Verheij et al [56] that patients' access to the EHR will affect the quality of information in the EHR and make it less useful to the following HCPs [56]. This is also in line with Heald's [47] question whether full transparency is always the best and whether it would be more rewarding and effective for both HCPs and patients if certain sensitive parts of the EHR had been visible to HCPs only.

The comments in both surveys indicate that HCPs change their writing to minimize what they believe are negative trade-offs of enhanced transparency. These trade-offs, as shown above, include loss of autonomy and control over the EHR; impaired confidentiality, privacy and integrity for HCPs, patients, and relatives; and loss of trust between HCPs and patients and more. The total loss of effectiveness in the work of HCPs, both when writing notes and in the care appointment/therapy, can be seen here as the effect of all other trade-offs put together. Postsurvey comments also indicate that many of patients cannot make use of the service, for reasons such as their mental condition, lack of equipment, or their social situation. The authorities' idea of the rational and capable patient making use of the enhanced transparency offered to them does thus not coincide with the perceptions HCPs have of many of their patients in psychiatry. This indicates that social justice may be hard to reach through governed individual real-time transparency because the possibility that patients can make use of the intended empowerment will always depend on individual situations regarding socioeconomic factors, level of education, interest, and health literacy. This implies a different conceptualization of transparency than the positive connotation it often has had in earlier research.

On the positive side, Open Notes has led to greater understanding and participation for some patients, according to some HCPs. The discourse of increased participation was more salient after implementation than before, indicating that this is an experience rather than a hope. This is in line with the political ambitions of the Open Notes reform and, for these patients, the aim of the reform may well be fulfilled. However, it has not been our intention to decide whether the Open Notes service (reform) has been successful but to broaden the understanding of how enhanced transparency into health care practice can affect HCPs and their work.

Limitations

This study has several limitations. First, we had no way of knowing if the same individuals answered the questions in the baseline survey and the postimplementation survey. Thus, it is only possible to compare the results from the two surveys on a group level; we do not know if and how individual employees have changed their perception about Open Notes. Second, the response rate was 28.86% to the baseline survey and 27.73% to the postimplementation survey. However, the aim of this article was not to generalize the results but rather to present a deeper understanding of the phenomenon by conducting a qualitative content analysis of all 1554 free-text answers from the two surveys. Third, as mentioned above, we have chosen to study perceptions of HCPs when Open Notes was implemented in psychiatric care because this medical specialty has a large percentage of vulnerable patients. It is therefore likely that concerns that in other parts of health care are less visible because the patients are thought to be capable enough to benefit from the Open Notes service are more salient in this context. Further research is needed to determine whether the results from this study are transferable to other medical specialties. Finally, the baseline and postimplementation surveys were designed in different ways, resulting in a different number of open-ended questions in the two surveys. To permit comparisons between expectations before implementation and experiences after, the postimplementation survey was based on the baseline survey [15]. However, in order to be able to capture the experiences of HCPs in the postimplementation survey, fixed-choice questions and open-ended questions were added. This methodological choice resulted in a different number of free-text answers in each survey. This was not considered to be an issue, however, as our aim was to do a qualitative content analysis [54]. We were not interested in counting frequencies or proportions of similar statements but rather in gaining a deeper understanding of the phenomenon.

Conclusion

The effects of Open Notes may well vary between different medical specialties relative to their sensitivity to both total and real-time transparency. Psychiatry has been viewed as a particularly vulnerable area when it comes to sensitivity to this type of transparency. This is confirmed by our results concerning both the content of the EHR and patients' ability to take advantage of the opportunities provided by the service. The results also show that if HCPs react by changing their way of writing notes, becoming less candid in their writing and/or omitting information, Open Notes can affect the efficiency of the work of HCPs and the service itself in a negative way. Additionally, we conclude that HCP reactions are aimed primarily at protecting patients and their relatives as well as their own relationship with the patients and secondly at protecting themselves.

In this paper, we have shown that governed individual real-time transparency that provides full transparency of an actual practice in health care may have the intended positive effects but can also result in negative trade-offs between transparency and the efficiency of the actual practice. This implies that full transparency is not always the most desirable and other options may be considered on the scale between none and full transparency.

Acknowledgments

The research presented in this paper was funded by AFA Insurance in Sweden via the project eHealth Services' Impact on the Working Environment of Health Professionals. We are grateful to Jan Walker and the OpenNotes team at Harvard Medical School for letting us use the OpenNotes survey in our study.

Conflicts of Interest

None declared.

References

1. Blomgren M. The drive for transparency: organizational field transformations in Swedish healthcare. Public Administration 2007 Mar;85(1):67-82. [doi: 10.1111/j.1467-9299.2007.00634.x]
2. Levay C. Health care transparency in organizational perspective. In: Ferlie E, Montgomery K, Pedersen AR, editors. The Oxford Handbook of Health Care Management. Oxford: Oxford University Press; 2016.
3. Kurtzman ET, Jennings BM. Trends in transparency: nursing performance measurement and reporting. J Nurs Adm 2008;38(7-8):349-354. [doi: 10.1097/01.NNA.0000323945.89640.89] [Medline: 18690126]
4. Sahlin-Andersson K. Transparensens former. In: Levay C, Waks C, editors. Strävan efter transparens? Granskning, styrning och organisering i sjukvårdens nätverk. Stockholm: Forlag; 2006:19-50.
5. Blomgren M, Sahlin-Andersson K. Quest for transparency: signs of a new institutional era in the health care field. In: Christensen T, editor. Transcending New Public Management: The Transformation of Public Sector Reform. Aldershot: Ashgate; 2007:155-177.
6. Jaffe R, Nash RA, Ash R, Schwartz N, Corish R, Born T, SAIMP Working Group on Healthcare Transparency. Healthcare transparency: opportunity or mirage. J Mgmt Devel 2006 Dec;25(10):981-995. [doi: 10.1108/02621710610708603]
7. Kajbjer K, Nordberg R, Klein G. Electronic health records in sweden: from administrative management to clinical decision support. In: Impagliazzo J, Lundin P, Wangler B, editors. History of Nordic Computing. Stockholm: Springer; 2011:74-82.
8. Ratwani RM, Hodgkins M, Bates DW. Improving electronic health record usability and safety requires transparency. JAMA 2018 Dec 25;320(24):2533-2534. [doi: 10.1001/jama.2018.14079] [Medline: 30489619]
9. Liebovitz D. Meaningful EHR attributes for an era of accountability, transparency, shared decision making, and value assessment. J Leg Med 2013;34(1):43-53. [doi: 10.1080/01947648.2013.768145] [Medline: 23550982]
10. Funck E. Ordination balanced scorecard: översättning av ett styrinstrument inom hälso- och sjukvården [Dissertation]. Växjö: Växjö University Press; 2009.
11. Levay C, Waks C. Professions and the pursuit of transparency in healthcare: two cases of soft autonomy. Organiz Stud 2009 May 20;30(5):509-527. [doi: 10.1177/0170840609104396]
12. Cresswell K, Sheikh A. Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review. Int J Med Inform 2013 May;82(5):e73-e86. [doi: 10.1016/j.ijmedinf.2012.10.007] [Medline: 23146626]
13. Erlingsdóttir G, Lindholm C. When patient empowerment encounters professional autonomy: the conflict and negotiation process of inscribing an eHealth service. Scand J Pub Admin 2015:27-48 [FREE Full text]
14. Petersson L, Erlingsdóttir G. Open Notes in Swedish psychiatric care (part 1): survey among psychiatric care professionals. JMIR Ment Health 2018 Feb 02;5(1):e11 [FREE Full text] [doi: 10.2196/mental.9140] [Medline: 29396386]
15. Petersson L, Erlingsdóttir G. Open Notes in Swedish psychiatric care (part 2): survey among psychiatric care professionals. JMIR Ment Health 2018 Jun 21;5(2):e10521 [FREE Full text] [doi: 10.2196/10521] [Medline: 29929946]
16. Peck P, Torous J, Shanahan M, Fossa A, Greenberg W. Patient access to electronic psychiatric records: a pilot study. Health Pol Technol 2017 Sep;6(3):309-315. [doi: 10.1016/j.hlpt.2017.06.003]
17. Walker J, Leveille SG, Ngo L, Vodicka E, Darer JD, Dhanireddy S, et al. Inviting patients to read their doctors' notes: patients and doctors look ahead: patient and physician surveys. Ann Intern Med 2011 Dec 20;155(12):811-819 [FREE Full text] [doi: 10.7326/0003-4819-155-12-201112200-00003] [Medline: 22184688]
18. Bell SK, Mejilla R, Anselmo M, Darer JD, Elmore JG, Leveille S, et al. When doctors share visit notes with patients: a study of patient and doctor perceptions of documentation errors, safety opportunities and the patient-doctor relationship. BMJ Qual Saf 2016 May 18;26:262-270. [doi: 10.1136/bmjqs-2015-004697] [Medline: 27193032]
19. Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, et al. Patient experiences with full electronic access to health records and clinical notes through the My HealtheVet Personal Health Record Pilot: qualitative study. J Med Internet Res 2013;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]
20. Denneson LM, Cromer R, Williams HB, Pisciotta M, Dobscha SK. A qualitative analysis of how online access to mental health notes is changing clinician perceptions of power and the therapeutic relationship. J Med Internet Res 2017 Jun 14;19(6):e208 [FREE Full text] [doi: 10.2196/jmir.6915] [Medline: 28615152]
21. Huvila I, Cajander A, Daniels M, Åhlfeldt R. Patients' perceptions of their medical records from different subject positions. J Assoc Inf Sci Tec 2015 Apr 02;66(12):2456-2470. [doi: 10.1002/asi.23343]
22. Grünloh C, Cajander A, Myreteg G. “The record is our work tool”—physicians’ framing of a patient portal in Sweden. J Med Internet Res 2016 Jun 27;18(6):e167 [FREE Full text] [doi: 10.2196/jmir.5705] [Medline: 27349531]
23. Dobscha SK, Denneson LM, Jacobson LE, Williams HB, Cromer R, Woods S. VA mental health clinician experiences and attitudes toward Open Notes. Gen Hosp Psychiatry 2016;38:89-93. [doi: 10.1016/j.genhosppsych.2015.08.001] [Medline: 26380876]

24. Klein JW, Jackson SL, Bell SK, Anselmo MK, Walker J, Delbanco T, et al. Your patient is now reading your note: opportunities, problems, and prospects. Am J Med 2016 Oct;129(10):1018-1021. [doi: 10.1016/j.amjmed.2016.05.015] [Medline: 27288854]
25. Cajander A, Grünloh C. Electronic health records are more than a work tool: conflicting needs of direct and indirect stakeholders. 2019 Presented at: ACM Conference on Human Factors in Computing Systems; 2019; Glasgow URL: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-235847 [doi: 10.1145/3290605.3300865]
26. Scandurra I, Jansson A, Forsberg-Fransson M, Ålander T. Patient accessible EHR is controversial: lack of knowledge and diverse perceptions among professions. Int J Reliab Qual E-Healthc 2017;6(1):29-45. [doi: 10.4018/ijrqeh.2017010103]
27. Leveille SG, Mejilla R, Ngo L, Fossa A, Elmore JG, Darer J, et al. Do patients who access clinical information on patient internet portals have more primary care visits? Med Care 2016 Jan;54(1):17-23. [doi: 10.1097/MLR.0000000000000442] [Medline: 26565525]
28. Lee BS, Walker J, Delbanco T, Elmore JG. Transparent electronic health records and lagging laws. Ann Intern Med 2016 Dec 02;165(3):219-220. [doi: 10.7326/M15-2827] [Medline: 27214266]
29. Delbanco T, Walker J. Benefits from destroying the black box (or are we opening Pandora's box?). 2012. URL: https://www.sgim.org/File%20Library/SGIM/Resource%20Library/Forum/2012/Feb2012-02.pdf [accessed 2019-08-28]
30. Walker J, Kahn MW, Delbanco T. Transparency in the delivery of mental health care—reply. JAMA 2014 Aug 13;312(6):650-651. [doi: 10.1001/jama.2014.7610] [Medline: 25117139]
31. Walker J, Meltsner M, Delbanco T. US experience with doctors and patients sharing clinical notes. BMJ 2015 Feb 10;350:g7785. [Medline: 25670182]
32. Walker J, Darer JD, Elmore JG, Delbanco T. The road toward fully transparent medical records. N Engl J Med 2014 Jan 02;370(1):6-8. [doi: 10.1056/NEJMp1310132] [Medline: 24304001]
33. Kahn MW, Bell SK, Walker J, Delbanco T. A piece of my mind. Let's show patients their mental health records. JAMA 2014 Apr 02;311(13):1291-1292. [doi: 10.1001/jama.2014.1824] [Medline: 24691603]
34. Farrell HM. Transparency in psychiatric care. Asian J Psychiatr 2012 Sep;5(3):273-274. [doi: 10.1016/j.ajp.2012.07.011] [Medline: 22981058]
35. Grünloh C, Myreteg G, Cajander A, Rexhepi H. “Why do they need to check me?” Patient participation through eHealth and the doctor-patient relationship: qualitative study. J Med Internet Res 2018 Jan 15;20(1):e11 [FREE Full text] [doi: 10.2196/jmir.8444] [Medline: 29335237]
36. Ouchi WG. A conceptual framework for the design of organizational control mechanisms. Manag Sci 1979 Sep;25(9):833-848. [doi: 10.1287/mnsc.25.9.833]
37. Freidson E. Professionalism, the Third Logic: On the Practice of Knowledge. Chicago: University Of Chicago Press; 2001.
38. Abbott A. The System of Professions: An Essay on the Division of Expert Labor. London: University Of Chicago Press; 1988.
39. Jonnergård K, Erlingsdóttir G. Variations in professions’ adaption of quality reforms: the cases of doctors and auditors in Sweden. Curr Sociol 2012 Apr 26;60(5):672-689. [doi: 10.1177/0011392112440440]
40. Hood C. Transparency in historical perspective. In: Hood C, Heald D, editors. Transparency: The Key to Better Governance. Oxford: Oxford University Press; 2006:3-23.
41. Blomgren M, Sundén E. Constructing a European healthcare market: the private healthcare company Capio and the strategic aspect of the drive for transparency. Soc Sci Med 2008 Nov;67(10):1512-1520. [doi: 10.1016/j.socscimed.2008.06.035] [Medline: 18703262]
42. Ball C. What is transparency? Pub Integ 2009 Sep 7;11(4):293-308. [doi: 10.2753/PIN1099-9922110400]
43. Florini A. The Right to Know: Transparency for an Open World. New York: Columbia University Press; 2007.
44. Hansen HK, Flyverbom M. The politics of transparency and the calibration of knowledge in the digital age. Organization 2014 Feb 28;22(6):872-889. [doi: 10.1177/1350508414522315]
45. Flyverbom M, Christensen LT, Hansen HK. The transparency–power nexus: observational and regularizing control. Manag Comm Q 2015 Jul 14;29(3):385-410. [doi: 10.1177/0893318915593116]
46. Hood C. Varieties of transparency. In: Hood C, Heald D, editors. Transparency: The Key to Better Governance. Oxford: Oxford University Press; 2006:25-43.
47. Hood C. Transparency as an instrumental value. In: Hood C, Heald D, editors. Transparency: The Key to Better Governance. Oxford: Oxford University Press; 2006:59-73.
48. Delbanco T, Walker J, Darer JD, Elmore JG, Feldman HJ, Leveille SG, et al. Open Notes: doctors and patients signing on. Ann Intern Med 2010 Jul 20;153(2):121-125. [doi: 10.7326/0003-4819-153-2-201007200-00008] [Medline: 20643992]
49. Delbanco T, Walker J, Bell SK, Darer JD, Elmore JG, Farag N, et al. Inviting patients to read their doctors' notes: a quasi-experimental study and a look ahead. Ann Intern Med 2012 Oct 2;157(7):461-470 [FREE Full text] [doi: 10.7326/0003-4819-157-7-201210020-00002] [Medline: 23027317]
50. Walker J, Delbanco T. Interval examination: moving toward Open Notes. J Gen Intern Med 2013 Jul;28(7):965-969 [FREE Full text] [doi: 10.1007/s11606-013-2407-3] [Medline: 23620188]
51. O'Cathain A, Thomas KJ. “Any other comments” Open questions on questionnaires—a bane or a bonus to research? BMC Med Res Methodol 2004 Nov 08;4:25 [FREE Full text] [doi: 10.1186/1471-2288-4-25] [Medline: 15533249]

52. Singer E, Couper M. Some methodological uses of responses to open questions and other verbatim comments in quantitative surveys. Meth Data Analys 2017;11(20):20. [doi: 10.12758/mda.2017.01]
53. Charmaz K. Constructing Grounded Theory: Introducing Qualitative Methods Series. Thousand Oaks: Sage Publications Ltd; 2014.
54. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
55. Graneheim UH, Lindgren B, Lundman B. Methodological challenges in qualitative content analysis: a discussion paper. Nurse Educ Today 2017 Sep;56:29-34. [doi: 10.1016/j.nedt.2017.06.002] [Medline: 28651100]
56. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. J Med Internet Res 2018 May 29;20(5):e185 [FREE Full text] [doi: 10.2196/jmir.9134] [Medline: 29844010]

Abbreviations

eHealth: electronic health
EHR: electronic health record
HCP: health care professional
NPM: New Public Management

Edited by G Eysenbach; submitted 11.04.19; peer-reviewed by Y Chu, R Verheij, K McKay, B Kohrt; comments to author 08.05.19; revised version received 28.06.19; accepted 18.08.19; published 25.09.19.

Please cite as:
Erlingsdóttir G, Petersson L, Jonnergård K
A Theoretical Twist on the Transparency of Open Notes: Qualitative Analysis of Health Care Professionals' Free-Text Answers
J Med Internet Res 2019;21(9):e14347
URL: https://www.jmir.org/2019/9/e14347
doi:10.2196/14347
PMID:31573905

Original Paper

Accuracy of Parent-Reported Child Height and Weight and Calculated Body Mass Index Compared With Objectively Measured Anthropometrics: Secondary Analysis of a Randomized Controlled Trial

Li Kheng Chai[1,2,3], BND (Hons); Clare E Collins[1,2,3], PhD; Chris May[1,4], PhD; Carl Holder[3], MBiostat; Tracy L Burrows[1,2,3], PhD

[1]School of Health Sciences, Faculty of Health and Medicine, The University of Newcastle, Callaghan, Australia
[2]Priority Research Centre in Physical Activity and Nutrition, The University of Newcastle, Callaghan, Australia
[3]Hunter Medical Research Institute, New Lambton, Australia
[4]Family Action Centre, The University of Newcastle, Callaghan, Australia

Corresponding Author:
Tracy L Burrows, PhD
School of Health Sciences
Faculty of Health and Medicine
The University of Newcastle
University Drive
Callaghan,
Australia
Phone: 61 249215514
Email: tracy.burrows@newcastle.edu.au

Abstract

Background: Electronic health (eHealth) interventions for children often rely on parent-reported child anthropometric measures. However, limited studies have assessed parental accuracy in reporting child height and weight via Web-based approaches.

Objective: The objective of this study was to determine the accuracy of parent-reported child height and weight, as well as body mass index and weight category that we calculated from these data. We also aimed to explore whether parent report was influenced by age, sex, weight status, or exposure to participation in a 12-week brief Web-based family lifestyle intervention.

Methods: This study was a secondary analysis of data from a 12-week childhood obesity pilot randomized controlled trial in families with children aged 4 to 11 years in Australia. We asked parents to report demographic information, including child height and weight, using an online survey before their child's height and weight were objectively measured by a trained research assistant at baseline and week 12. We analyzed data using the Lin concordance correlation coefficient (ρc, ranging from 0 [poor] to ±1 [perfect] concordance), Cohen kappa coefficient, and multivariable linear regression models.

Results: There were 42 families at baseline and 35 families (83%) at week 12. Overall, the accuracy of parent-reported child height was moderate (ρc=.94), accuracy of weight was substantial (ρc=.96), and accuracy of calculated body mass index was poor (ρc=.63). Parents underreported child height and weight, respectively, by 0.9 cm and 0.5 kg at baseline and by 0.2 cm and 1.6 kg after participating in a 12-week brief Web-based family lifestyle intervention. The overall interrater agreement of child body mass index category was moderate at baseline (κ=.59) and week 12 (κ=.54). The weight category calculated from 74% (n=31) and 70% (n=23) of parent-reported child height and weight was accurate at baseline and week 12, respectively. Parental age was significantly (95% CI –0.52 to –0.06; P=.01) associated with accuracy of reporting child height. Child age was significantly (95% CI –2.34 to –0.06; P=.04) associated with reporting of child weight.

Conclusions: Most Australian parents were reasonably accurate in reporting child height and weight among a group of children aged 4 to 11 years. The weight category of most of the children when calculated from parent-reported data was in agreement with the objectively measured data despite the body mass index calculated from parent-reported data having poor concordance at both time points. Online parent-reported child height and weight may be a valid method of collecting child anthropometric data ahead of participation in a Web-based program. Future studies with larger sample sizes and repeated measures over time in the context of eHealth research are warranted. Future studies should consider modeling the impact of calibration equations applied to parent-reported anthropometric data on study outcomes.

(J Med Internet Res 2019;21(9):e12532) doi:10.2196/12532

KEYWORDS

telemedicine; parents; child; body height; body weight; body mass index; self report; dimensional measurement accuracy

Introduction

Background

The wide coverage of the internet and the increase in technology use worldwide have led to the emergence of electronic health (eHealth) for lifestyle interventions [1,2]. Web-based platforms are increasing in popularity and used for data collection and delivery of eHealth interventions [1,3]. In Australia, technology use is increasing and not limited by socioeconomic status or location, with more than 97% of households with children under age 15 years having access to the internet via computer, smartphone, or tablet [4]. Research suggests that online data collection and delivery of interventions is more cost effective than conventional face-to-face modes [5], allows providers to connect with a large number of people simultaneously, and enhances access to services for communities living in rural and remote locations [6].

One limitation, however, is that eHealth interventions and non–face-to-face programs are usually delivered over distance. Hence, interventionists have had to rely on self-reported measures (eg, anthropometrics, diet, physical activity) when objective measurements were not possible. For young children in eHealth lifestyle interventions, these measures often include parent-reported child height and weight data. The risk associated with self-reported height and weight data is that discrepancy with other objective measures can result in miscalculation of weight trajectories and weight category. Parental underestimation of their child's weight has important clinical implications due to the health consequences of childhood obesity [7] and the importance of early identification of a weight gain trajectory in order to seek early intervention. Misreporting may also influence a child's actual eligibility for research or treatment programs that recruit participants using self-reported screening surveys.

Studies have used face-to-face interviews or surveys completed at home visits or during clinic visits to collect parent-reported child data. However, data collection may differ between remote non–person-to-person methods (eg, Web-based, posted paper surveys) and direct person-to-person methods (eg, home visits, clinic visits, telephone interviews). Very few studies have evaluated parental accuracy in reporting the height and weight of their children remotely without the presence of clinicians or researchers (ie, online surveys). Furthermore, most of the previous research on parental reporting of their child's height and weight was conducted in Canada, Western Europe, or the United States [8-12], and no studies have included Australian children. Therefore, it is unknown whether Australian parents would perceive their children's height and weight in similar ways to parents in other countries. Moreover, previous studies have used limited measures to assess agreement, such as Pearson correlation coefficients or paired *t* tests. These measures were unable to adequately detect levels of agreements (ie, accuracy and precision) and, rather, they are associations between parent-reported and researcher- or clinician-measured anthropometric data only [13].

Objective

We aimed to determine (1) the accuracy of parent-reported child height and weight, as well as the body mass index (BMI) and weight category that we calculated from these data, compared with data measured objectively by researchers as the reference standard, and (2) whether parent report was influenced by age, sex, weight category, or exposure to participation in a 12-week brief Web-based family lifestyle intervention.

Methods

Data Source

This was a secondary analysis of data from a pilot randomized controlled trial that aimed to investigate the feasibility of a 12-week Web-based family lifestyle intervention to support parents in improving their child's weight status and eating habit [14]. The intervention group received 2 semistructured telehealth consultations (online video consultations, attended by at least one parent and the index child) with Accredited Practising Dietitian, access to the Back2basics Family program website [15], a Facebook group, and additional evidence-based short message service (SMS) text messaging targeted to the parents. The waitlist control group received all the intervention components at 3-month follow-up from baseline.The pilot trial received ethics approval from the Hunter New England Human Research Ethics Committee (16/07/20/4.04), New Lambton, Australia, and University of Newcastle Human Research Ethics Committee (H-2016-0329), Callaghan, Australia. The trial is reported in accordance with the Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth (CONSORT-EHEALTH) checklist (Multimedia Appendix 1 [16]).

Participants

We recruited families from New South Wales, Australia through clinician referrals, school newsletters, flyers, and word-of-mouth. Eligible families were those who had a child aged 4 to 11 years with a BMI of 21.5 kg/m^2 or greater (International Obesity Task Force [IOTF] child cutoffs) [17], had access to the internet, and were able to attend laboratory measurement sessions at 1 of the 3 study sites (Newcastle, Tamworth, and Armidale). Parents' written consents and children's assents were procured prior to the baseline laboratory measurement session.

Data Collection

We asked parents to report demographic information of themselves and an identified index child from the family participating in the intervention. Demographic information included age, sex, height, weight, highest education attained (parent only), and postcode (parent only). Parents provided these details using an online survey before their child's anthropometric measurements were clinically assessed at baseline and week 12, which is a 3-month follow-up from baseline.

We provided no specific instructions to the parents when asked about reporting their child's height and weight (eg, use a tape or scale or time of day to measure). The questions were "What is your child's height in cm (if unsure, please estimate)" and "What is your child's weight in kg (if unsure, please estimate)". Subsequently, the child's height and weight were measured at baseline and week 12 using standard protocols by blinded research assistants.

In the laboratory sessions, child height was measured to the nearest 0.1 cm while the child was standing with their head and chin up, looking straight ahead (ie, held in the Frankfurt plane) using the BSM370 Automatic Stadiometer (Biospace Co Ltd, Seoul, South Korea). Child weight was measured to the nearest 0.1 kg without shoes and in light clothing using the 720 body composition analyzer (InBody Co, Ltd, Seoul, South Korea). Measurements were taken twice, and the difference between measures was required to be 0.3 cm or less (height) and 0.4 kg or less (weight). Otherwise, a third reading was obtained, and the 2 closest readings were used to compute an average height or weight measurement. Families were offered an Aus $10 gift voucher for participation in each laboratory measurement session. Analyses were conducted by a researcher not involved in the laboratory measurements.

Statistical Analysis

All data manipulation and statistical analyses were undertaken using Stata version 12 (StataCorp LLC). We calculated descriptive statistics for baseline participant characteristics. We used parent-reported and researcher-measured child height and weight data to calculate child BMI, with weight category based on IOTF child cutoffs [17]. We used the Lin concordance correlation coefficient (CCC) to assess the level of agreement of parent-reported child height and weight and the calculated BMI compared with researcher-measured data at baseline and week 12. We chose this as a superior method because it measures both precision (ie, Pearson correlation coefficient) and accuracy, thus indicating how well a set of bivariate data compares with the reference standard, measured data. Lin CCC (ρc, ranging from 0 to ±1) is interpreted as almost perfect agreement (ρc>.99), substantial agreement (ρc range .951-.99), moderate agreement (ρc range .90-.950), and poor agreement (ρc<.90) [18]. We used Cohen kappa coefficient to ascertain interrater agreement between the child weight category calculated from parent-reported child height and weight and the researcher-measured data [19]. Cohen kappa coefficient (κ, ranging from 0 to 1) is interpreted as almost perfect (κ>.80), substantial (.61≤κ≤.80), moderate (.41≤κ≤.60), fair (.21≤κ≤.40), slight (.00≤κ≤.20), and poor (κ<.00) [20]. We used multivariable linear regression models for the outcomes of difference between parent-reported and researcher-measured height and weight (calculated by subtracting researcher-measured data from parent-reported data) to further investigate relationships between outcomes and age (years), sex, and weight status of parents and the index child. We calculated a sample size of 7 and 29 participants per group as needed to detect expected correlation coefficients of .9 and .5, respectively, at an alpha of .05 and with 80% power for height and weight.

Results

Participant Characteristics

Parent-reported and researcher-measured child height and weight data were available from 42 families at baseline and 35 families (83%) at week 12. Baseline characteristics of children and parents who were lost to follow-up, defined as not responding after 3 reminders to complete assessments, were not significantly different from the baseline characteristics of those who completed the follow-up at week 12 [14].

Most parents were female (40/42, 95%) with a mean age 40.5 years and mean BMI 29.9 kg/m^2, were from middle socioeconomic background (n=28, 67%), and attained a certificate or diploma level of education (n=13, 31%) or a university degree (n=11, 26%). Parents were classified into the obese (n=19, 46%), overweight (n=14, 33%), and healthy weight (n=9, 21%) categories based on the IOTF adult cutoffs [17]. Children were fairly evenly represented by sex (n=24 male, 57%) with a mean age 8.5 years, mean BMI 22.9 kg/m^2, and weight category of obese (n=22, 52%), overweight (n=9, 21%), and healthy weight (n=11, 26%) based on the IOTF child cutoffs [17]. Table 1 presents detailed participant characteristics.

Table 1. Baseline characteristics of parents and their children.

Characteristics	Intervention group (n=28)	Control group (n=14)	Combined (n=42)
Parents			
Age (years), mean (SD)	41 (7)	39 (8)	41 (7)
Female sex, n (%)	26 (93)	14 (100)	40 (95)
BMI[a] (self-reported), mean (SD), kg/m^2	28.8 (5.2)	32.0 (7.8)	29.9 (6.3)
Weight category[b] (self-reported), n (%)			
Healthy weight	6 (21)	3 (21)	9 (21)
Overweight	11 (39)	3 (21)	14 (33)
Obese	11 (39)	8 (57)	19 (46)
Education level, n (%)			
School certificate	1 (4)	2 (14)	3 (7)
Higher school certificate	4 (14)	2 (14)	6 (14)
Certificate or diploma	11 (39)	2 (14)	13 (31)
Undergraduate degree	7 (25)	4 (29)	11 (26)
Postgraduate degree	5 (18)	4 (29)	9 (21)
Socioeconomic status, n (%)			
Low (IRSAD[c] 1-3)	4 (14)	2 (14)	6 (14)
Mid (IRSAD 4-7)	17 (61)	11 (79)	28 (67)
High (IRSAD 8-10)	7 (25)	1 (7)	8 (19)
Children			
Age (years), mean (SD)	9 (2)	9 (2)	9 (2)
Female sex, n (%)	13 (46)	5 (36)	18 (43)
Anthropometry, mean (SD)			
Height (measured), cm	138 (16)	135 (16)	137 (16)
Weight (measured), kg	44 (17)	46 (19)	45 (17)
BMI (measured), kg/m^2	22.4 (4.7)	23.8 (5.9)	22.9 (5.1)

[a] BMI: body mass index.

[b] Weight categories as per International Obesity Task Force age-appropriate cutoffs: healthy weight, BMI 18.5-24.9 kg/m^2; overweight, BMI 25-29.9 kg/m^2; obese, BMI ≥30 kg/m^2.

[c] IRSAD: Index of Relative Socio-Economic Advantage and Disadvantage.

Agreement Between Parent-Reported and Researcher-Measured Data

Table 2 summarizes the level of agreement between the parent-reported child height and weight and calculated BMI and the researcher-measured data. At baseline, the level of agreement between parent-reported and researcher-measured data as determined by Lin CCC was moderate (ρc=.94) for parent-reported child height and substantial (ρc=.96) for weight, and poor (ρc=.63) for calculated BMI. In this study, parents tended to underreport their child's height and weight, with a mean (SD) difference of –0.9 (SD 6.0) cm and –0.5 (SD 4.9) kg, respectively, compared with researcher-measured data. Parents were more accurate in reporting the height of children who were taller than 140 cm (Figure 1) and were overall better reporters of the weight of children who weighed between 30 and 50 kg (Figure 2). BMI calculated using parent-reported data was higher than researcher-measured data with a mean (SD) difference of 0.7 (SD 4.7) kg/m^2. Figure 3 demonstrates that child BMI calculated from parent-reported data was more accurate (ie, closer to researcher-measured data) for children whose BMI was between 15 and 25 kg/m^2 compared with those whose BMI was at either end of the spectrum (ie, <15 kg/m^2 or >25 kg/m^2).

Table 2. Level of agreement between parent-reported and researcher-measured child height, weight, and calculated body mass index (BMI).

Time point	Intervention group			Control group			Combined		
	MD[a] (SD)	ρc[b]	95% CI	MD (SD)	ρc	95% CI	MD (SD)	ρc	95% CI
Baseline	**n=28**			**n=14**			**n=42**		
Height (cm)	–1.9 (5.7)	.94	0.90-0.98	1.0 (6.3)	.94	0.87-0.99	–0.9 (6.0)	.94	0.91-0.97
Weight (kg)	–1.1 (5.3)	.94	0.90-0.98	0.7 (4.0)	.98	0.95-1.00	–0.5 (4.9)	.96	0.93-0.98
BMI (kg/m^2)	0.7 (4.9)	.53	0.26-0.80	0.7 (4.2)	.76	0.52-1.00	0.7 (4.7)	.63	0.45-0.81
Week 12	**n=21**			**n=14**			**n=35**		
Height (cm)	–0.2 (6.4)	.93	0.87-0.99	–3.9 (8.8)	.86	0.74-0.99	–1.7 (7.5)	.90	0.84-0.96
Weight[c] (kg)	–1.6 (3.3)	.98	0.95-1.00	–2.9 (2.6)	.98	0.96-1.00	–2.1 (3.0)	.98	0.96-0.99
BMI[c] (kg/m^2)	–0.9 (2.3)	.85	0.73-0.98	0.2 (3.6)	.84	0.68-1.01	–0.5 (2.9)	.86	0.76-0.95

[a] MD: mean difference (parent-reported value subtract researcher-measured value).

[b] ρc: Lin concordance correlation coefficient, ranging from 0 to ±1, where a value close to 1.0 (and a 45° fitted line of perfect concordance) suggests a perfect level of agreement, .951-.99 is substantial agreement, .90-.950 is moderate agreement, and <.90 is poor agreement).

[c] Data were available from n=19 intervention families due to missing parent-reported weight.

Figure 1. Concordance between parent-reported and researcher-measured height at baseline.

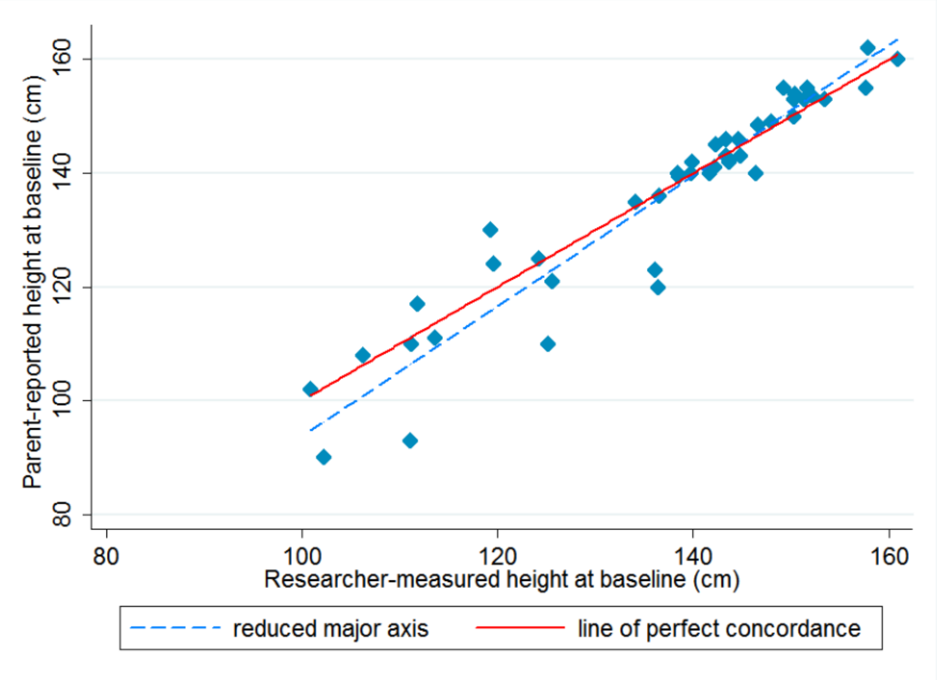


Figure 2. Concordance between parent-reported and researcher-measured weight at baseline.

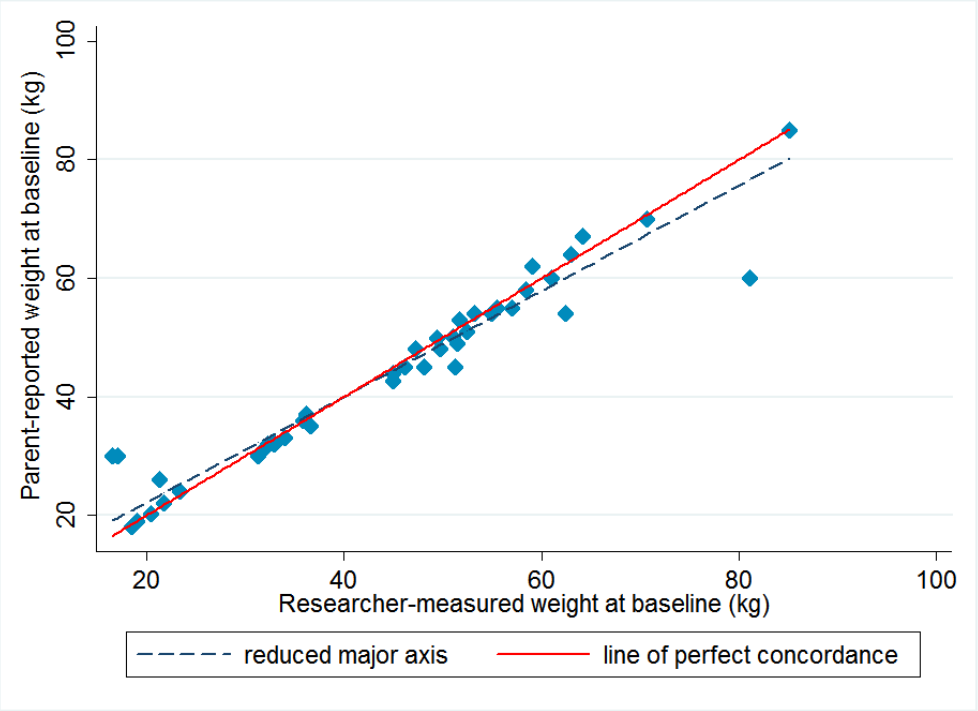
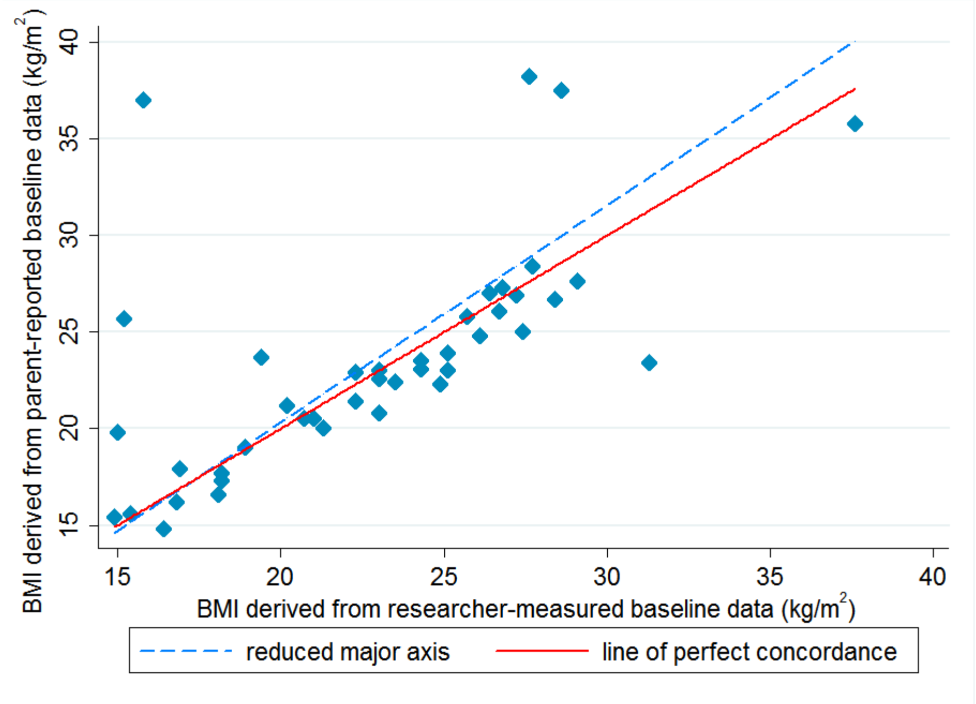


Figure 3. Concordance between body mass index (BMI) calculated from parent-reported data and from researcher-measured data at baseline.



At week 12, the level of agreement between parent-reported and researcher-measured child height remained moderate (ρc=.93) in the intervention group but declined from moderate to poor agreement (ρc=.86) in the control group. Parents' accuracy in reporting their child's weight improved in the intervention group, demonstrating substantial agreement (ρc=.98) with researcher-measured data, whereas control group parents remained at the same level of agreement as at baseline (ρc=.98). At week 12, parents in the intervention group continued to underreport their child's height and weight with a mean difference (SD) of –0.2 cm (SD 6.4) and –1.6 kg (SD 3.3), respectively, compared with researcher-measured data. However, these parents demonstrated improved accuracy in reporting their child's height (–0.9 cm to –0.2 cm), but not in reporting their child's weight (–0.5 kg to –1.6 kg). Parents in the control group overreported their child's height and weight at baseline by 1 cm (SD 6.3) and 0.7 kg (SD 4.0), respectively. However, at week 12 the control group parents also underreported their child's height and weight and to a greater extent than the intervention group, and –3.9 cm (SD 8.8) and –2.9 kg (2.6), respectively, lower than researcher-measured data. At week 12, the level of agreement for calculated BMI using parent-reported and researcher-measured data improved but remained poor for both the intervention group (ρc=.85 vs ρc=.53) and the control group (ρc=.84 vs ρc=.76). However, the mean differences of parent-reported and researcher-measured height, weight, and BMI calculated from these data at baseline and week 12 were not statistically significant between the intervention and control groups.

Interrater Agreement

Table 3 summarizes interrater agreement of child weight category based on calculated BMI using parent-reported and researcher-measured child height and weight. At baseline, the level of agreement was moderate (κ=.59). Overall, the weight category (ie, healthy weight, overweight, obese) calculated from 74% (31/42) of parent-reported child height and weight at baseline was accurate (ie, within the same category calculated based on objectively measured data). Of the 11 children in the healthy weight category, the weight category calculated from 55% (6/11) of parent-reported data was accurate, whereas the weight category calculated from 45% (5/11) of parent-reported data misclassified their child in the overweight (3/11, 27%) or obese (2/11, 18%) category. Of the 9 children in the overweight category, the weight category calculated from 89% (8/9) of parent-reported data was accurate, whereas 11% (1/9) of parent-reported data misclassified children who were overweight into the healthy weight category. Among the 22 children in the obese category, the weight category calculated from 77% (17/22) of parent-reported data was accurate, whereas 23% (5/22) of parent-reported data misclassified children as overweight. At week 12, the interrater agreement for child weight category decreased in both the intervention (71% vs 68%) and control groups (79% vs 71%). The level of agreement remained moderate in the intervention (κ=.54) and control groups (κ=.51), as well as for both groups combined (κ=.54). Among the children in the healthy weight category, the weight category calculated from 100% (5/5) of parent-reported data in the intervention group (vs 2/3, 67% in the control group) was accurate at week 12. Among children who were overweight, the weight category calculated from 67% (4/6 and 2/3) of parent-reported data in the intervention and control groups was accurate at week 12, respectively. Among children who were obese, the weight category calculated from 50% (4/8) of parent-reported data in the intervention group (vs 75% (6/8) in the control group) was accurate.

In all families except 1, the same parent reported child height and weight at both time points. Overall, the weight category calculated from 55% (23/42) of parent-reported data was accurate (ie, closer to researcher-measured data) at both the baseline and week 12 time points. The weight category calculated from a small number of parent-reported data was consistently 1 category under (3/42, 7%) or above (1/42, 2%) their child’s correct weight category at both time points. Further analysis did not find statistically significant differences in demographic characteristics (eg, age and sex of parent and child, parental BMI, education, and socioeconomic status) between underreporters and overreporters. A multivariable linear regression model identified that parental age was the only variable that had a significant association (*P*=.01) with accuracy of reporting child height. Every 1-unit (in years) increase in the parent’s age resulted in an underreporting of 0.29 cm (95% CI –0.52 to –0.06; *P*=.01) for child height. Child age was the only variable that was significantly associated with reporting of child weight. Every 1-unit (in years) increase in the child’s age resulted in an underreporting of 1.2 kg (95% CI –2.34 to –0.06; *P*=.04) for child weight.

Table 3. Interrater agreement for child weight category[a] calculated using parent-reported versus researcher-measured child height and weight.

Time point and parent-reported data	Calculated from researcher-measured data							
	Intervention group[b]				Control group[c]			
	Healthy weight	Overweight	Obese	Total	Healthy weight	Overweight	Obese	Total
Baseline, n (%)	**n=28**				**n=14**			
Healthy weight	4 (14)	1 (4)	0 (0)	5 (18)	2 (14)	0 (0)	0 (0)	2 (14)
Overweight	3 (11)	6 (21)	3 (11)	12 (43)	0 (0)	2 (14)	2 (14)	4 (29)
Obese	1 (4)	0 (0)	10 (36)	11 (39)	1 (7)	0 (0)	7 (50)	8 (57)
Total	8 (29)	7 (25)	13 (46)	28 (100)	3 (21)	2 (14)	9 (64)	14 (100)
Week 12, n (%)	**n=19[d]**				**n=14**			
Healthy weight	5 (26)	2 (21)	0 (0)	7 (37)	2 (14)	0 (0)	0 (0)	2 (14)
Overweight	0 (0)	4 (21)	4 (21)	8 (42)	0 (0)	2 (14)	2 (14)	4 (29)
Obese	0 (0)	0 (0)	4 (21)	4 (21)	1 (7)	1 (7)	6 (43)	8 (57)
Total	5 (26)	6 (32)	8 (42)	19 (100)	3 (21)	3 (21)	8 (57)	14 (100)

[a] Weight categories as per International Obesity Task Force age-appropriate cutoffs: healthy weight, body mass index (BMI) 18.5-24.9 kg/m^2; overweight, BMI 25-29.9 kg/m^2; obese, BMI ≥30 kg/m^2.

[b] Baseline: Cohen kappa coefficient= .57, SE 0.13, and % agreement=71; week 12: Cohen kappa coefficient=.54, SE 0.15, and % agreement=68 (Cohen kappa coefficient ranges from 0 to 1, where κ>.80 is almost perfect, .61≤κ≤.80 is substantial, .41≤κ≤.60 is moderate, .21≤κ≤.40 is fair, .00≤κ≤.20 is slight, and κ<.00 is poor).

[c] Baseline: Cohen kappa coefficient=.62, SE 0.19, and % agreement=79; week 12: Cohen kappa coefficient=.51, SE 0.19, and % agreement=71.

[d] Data were available from 19 intervention families due to missing parent-reported weight.

Discussion

Principal Findings

We evaluated the accuracy of online parent-reported child height and weight, as well as BMI and weight category that we calculated from these data, compared with researcher-measured data in a sample of Australian children aged 4 to 11 years. The study also examined whether accuracy of parental-reporting was influenced by age, sex, BMI, and participation in a 12-week brief Web-based family lifestyle intervention.

Key findings indicated that parents were relatively accurate in reporting their child's height and weight as shown by the overall high CCCs (ρc≥.9). Results indicated that parents in this study underreported their child's height and weight at both baseline and after participating in a 12-week Web-based family lifestyle intervention. These findings were similar to previous studies that indicated that parents tended to underreport the height and weight of American children aged 6 to 12 years [21]. Previous studies found that parents underreported their child's height and weight by –1.4 cm and –2.3 kg, respectively (n=475 American children aged 11-12 years) [22], and by –1 cm and –1.6 kg, respectively (n=116 Belgian children aged 7-9 years) [23]. This compares with a study in 662 children in the United States, which found that 35% of parents underreported their child's height by at least 1 inch (2.54 cm) and 26% by at least 2 inches (5.08 cm) [24]. In that study, 22% of parents of children aged 3 to 5 years (n=343) and 39% of parents of children aged 6 to 12 years (n=452) underestimated their child's weight by at least 2 lbs (0.9 kg) [24]. It is evident across previous research that parents' inaccuracy in reporting their child's height and weight, though varying in extent and by country, was commonly due to underreporting instead of overreporting, regardless of the measurement systems used (metric vs imperial system).

Previous studies have highlighted that parents were more likely to underreport their child's height than their weight [21,24]. Our study arrived at similar findings, where parents underreported their child's height and weight, and were generally less accurate in reporting their child's height (ρc range .86-.94) than their weight (ρc range .94-.98) as demonstrated by the consistently higher CCC for child weight over time. This suggests that children may be weighed more regularly or accurately than they are measured for height. It is possible that parents measured their child's weight at home using a weighing scale, which is a common household item. Furthermore, enrolling in a weight management program may make them more aware of their child's weight than their height. In contrast, child height may not be measured as regularly due to not having a stadiometer at home, or as accurately due to not using the Frankfurt plane position, which is the standard protocol for measuring height. However, there are discrepancies between this study and a previous study that found poor concordance between parent-reported and researcher-measured child height (ρc=.007), weight (ρc=–.039), and BMI (ρc=–.005) [13]. It was suggested that a sample of parents in California, USA may have been inclined to report child height in whole inches, resulting in a greater degree of underreporting, or overreporting by 2.54 cm [13]. Using the smaller increments of the metric system may therefore enhance parents' accuracy and precision in reporting their child's height in centimeters, which is a smaller unit [13]. Due to the differences between the metric and imperial systems, study findings in US populations may not be generalizable to those of countries using the metric measurement system [21].

Despite our finding of a consistent trend in underreporting over time, parents who completed a 12-week brief Web-based family lifestyle intervention demonstrated improved accuracy in reporting their child's weight (ρc increased from .94 to .98) across time points in the study, whereas the control group maintained their high accuracy from baseline to 12 weeks (ρc remained at .98). Parents may have become more attentive to child weight information received at clinic or intervention visits or from other sources, or may have recorded height and weight measures at home more regularly after participating in the baseline survey and the intervention program. Studies also suggest that parental accuracy in reporting their child's height and weight may be influenced by whether the parents know that their child's height and weight will be measured by treating clinicians at a later time [8-10,25], and whether parents were asked to self-measure their child's height and weight before reporting [26]. Hence, suggesting to parents that their child's measurements will be validated or providing instructions to parents to measure their child's height and weight themselves may improve accuracy in parental reporting [13].

The underreported child height and weight in this study resulted in poor concordance (ρc=.86) between BMI calculated from parent-reported and BMI calculated from researcher-measured data. Overall, BMI was underestimated by 0.5 kg/m^2 when calculated from parent-reported data. Similar findings were reported in other studies in which BMI calculated by researchers from parent-reported child height and weight data was 0.5 kg/m^2 lower than BMI calculated from objective measures (n=475 children aged 12-13 years) [22]. In another study, BMI was 0.6 kg/m^2 lower than the BMI calculated from researcher-measured data (n=116 children aged 7 to 9 years) [23]. In this study, we found that child BMI calculated from parent-reported data was more accurate for children whose BMI was between 15 and 25 kg/m^2 than for those whose BMI was at either end of the spectrum. Similar findings were reported in a study (n=864 Dutch children aged 4 years) in which parents tended to misreport the weight of their children in the lowest and highest BMI quartiles, and the authors suggested that the turning point for overreporting and underreporting of child BMI appeared to be around 15.4 kg/m^2 [11]. The significance of a misreported BMI depends on whether the BMI is close to the lower or upper range of a weight category. For example, for a 9-year-old boy whose measured BMI is 19 kg/m^2 (overweight category), an underreport by 0.5 kg/m^2 would result in a reported BMI of 18.5 kg/m^2, incorrectly placing the child into the healthy weight category. Studies indicated that misclassification of children as obese based on parent-reported data was associated with underreporting of child height [12,24], as the misreporting was magnified through the BMI calculation formula (ie, weight in kilograms divided by height in meters squared). For this reason, the use of a height or weight percentile might be useful in future studies when interpreting parent-reported child height and weight, instead of calculating BMI to determine child weight category. Future studies may consider modeling a calibration equation for adjusting BMI calculated from parent-reported data to improve accuracy.

Parental underreporting of child height and weight resulted in underestimation of child BMI and misclassification of weight category among 30% (10/33) of children in our study. Overall, child weight category calculated using parent-reported child height and weight at baseline was accurate for 74% (31/42) of families, and this was reduced to 70% (23/33) at week 12. Similar findings were reported in 2 other studies in which child weight category was calculated by researchers using parent-reported child height and weight data, with the BMI category being accurate for 80% (n=558) [24] and 76% (n=600 Austrian children aged 0-15 years) [27]. Among the overweight children in our study, the overall proportion of parents who underestimated their child's weight category ranged from 11% (1/9) to 22% (2/9) over time. Among children in the obese category, the overall proportion of parents who underestimated their child's weight category ranged from 23% (5/22) to 38% (6/16) over time. Previous studies regularly reported misclassifications of overweight and obese children [11,27]. One study reported that 46% of 116 overweight children were misclassified as healthy weight when parent-reported data were used [11]. In another study of 600 children aged 0 to 15 years, parents reported that 37% of obese children were incorrectly classified in the overweight category [27]. Such misclassifications, if not addressed and corrected, or accounted for in interpretation, could have an impact on obesity prevalence statistics or intervention programs calculated using parent-reported child height and weight data.

Strengths and Limitations

A particular limitation of this study was the small sample size, which impeded the modeling of calibration equations to improve the validity of parent-reported data. A large sample would be needed to generate a viable predictive model. Parents tend to be less accurate in reporting for children with excess body weight. Therefore, the study sample comprised children with a BMI above the midpoint of the healthy weight category (≥21.5 kg/m^2) in order to assess parents' accuracy in reporting the height and weight of children with a higher weight. This means that our results may not be generalizable or applicable to other populations and ethnicities and, hence, results should be interpreted with caution. The study, although not population based, is to our knowledge the first Australian study to assess parental accuracy in online reporting of their child's height and weight, and weight status determined by BMI calculated using parent-reported data compared with objective researcher-measured data in a sample of children aged 4 to 11 years. Given that no specific instructions about how to take height and weight measures were provided to parents, a limitation is that parents may or may not have measured their child before reporting the measures. Future studies should explore whether parents' accuracy in reporting their child's anthropometrics improves when specific guidance [28] on when and how to perform the measurements is provided. However, it could be challenging to assess parents' adherence to the specific guidance. Moreover, parents may be less likely to measure a child who is sensitive about weight and body image.

Future studies should collect information on whether parent-reported data were based on estimates or measurements, and whether the measurements were done at home, school, or a clinic, to further the understanding of parents' accuracy in measuring or estimating child height and weight.

A key strength of our study was the use of Lin CCC to assess the level of agreement between parent-reported and researcher-measured data and, hence, offer some confidence in the findings. Many previous studies measured correlations between data, which is insufficient in assessing levels of agreements (ie, accuracy and precision). For example, Pearson correlation coefficients only measure the extent to which the parent-reported data conform to the best-fitting straight line, but not how close or far the data fall from the line that represents perfect agreement [20].

Conclusions

We found that Australian parents of children aged 4 to 11 years were reasonably accurate in reporting their child's height and weight online. The weight category for the majority of children calculated using parent-reported data was in agreement with the objectively measured data despite the BMI calculated from parent-reported data having poor concordance at both time points. It appears that online parent-reported child height and weight may be a valid method of collecting child anthropometric data ahead of participation in a Web-based health, diet, and lifestyle program. Future studies with larger sample sizes and repeated measures over time in the context of eHealth research are warranted.

Acknowledgments

The pilot study received funding from nib foundation through Hunter Medical Research Institute. The funding body was not involved in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript. LKC is supported by the University of Newcastle International Postgraduate Research Scholarships, Barker PhD Award Top-up Scholarship, and Emlyn and Jennie Thomas Postgraduate Medical Research Scholarship through the Hunter Medical Research Institute. CEC is supported by an NHMRC Senior Research Fellowship and a Faculty of Health and Medicine, Gladys M Brawn Senior Research Fellowship, the University of Newcastle. TLB is supported by a Faculty of Health and Medicine, Early Career Brawn Fellowship, the University of Newcastle.

Conflicts of Interest

None declared.

The randomized study analyzed in this paper was not registered, explained by authors because it is a pilot study as advised by the research ethics committee. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2MB - jmir_v21i9e12532_app1.pdf]

References

1. Hammersley ML, Jones RA, Okely AD. Parent-focused childhood and adolescent overweight and obesity ehealth interventions: a systematic review and meta-analysis. J Med Internet Res 2016 Jul 21;18(7):e203 [FREE Full text] [doi: 10.2196/jmir.5893] [Medline: 27443862]
2. Nguyen B, Kornman KP, Baur LA. A review of electronic interventions for prevention and treatment of overweight and obesity in young people. Obes Rev 2011 May;12(5):e298-e314. [doi: 10.1111/j.1467-789X.2010.00830.x] [Medline: 21348921]
3. NSW Ministry of Health. NSW Health Telehealth Framework and Implementation Strategy. North Sydney, Australia: NSW Ministry of Health; 2016.
4. Australian Bureau of Statistics. 8146.0 - Household Use of Information Technology, Australia, 2016-17. Canberra, Australia: Australian Bureau of Statistics; 2018 Mar 28. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/8146.0 [accessed 2018-10-22] [WebCite Cache ID 73LpI27Nc]
5. Harvey-Berino J, Krukowski RA, Buzzell P, Ogden D, Skelly J, West DS. The accuracy of weight reported in a web-based obesity treatment program. Telemed J E Health 2011 Nov;17(9):696-699 [FREE Full text] [doi: 10.1089/tmj.2011.0032] [Medline: 21882997]
6. Mehring M, Haag M, Linde K, Wagenpfeil S, Frensch F, Blome J, et al. Effects of a general practice guided web-based weight reduction program--results of a cluster-randomized controlled trial. BMC Fam Pract 2013;14:76 [FREE Full text] [doi: 10.1186/1471-2296-14-76] [Medline: 23981507]

7. Reilly JJ, Kelly J. Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review. Int J Obes (Lond) 2011 Jul;35(7):891-898. [doi: 10.1038/ijo.2010.222] [Medline: 20975725]
8. Dubois L, Girad M. Accuracy of maternal reports of pre-schoolers' weights and heights as estimates of BMI values. Int J Epidemiol 2007 Feb;36(1):132-138. [doi: 10.1093/ije/dyl281] [Medline: 17510077]
9. Garcia-Marcos L, Valverde-Molina J, Sanchez-Solis M, Soriano-Pérez MJ, Baeza-Alcaraz A, Martinez-Torres A, et al. Validity of parent-reported height and weight for defining obesity among asthmatic and nonasthmatic schoolchildren. Int Arch Allergy Immunol 2006;139(2):139-145. [doi: 10.1159/000090389] [Medline: 16374023]
10. Huybrechts I, De Bacquer D, Van Trimpont I, De Backer G, De Henauw S. Validity of parentally reported weight and height for preschool-aged children in Belgium and its impact on classification into body mass index categories. Pediatrics 2006 Nov;118(5):2109-2118. [doi: 10.1542/peds.2006-0961] [Medline: 17079585]
11. Scholtens S, Brunekreef B, Visscher TLS, Smit HA, Kerkhof M, de Jongste JC, et al. Reported versus measured body weight and height of 4-year-old children and the prevalence of overweight. Eur J Public Health 2007 Aug;17(4):369-374. [doi: 10.1093/eurpub/ckl253] [Medline: 17124194]
12. Shields M, Connor Gorber S, Janssen I, Tremblay MS. Obesity estimates for children based on parent-reported versus direct measures. Health Rep 2011 Sep;22(3):47-58 [FREE Full text] [Medline: 22106789]
13. Wright DR, Glanz K, Colburn T, Robson SM, Saelens BE. The accuracy of parent-reported height and weight for 6-12 year old U.S. children. BMC Pediatr 2018 Dec 12;18(1):52 [FREE Full text] [doi: 10.1186/s12887-018-1042-x] [Medline: 29433464]
14. Chai LK, Collins CE, May C, Ashman A, Holder C, Brown LJ, et al. Feasibility and efficacy of a web-based family telehealth nutrition intervention to improve child weight status and dietary intake: a pilot randomised controlled trial. J Telemed Telecare 2019 Jul 31:1357633X19865855. [doi: 10.1177/1357633X19865855] [Medline: 31364474]
15. Back2Basics Family. Callaghan, Australia: The University of Newcastle URL: http://www.cookingmattersaustralia.com.au/b2bfamily/ [accessed 2019-08-26]
16. 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]
17. Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatr Obes 2012 Aug;7(4):284-294. [doi: 10.1111/j.2047-6310.2012.00064.x] [Medline: 22715120]
18. McBride GB. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. NIWA Client Report: HAM2005-0622005. Hamilton, New Zealand: National Institute of Water & Atmospheric Research Ltd; 2005 May. URL: https://www.medcalc.org/download/pdf/McBride2005.pdf [accessed 2019-08-21]
19. Lipsey M, Wilson D. Practical Meta-Analysis (Applied Social Research Methods). Thousand Oaks, CA: Sage Publications, Inc; 2019.
20. Watson PF, Petrie A. Method agreement analysis: a review of correct methodology. Theriogenology 2010 Jun;73(9):1167-1179 [FREE Full text] [doi: 10.1016/j.theriogenology.2010.01.003] [Medline: 20138353]
21. Weden MM, Brownell PB, Rendall MS, Lau C, Fernandes M, Nazarov Z. Parent-reported height and weight as sources of bias in survey estimates of childhood obesity. Am J Epidemiol 2013 Aug 01;178(3):461-473 [FREE Full text] [doi: 10.1093/aje/kws477] [Medline: 23785115]
22. Ghosh-Dastidar MB, Haas AC, Nicosia N, Datar A. Accuracy of BMI correction using multiple reports in children. BMC Obes 2016;3(1):37 [FREE Full text] [doi: 10.1186/s40608-016-0117-1] [Medline: 27648293]
23. Van Cauwenberghe J, Delvaux I, Michels N, Den Hond E, Schoeters G, Nelen V, et al. Validity of parentally reported versus measured weight, length and waist in 7- to 9-year-old children for use in follow-up studies. Eur J Pediatr 2014 Jul;173(7):921-928. [doi: 10.1007/s00431-014-2274-x] [Medline: 24497184]
24. Gordon NP, Mellor RG. Accuracy of parent-reported information for estimating prevalence of overweight and obesity in a race-ethnically diverse pediatric clinic population aged 3 to 12. BMC Pediatr 2015 Feb 12;15:5 [FREE Full text] [doi: 10.1186/s12887-015-0320-0] [Medline: 25886135]
25. O'Connor DP, Gugenheim JJ. Comparison of measured and parents' reported height and weight in children and adolescents. Obesity (Silver Spring) 2011 May;19(5):1040-1046 [FREE Full text] [doi: 10.1038/oby.2010.278] [Medline: 21127476]
26. Huybrechts I, Himes JH, Ottevaere C, De Vriendt T, De Keyzer W, Cox B, et al. Validity of parent-reported weight and height of preschool children measured at home or estimated without home measurement: a validation study. BMC Pediatr 2011 Jul 07;11:63 [FREE Full text] [doi: 10.1186/1471-2431-11-63] [Medline: 21736757]
27. Nemecek D, Sebelefsky C, Woditschka A, Voitl P. Overweight in children and its perception by parents: cross-sectional observation in a general pediatric outpatient clinic. BMC Pediatr 2017 Dec 22;17(1):212 [FREE Full text] [doi: 10.1186/s12887-017-0964-z] [Medline: 29273009]
28. Huybrechts I, Beirlaen C, De Vriendt T, Slimani N, Pisa PT, Schouppe E, et al. Validity of instruction leaflets for parents to measure their child's weight and height at home: results obtained from a randomised controlled trial. BMJ Open 2014 Feb 07;4(2):e003768 [FREE Full text] [doi: 10.1136/bmjopen-2013-003768] [Medline: 24508849]

Abbreviations

BMI: body mass index
CCC: concordance correlation coefficient
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
eHealth: electronic health
IOTF: International Obesity Task Force
SMS: short message service

Edited by G Eysenbach; submitted 23.10.18; peer-reviewed by C Schnelle, J Recio-Rodriguez, P Wark, J Scott; comments to author 21.03.19; revised version received 20.06.19; accepted 02.08.19; published 16.09.19.

Please cite as:
Chai LK, Collins CE, May C, Holder C, Burrows TL
Accuracy of Parent-Reported Child Height and Weight and Calculated Body Mass Index Compared With Objectively Measured Anthropometrics: Secondary Analysis of a Randomized Controlled Trial
J Med Internet Res 2019;21(9):e12532
URL: https://www.jmir.org/2019/9/e12532/
doi:10.2196/12532
PMID:31538954

Original Paper

Predictors of Patients' Loyalty Toward Doctors on Web-Based Health Communities: Cross-Sectional Study

Tailai Wu[1], PhD; Zhaohua Deng[1], PhD; Zhuo Chen[2,3], PhD; Donglan Zhang[2], PhD; Xiang Wu[1], PhD; Ruoxi Wang[1], PhD

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[2]Department of Health Policy and Management, College of Public Health, University of Georgia, Athens, GA, United States
[3]School of Economics, Faculty of Humanities and Social Sciences, University of Nottingham Ningbo China, Ningbo, China

Corresponding Author:
Xiang Wu, PhD
School of Medicine and Health Management
Tongji Medical College
Huazhong University of Science and Technology
13 Hangkong Road
Wuhan, 430030
China
Phone: 86 13071253919
Email: wuhsiang@hust.edu.cn

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2019/11/e16942

Abstract

Background: Web-based health communities provide means for patients to not only seek care but also to promote their relationship with doctors. However, little is known about the predictors of patients' loyalty toward doctors in Web-based health communities.

Objective: This study aimed to investigate the predictors of patients' loyalty toward doctors in Web-based health communities.

Methods: On the basis of sociotechnical systems theory and attachment theory, we propose that social factors including emotional interaction, perceived expertise, and social norm influence patients' loyalty through their emotional attachment, whereas technical factors including sociability, personalization, and perceived security affect patients' loyalty through functional dependence. To validate our proposed research model, we used the survey method and collected 373 valid answers. Partial least square was used to analyze the data.

Results: Our empirical analysis results showed that all the social factors including emotional interaction (beta=.257, t_{350}=2.571; *P*=.01), perceived expertise (beta=.288, t_{350}=3.412; *P*=.001), and social norm (beta=.210, t_{350}=2.017; *P*=.04) affect patients' emotional attachment toward doctors significantly, whereas except sociability (beta=.110, t_{350}=1.152; *P*=.25), technical factors such as personalization (beta=.242, t_{350}=2.228; *P*=.03) and perceived security (beta=.328, t_{350}=3.438; *P*=.001) impact functional dependence significantly. Considering the effect of working mechanisms, both emotional attachment (beta=.443, t_{350}=4.518; *P*<.001) and functional dependence (beta=.303, t_{350}=2.672; *P*=.008) influence patients' loyalty toward doctors in Web-based health communities significantly.

Conclusions: Patients' loyalty toward doctors in Web-based health communities is important for the effectiveness of doctors' advice or service in Web-based health communities. The research results not only fill the gaps in the literature of the patient-doctor relationship and Web-based health communities but also has many implications for establishing patients' loyalty on Web-based health communities and in physical context.

(J Med Internet Res 2019;21(9):e14484) doi:10.2196/14484

KEYWORDS

medical informatics; telemedicine; patients; physicians; community network; psychological theory; social theory; health services

Introduction

Background

Many people use the internet to search for health-related information and use online health services. According to a report from the Pew Research Center, almost two-thirds of US adults seek health information online, and one-third of US adults self-diagnose using information from the internet [1]. Meanwhile, many Chinese consumers have used health services online with the online health market estimated to reach 90 billion RMB in 2020 [2]. Therefore, electronic health has become an important means for people to seek care. Among electronic health applications, Web-based health communities have become increasingly important with growing utilization [3].

Web-based health communities facilitate the communication between patients and doctors online [4]. Patients and doctors could discuss health-related topics in different discussion groups using communication tools embedded in Web-based health communities [5]. Therefore, Web-based health communities could change the traditional 1-way communication from doctors to patients to bidirectional communication between patients and doctors [6].

In a patient-physician relationship, patients' loyalty toward doctors is defined as their intention to revisit the doctor [7]. It has been shown to correlate with patients' compliance with doctors' medical advice, great use of medical services, and disclosure of medical information by the patient for doctors' diagnosis [8]. Therefore, it is critical to explore the predictors of patients' loyalty toward doctors. Previous literature has discussed this topic. For example, Rundle-Thiele and Russell-Bennett [9] found age and frequency of visits had significant effects on patients' loyalty through satisfaction. Sutharjana et al [10] showed that the effect of organizational citizen behaviors and service quality influenced patients' loyalty. Wu [11] demonstrated that hospital brand image had both direct and indirect effects on patients' loyalty through service quality and patient satisfaction. Platonova et al [12] concluded that patients' satisfaction, trust, and interpersonal relationship significantly affected patients' loyalty toward doctors. However, few studies have considered patients' loyalty in a Web-based health community context and studied the effect of technical environment on the formation of patients' loyalty. Meanwhile, although previous literature studied users' activities on the Web-based health communities, such as knowledge generation [13], knowledge sharing [14], and personal health information communication [15], few of them investigated patients' loyalty toward doctors in Web-based health communities. Therefore, we propose our research question *What are the predictors of patients' loyalty toward doctors in Web-based health communities?*

The remainder of this paper is organized as follows: we explore the predictors of patients' loyalty toward doctors in Web-based health communities using attachment theory and sociotechnical system theory. Our research model hypothesizes a causal effect of the explored predictors on patients' loyalty toward doctors in Web-based health communities. Then, consumer survey is used to collect data and validate the research model. The analysis results are discussed, and the implications are provided. The last section discusses the limitations of this work and future research directions.

Theoretical Foundation

To examine the research questions, we used sociotechnical systems theory to explore the predictors of patients' loyalty toward doctors in Web-based health communities. Meanwhile, to understand the working mechanisms of these predictors, we applied attachment theory to locate the mediating constructs for the predictors.

Sociotechnical systems theory was originally used to understand the organizational outcomes by dividing organizational systems into social subsystems and technical subsystems [16]. Social subsystems are mainly about people and their attributes such as values, attitudes, and skills. Meanwhile, technical subsystems concern the processes, tasks, and technologies to produce desired organizational outcomes. Therefore, the organizational outcomes are the results of the interaction between social subsystems and technical subsystems. The organizationally designated outcomes may come from the successful fit between the 2 subsystems [17]. Given that patients and doctors could constitute temporal discussion groups on Web-based health communities, patients' loyalty toward doctors in Web-based health communities could be the outcome of patient-doctor group interaction, and we could leverage sociotechnical systems theory to understand the drivers of patients' loyalty toward doctors on the Web-based health communities. On the Web-based health communities, the technical attributes could correspond to the technical subsystems, whereas the interaction between patients and doctors becomes the source of social subsystems. Therefore, we propose that emotional interaction between patients and doctors, perceived expertise of doctors, and social norm for patients are the social factors based on the social systems, whereas sociability, personalization, and perceived security of Web-based health communities are the technical factors according to the technical systems.

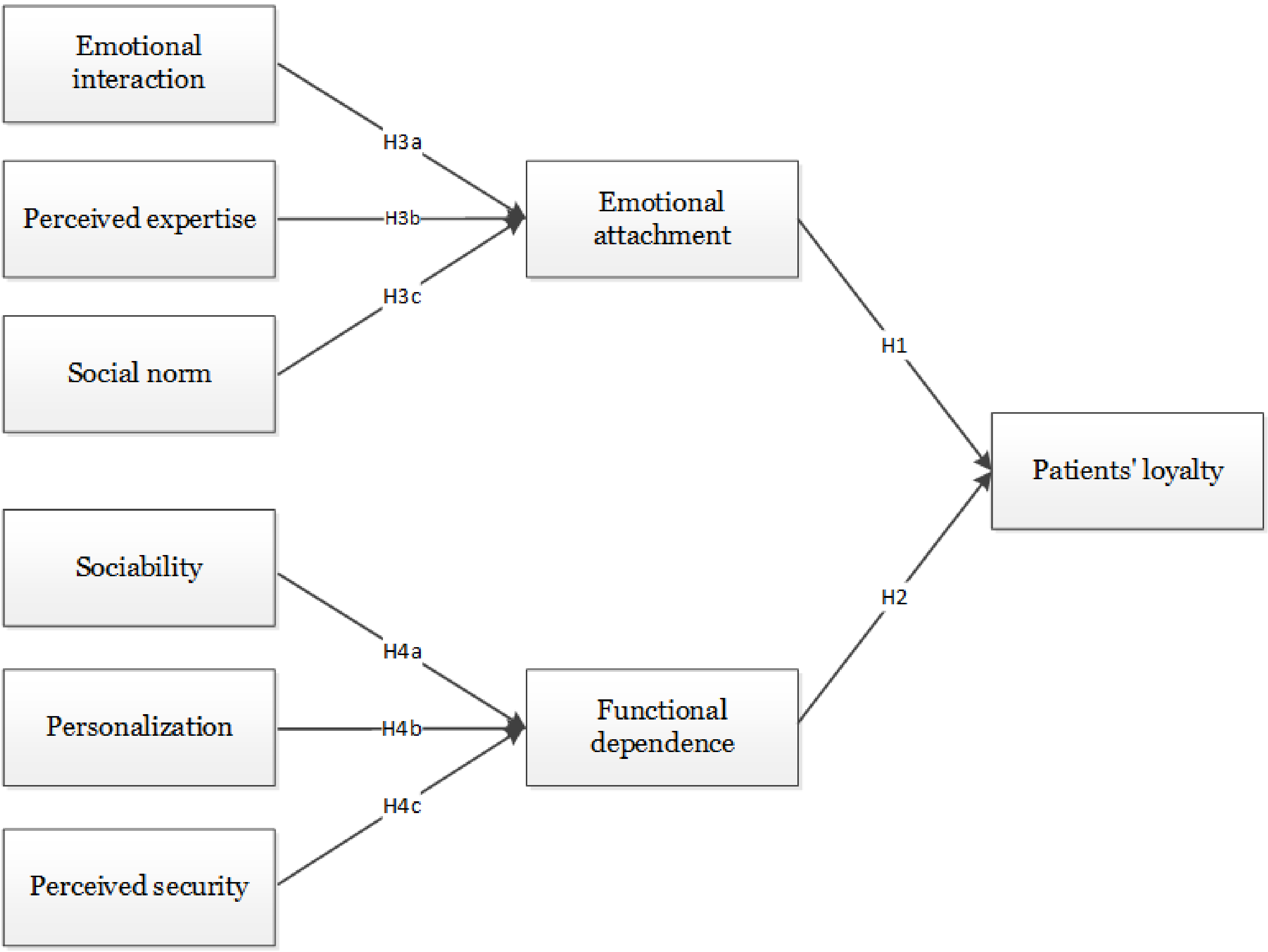
Attachment theory initially explained the reason why infants attach to their caregivers. This theory states that the attachment is described as the emotional-laden bond between infants and caregivers [18]. When the attachment is impaired, infants will be distressed or in anxiety. Several studies expanded the use of attachment theory to understand the relationship between product and consumers [19], brand and consumers [20], technology and consumers [21], and patients and providers [22]. As patients' loyalty toward doctors in Web-based health communities is the result of the physician-patient relationship, we can use attachment theory to understand the influence of predictors of patients' loyalty [12]. To be specific, we speculate that patients' emotional attachment and functional dependence on doctors in Web-based health communities impact patients' loyalty toward doctors in Web-based health communities.

On the basis of the above theories, we formulate our proposed relationships in a research model and develop corresponding hypotheses in the following section.

Research Model and Hypotheses

In line with the above reasoning, we hypothesize that emotional interaction between patients and doctors, perceived expertise of doctors, and social norm for patients influence patients' loyalty through emotional attachment, whereas sociability, personalization, and perceived security of Web-based health communities affect patients' loyalty through functional dependence. The research model is presented in Figure 1.

Figure 1. Research model.



Emotional Attachment and Patients' Loyalty

Patients' emotional attachment to doctors in Web-based health communities is patients' emotional bond with doctors in Web-based health communities [23]. Patients' emotional attachment to doctors predicts patients' commitment toward doctors [24]. Once a patient commits to a doctor, he visits the doctor persistently. Then, patients' emotional attachment to doctors may make them revisit the doctors. Therefore, we hypothesize as follows:

> *H1: Patients' emotional attachment to doctors positively influences patients' loyalty.*

Functional Dependence and Patients' Loyalty

Patients' functional dependence on doctors in Web-based health communities is the interdependence between patients and doctors for solving patients' health problems [21]. The interdependence between patients and doctors can also increase patients' commitment to doctors [25]. The increased commitment to doctors leads a patient to revisit the doctor. Therefore, we hypothesize as follows:

> *H2: Patients' functional dependence on doctors positively influences patients' loyalty.*

Social Factors and Emotional Attachment

In this study, we identify emotional interaction, perceived expertise, and social norm as social factors. Toward patients' emotional interaction with doctors in Web-based health communities, it is the interaction between patients and doctors that cares about patients' emotions [26]. Compared with the interaction that only focuses on solving patients' health problems, an interaction that cares about patients' emotions is more important because patients cannot judge professional treatment and their negative emotions that are caused by the diseases they are suffering needs to be cared for [27]. Meanwhile, through emotional interaction with doctors, patients may form trust toward doctors in Web-based health communities [28]. Patients' trust toward doctors helps build their emotional attachment to the doctors [24]. Therefore, we hypothesize as follows:

> *H3a: Patients' emotional interaction with doctors positively influences patients' emotional attachment.*

For patients' perceived expertise of doctors, it is patients' perception that the doctors are valid medical professionals [29]. Perceived expertise has been defined as 1 of the 3 dimensions of perceived credibility [29]. Given trust is a reflection of perceived credibility and one of the predictors of emotional attachment [30], we hypothesize as follows:

> *H3b: Patients' perceived expertise of doctors positively influences patients' emotional attachment.*

With regard to social norm, it is patients' perception that important others approve their interaction with doctors in Web-based health communities [31]. The approval from important others represents the identification of doctors by the social group patients belong to [32]. Patients may comply with the social norm to identify the patents. The identification of doctors in Web-based health communities can drive the patients' attachment to doctors [33]. Therefore, we hypothesize as follows:

> *H3c: Social norm positively influences patients' emotional attachment.*

Technical Factors and Functional Dependence

A total of 3 factors have also been figured out to represent the technical systems: sociability, personalization, and perceived security. These factors are all functional features of Web-based health communities. Considering sociability, it is the extent to which Web-based health communities facilitate the socialization between doctors and patients [34]. The socialization between doctors and patients underpinned by Web-based health communities can make patients feel the flow in interacting with doctors in Web-based health communities [35]. The sense of flow can give rise to patients' dependence on doctors functionally. Therefore, we hypothesize as follows:

> *H4a: Sociability positively influences patients' functional dependence.*

With respect to personalization, it is the extent to which Web-based health communities reflect patients' personal needs [36]. Different patients may have different conditions, symptoms, or treatment needs, whereas they may also have different feelings and perceptions toward their diseases based on their own experience. Web-based health communities with a high degree of personalization could satisfy different patients' needs to help them find suitable doctors for their health problems. Through interacting with a doctor who provides effective advice and support, a patient will build trust toward the doctor and rely on them [36]. Therefore, we hypothesize as follows:

> *H4b: Personalization positively influences patients' functional dependence.*

As to perceived security, it is the degree to which patients believe that their personal assets, such as their private information or money, will be safe when they interact with doctors in Web-based health communities [37]. When patients perceive secureness in interacting with doctors in Web-based health communities, they perceive less risky in the interaction. The lower degree of sense of risk results in the higher trust toward the doctors [38]. Therefore, we hypothesize as follows:

> *H4c: Perceived security positively influences patients' functional dependence.*

Methods

Measurement Instrument

A consumer survey was implemented in this study to test our model. To make the measurement instrument of survey, we adapted previously validated scales in our research context. Items for patients' loyalty were from Cyr [7]. Items for emotional attachment were adapted from Vlachos et al [24], whereas items for functional dependence were based on Wu et al [21]. Toward scales for social factors, items for emotional interaction were from Ben-Sira [39], items for perceived expertise were from Ohanian [29], and items for social norm were from Venkatesh et al [31]. With regards to scales for technical factors, items for sociability were from Animesh et al [35], items for personalization were from Zhou et al [40], and items for perceived security were from Chang and Chen [41]. We used the 5-point Likert scale to measure the items with anchors ranging from 1 (strongly disagree) to 5 (strongly agree).

Our survey was conducted in China. As our scales were originally developed in English, we used back-translation method to translate the English scales into Chinese. To be specific, 1 bilingual investigator translated the English scales into Chinese at first, and another bilingual investigator translated the Chinese scales into English. Then, the 2 investigators compared the 2 English versions to check whether there were significant differences between them. To ensure the validity of our measurement instrument, we conducted a pilot study by interviewing 7 experts in the area of medical informatics and 16 users of Web-based health communities. After collecting comments from them, we revised and decided our questionnaire for the survey. The measurement instrument is showed in Multimedia Appendix 1.

Besides the constructs in our hypotheses, several control variables were included such as gender, age, education, length and experience of using Web-based health communities, and perceived health [42]. To explore the non-naive effect of respondents, respondents' experience of filling online questionnaire was also considered [43].

Data Collection

Given that China has the largest number of internet users and almost two-thirds of them search health information online, we chose Chinese Web-based health community users as our respondents [44]. We adopted the data collection service from a leading Chinese online survey site to administrate the survey. To ensure the data quality and reduce social desirability bias, we took several measures in designing the questionnaire and cleaning the collected data. For example, we added attention-trap and reverse-coded questions to decrease monomethod bias, whereas we set screening questions to make sure the answers from the right respondents. The screening questions included which Web-based health communities had been used, whether you were the registered member, and whether you had interacted with doctors on the Web-based health communities. Finally, we excluded cases with missing values or similar values for all questions. Our study procedures had been approved by the Institutional Review Board of Tongji Medical College, Huazhong University of Science and Technology (No. 2017S319).

Results

Demographic Information

Through 2 weeks' survey, we obtained 373 complete and valid responses. In this sample, most of the respondents were in the age group of 30 years and older, females, with college degrees, and were familiar with using Web-based health communities. These results are reasonably consistent with the report of the China Internet Network Information Center on demographics of Chinese internet users [44]. The specific demographic information of this sample is summarized in Table 1.

Table 1. Demographic information.

Characteristics		Statistics, n (%)
Age (years)		
	18-25	67 (18)
	25-30	138 (37)
	>30	168 (45)
Gender		
	Male	160 (42.9)
	Female	213 (57.1)
Education		
	High school	8 (2.1)
	College	330 (88.5)
	Master's degree and above	35 (9.4)
Intensity of using Web-based health communities (hours/day)		
	<0.5	305 (81.8)
	0.5-1	59 (15.8)
	>1	9 (2.4)
Length of using Web-based health communities (years)		
	<1	137 (36.7)
	1-5	228 (61.1)
	>5	8 (2.2)

Reliability and Validity

To test the reliability and validity of our measurement instrument, we used partial least square (PLS) technique to analyze the collected data in SmartPLS 2.0.3M from SmartPLS GmbH in Germany [45]. Using PLS technique to conduct the confirmatory factor analysis, the values of indicators for reliability and validity are listed in Tables 2 and 3. As shown in Table 2, all the values of Cronbach alpha and composite reliabilities were above 0.7, indicating acceptable reliability for all constructs [46]. Meanwhile, the value of average variance extracted (AVE) of each construct was above 0.5, and loading weights for each item were above 0.7, thus confirming good convergent validity [46]. As shown in Table 3, the values in the square roots of AVEs were all greater than the interconstruct correlations, thus showing good discriminant validity [47]. Therefore, we concluded that our measurement instrument has enough reliability and validity to test the structural model of our proposed hypotheses.

Table 2. Construct reliability and convergent validity.

Construct	Items	Factor loadings	Composite reliability	Average variance extracted	Cronbach alpha
Patient's loyalty					
	LYT1	0.7904	0.836	0.6299	.7062
	LYT2	0.755	—[a]	—	—
	LYT3	0.8336	—	—	—
Emotional attachment					
	EA1	0.8374	0.8676	0.6859	.771
	EA2	0.8128	—	—	—
	EA3	0.8343	—	—	—
Functional dependence					
	FD1	0.7761	0.8337	0.626	.7016
	FD2	0.7586	—	—	—
	FD3	0.8368	—	—	—
Emotional interaction					
	EI1	0.8635	0.8702	0.7703	.7027
	EI2	0.8917	—	—	—
Perceived expertise					
	PE1	0.7863	0.8436	0.6429	.7228
	PE2	0.7733	—	—	—
	PE3	0.8441	—	—	—
Social norm					
	SN1	0.8802	0.8862	0.7957	.7438
	SN2	0.9036	—	—	—
Sociability					
	SBY1	0.7974	0.833	0.6248	.7066
	SBY2	0.8255	—	—	—
	SBY3	0.7464	—	—	—
Personalization					
	PLN1	0.8905	0.8838	0.7919	.7372
	PLN2	0.8892	—	—	—
Perceived security					
	PS1	0.8808	0.8815	0.7881	.7313
	PS2	0.8946	—	—	—

Table 3. Discriminant validity.

Constructs	LYT[a]	EA[b]	FD[c]	EI[d]	PE[e]	SN[f]	SBY[g]	PLN[h]	PS[i]
LYT	0.7937[j]	—[k]	—	—	—	—	—	—	—
EA	0.5511	0.8282[j]	—	—	—	—	—	—	—
FD	0.4668	0.379	0.7912[j]	—	—	—	—	—	—
EI	0.4208	0.4427	0.3711	0.8777[j]	—	—	—	—	—
PE	0.4662	0.4254	0.3766	0.3013	0.8018[j]	—	—	—	—
SN	0.4812	0.3895	0.362	0.3832	0.257	0.892[j]	—	—	—
SBY	0.3407	0.3628	0.2317	0.3822	0.2085	0.4218	0.7904[j]	—	—
PLN	0.2677	0.168	0.3459	0.1957	0.1542	0.1212	0.2164	0.8899[j]	—
PS	0.4388	0.3564	0.4101	0.3778	0.3454	0.3177	0.211	0.2448	0.8878[j]

[a] LYT: loyalty.

[b] EA: emotional attachment.

[d] EI: emotional interaction.

[c] FD: functional dependence.

[e] PE: perceived expertise.

[f] SN: social norm.

[g] SBY: sociability.

[h] PLN: personalization.

[i] PS: perceived security.

[j] The square roots of average variance extracted.

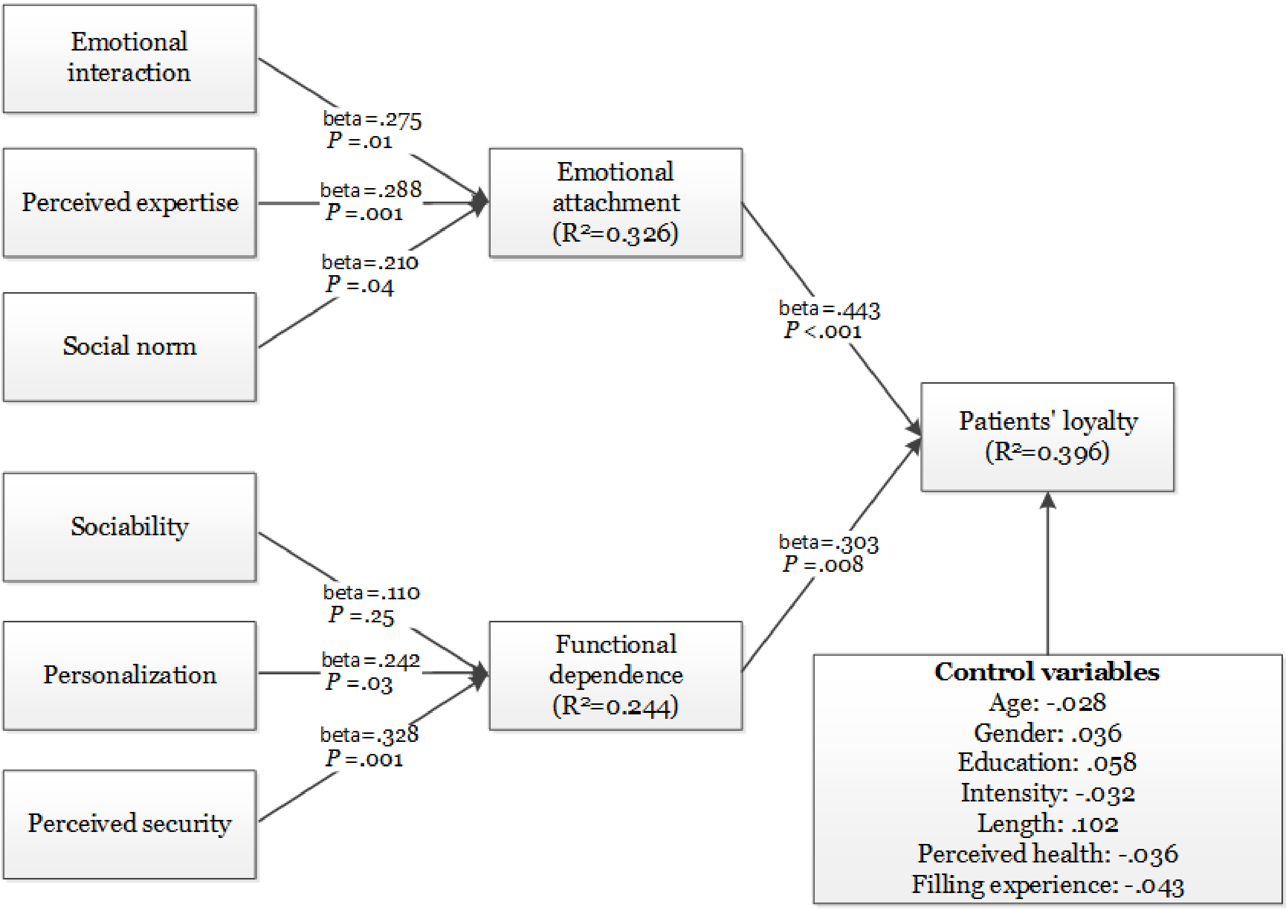
[k] —: Not applicable

Given that we used the measurement instrument to collect data for all the constructs, the possibility of common method bias was tested. First, we checked the values of correlation coefficient among constructs in Table 3 and found that there were no pairs with very high correlation (r>0.90). Second, Harman single factor test was conducted by principal component analysis in SPSS 22.0. The first extracted factor in the unrotated solution accounted for 29.44%, which is less than 50% [48]. Finally, the marker variable technique was used to test the bias. Perceived organizational support, which is not relevant to our study theoretically, was set as the marker variable [49]. The average value of correlation coefficients between perceived organizational support and other variables was 0.196. Therefore, we rule out the common method bias in our study.

Hypotheses Testing

Through using the bootstrapping procedures in PLS, we tested the hypothesized relationships in our research model by computing the *t* values of each path. The analysis results are listed in Figure 2. According to Figure 2, both emotional attachment (beta=.443; *P*<.001) and functional dependence (beta=.303; *P*=.008) affected patients' loyalty toward doctors in Web-based health communities significantly. Therefore, both H1 and H2 were supported. These results suggested that attachment theory provides a useful theoretical perspective to understand patients' loyalty. Toward the effect of social factors, emotional interaction (beta=.275; *P*=.01), perceived expertise (beta=.288; *P*=.001) and social norm (beta=.210; *P*=.04) all influenced emotional attachment significantly. Therefore, H3a, H3b, and H3c were all supported. These results suggested that factors in social systems are important for the formation of emotional attachment. With regards to the effect of technical system, sociability (beta=.110; *P*=.25) did not have significant influence on functional dependence, whereas personalization (beta=.242; *P*=.03) and perceived security (beta=.328; *P*=.001) had significant influence on functional dependence. Therefore, we cannot reject H4b and H4c, whereas H4a was not supported. These results reveal that factors in technical systems are also important for functional dependence. Besides, none of the control variables had a statistically significant effect on patients' loyalty. We think our findings are sound and strong for the following reasons: first, our measurement instrument is tested reliable and valid, and the effect of common method bias is insignificant. Second, the quality of data collection is guaranteed because several measures such as screening questions and attention-trap and reverse-rode questions are used in the questionnaires.

Figure 2. Analysis results of hypothesized model.



Discussion

Principal Findings

In this paper, we investigated patients' loyalty toward doctors in Web-based health communities based on attachment theory and sociotechnical systems theory. According to sociotechnical systems, we proposed emotional interaction, perceived expertise, and social norm can represent the social systems, which lead to patients' loyalty, whereas sociability, personalization, and perceived security comprise the technical systems, which help the formation of patients' loyalty. Meanwhile, to explore the working mechanisms of these predictors, we leveraged attachment theory. On the basis of attachment theory, we proposed that patients' emotional attachment corresponds to social factors and functional dependence link to technical factors. All hypothesized relationships were confirmed, except the relationship between sociability and functional dependence.

Among our proposed predictors, perceived security is shown to have the greatest impact on functional dependence, whereas perceived expertise seems to have the greatest impact on emotional attachment. This result uncovers that perceived expertise of doctors and perceived security of using the Web-based health platform are the most concerning factors when patients decide to interact with doctors in Web-based health platforms.

For the insignificant effect of sociability on functional dependence, the possible reason is that the purpose of interaction between patients and doctors is single, which is to solve consumers' health problems to some degree. Therefore, patients may not pay attention to social tools on Web-based health communities and only use few tools to interact and communicate with doctors.

Implications

This study has important theoretical implications. First, the application of the theoretical model on survey data provides a better understanding of patients' loyalty to doctors in Web-based health communities. Second, this study contributes to the patient-doctor relationship literature by studying the Web-based health community context. Although previous literature about patient-doctor relationship has also studied patients' loyalty, they study it mostly in the offline context while considering less about the influence of a Web-based health community, which is becoming a more important application in the field of digital health. Finally, this study contributes to the Web-based health community literature by studying patients' loyalty. Although previous literature about Web-based health communities also studied patients' activities, few have examined patients' loyalty. Patients' loyalty reflects the relationship between patients and doctors, which may determine the development of Web-based health communities.

Meanwhile, this study also has important implications for practice. First, this study suggests that Web-based health communities could be an important channel to improve the patient-physician relationship. Therefore, health policy makers may encourage the use of Web-based health communities.

Second, predictors that represent both social and technical systems could be the basis to develop many actionable strategies to increase patients' loyalty for clinical practice. For example, based on social factors including emotional interaction, perceived expertise, and social norm, doctors could provide more emotional support to patients, manifesting their expertise by displaying their qualification certificates and collaborate with patient group leaders to organize their virtual relationships with patients. Meanwhile, based on technical factors including personalization and perceived security, designers and managers of the Web-based health communities should understand patients' needs to provide personalized functions and services and protect private assets including patient information and budget. These strategies can also be used by doctors, managers, and policy makers in offline context to strengthen patients' loyalty. Finally, given that emotional attachment and functional dependence have significant influence on patients' loyalty, the measurement scales of emotional attachment and functional dependence could be effective indicators for the formation of patients' loyalty.

Limitation and Future Research Direction

We point to several future research directions based on the limitations of this study. First, a technical factor–sociability does not have significant influence on functional dependence. Therefore, more factors can be explored to improve the explanatory power of this empirical model. Second, we only investigated Chinese Web-based health community users. Chinese patients' attitudes, beliefs, and values about physician may be different from patients elsewhere [50]. Therefore, Chinese patients may behave differently when they interact with physicians. Meanwhile, culture could influence several consequences of patient-physician relationship such as satisfaction, accessibility, and continuity of care [51]. Future studies could validate this model by examining the patient-physician relationship in other cultural contexts. Finally, we used the cross-sectional data for our research model and ignored the time change effect of variables in our model. Future research can consider using the longitudinal data to capture the dynamics of our model.

Conclusions

This study explores the predictors of patients' loyalty toward doctors in Web-based health communities based on sociotechnical systems theory and attachment theory. Several social and technical factors are revealed, and their working mechanisms are confirmed in our empirical study. This study not only provides direct guidance to establish patients' loyalty in the Web-based health community context but also conveys the implications for building patients' loyalty physically.

Acknowledgments

This work was partially supported by grants from the National Natural Science Foundation of China (No 71801100, 71671073, 71602063, 71704059, 71701075, and 71810107003).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey instrument.
[PDF File (Adobe PDF File), 55KB - jmir_v21i9e14484_app1.pdf]

References

1. Fox S, Duggan M. Pew Research Center. 2013. Health Online 2013 URL: https://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2019-08-08]
2. Li N. Billion Euro. 2018. 90 Billion Yuan Market Size! The Internet Medical Industry Has Entered an Expansion Phase URL: https://www.iyiou.com/p/71492.html [accessed 2019-02-13] [WebCite Cache ID 769AZlvF7]
3. Goh JM, Gao GG, Agarwal R. The creation of social value: can an online health community reduce rural-urban health disparities? MIS Q 2016 Jan 1;40(1):247-263. [doi: 10.25300/MISQ/2016/40.1.11]
4. Maloney-Krichmar D, Preece J. A multilevel analysis of sociability, usability, and community dynamics in an online health community. ACM T Comput-Hum Int 2005 Jun 1;12(2):201-232. [doi: 10.1145/1067860.1067864]
5. Ridings CM, Gefen D. Virtual community attraction: why people hang out online. J Comput Mediat Commun 2004;10(1):-. [doi: 10.1111/j.1083-6101.2004.tb00229.x]
6. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471. [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
7. Cyr D. Modeling web site design across cultures: relationships to trust, satisfaction, and e-loyalty. J Manag Inf Syst 2014 Dec 8;24(4):47-72. [doi: 10.2753/MIS0742-1222240402]
8. MacStravic S. Patient loyalty to physicians. J Health Care Mark 1994;14(4):53-56. [Medline: 10154637]
9. Rundle-Thiele S, Russell-Bennett R. Patient influences on satisfaction and loyalty for GP services. Health Mark Q 2010 Apr;27(2):195-214. [doi: 10.1080/07359681003745162] [Medline: 20446141]

10. Sutharjana NW, Thoyib A, Taroena EA, Rahayu M. Organizational citizenship behavior effect on patient satisfaction and loyalty through service quality (study on maternity hospitals in Indonesia). Int J Sci Technol Res 2013;2(5):288-299 [FREE Full text]
11. Wu CC. The impact of hospital brand image on service quality, patient satisfaction and loyalty. Afr J Bus Manage 2011;5(12):4873-4882. [doi: 10.5897/AJBM10.1347]
12. Platonova EA, Kennedy KN, Shewchuk RM. Understanding patient satisfaction, trust, and loyalty to primary care physicians. Med Care Res Rev 2008 Dec;65(6):696-712. [doi: 10.1177/1077558708322863] [Medline: 18723450]
13. Zhao J, Ha S, Widdows R. The influence of social capital on knowledge creation in online health communities. Inf Technol Manag 2015 Jan 4;17(4):311-321. [doi: 10.1007/s10799-014-0211-3]
14. Zhang X, Liu S, Deng Z, Chen X. Knowledge sharing motivations in online health communities: a comparative study of health professionals and normal users. Comput Human Behav 2017 Oct;75:797-810. [doi: 10.1016/j.chb.2017.06.028]
15. Kordzadeh N, Warren J. Communicating personal health information in virtual health communities: an integration of privacy calculus model and affective commitment. J Assoc Inf Syst 2017 Feb;18(1):45-81. [doi: 10.17705/1jais.00446]
16. Bostrom RP, Heinen JS. MIS problems and failures: a socio-technical perspective, part II: the application of socio-technical theory. MIS Q 1977 Dec;1(4):11-28. [doi: 10.2307/249019]
17. Katz D, Kahn RL. The Social Psychology of Organizations. New York: Wiley; 1972.
18. Bowlby J. Attachment: Attachment and Loss. First Edition. New York: Basic Books; 1969.
19. Schifferstein HN, Zwartkruis-Pelgrim EP. Consumer-product attachment: measurement and design implications. Int J Des 2008;2(3):1-13 [FREE Full text]
20. Park CW, Macinnis DJ, Priester J, Eisingerich AB, Iacobucci D. Brand attachment and brand attitude strength: conceptual and empirical differentiation of two critical brand equity drivers. J Mark 2010 Nov;74(6):1-17. [doi: 10.1509/jmkg.74.6.1]
21. Wu T, Lu Y, Gong X, Gupta S. A study of active usage of mobile instant messaging application. Inf Dev 2016 Jul 9;33(2):153-168. [doi: 10.1177/0266666916646814]
22. Ciechanowski PS, Katon WJ, Russo JE, Walker EA. The patient-provider relationship: attachment theory and adherence to treatment in diabetes. Am J Psychiatry 2001 Jan;158(1):29-35. [doi: 10.1176/appi.ajp.158.1.29] [Medline: 11136630]
23. Thomson M, MacInnis DJ, Park CW. The ties that bind: measuring the strength of consumers' emotional attachments to brands. J Consumer Psychol 2005 Jan;15(1):77-91. [doi: 10.1207/s15327663jcp1501_10]
24. Vlachos PA, Theotokis A, Pramatari K, Vrechopoulos A. Consumer-retailer emotional attachment: some antecedents and the moderating role of attachment anxiety. Eur J Mark 2010 Sep 21;44(9/10):1478-1499. [doi: 10.1108/03090561011062934]
25. Teng CI, Chen MY, Chen YJ, Li YJ. Loyalty due to others: the relationships among challenge, interdependence, and online gamer loyalty. J Comput-Mediat Comm 2012 Sep 13;17(4):489-500. [doi: 10.1111/j.1083-6101.2012.01586.x]
26. Wu T, Deng Z, Feng Z, Gaskin DJ, Zhang D, Wang R. The effect of doctor-consumer interaction on social media on consumers' health behaviors: cross-sectional study. J Med Internet Res 2018 Feb 28;20(2):e73 [FREE Full text] [doi: 10.2196/jmir.9003] [Medline: 29490892]
27. Ben-Sira Z. The function of the professional's affective behavior in client satisfaction: a revised approach to social interaction theory. J Health Soc Behav 1976 Mar;17(1):3-11. [doi: 10.2307/2136462] [Medline: 1270783]
28. Hill NS, Bartol KM, Tesluk PE, Langa GA. Organizational context and face-to-face interaction: influences on the development of trust and collaborative behaviors in computer-mediated groups. Organ Behav Hum Dec 2009 Mar;108(2):187-201. [doi: 10.1016/j.obhdp.2008.10.002]
29. Ohanian R. Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. J Advert 1990 Oct;19(3):39-52. [doi: 10.1080/00913367.1990.10673191]
30. Everard A, Galletta DF. How presentation flaws affect perceived site quality, trust, and intention to purchase from an online store. J Manage Inform Syst 2014 Dec 9;22(3):56-95. [doi: 10.2753/MIS0742-1222220303]
31. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
32. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
33. Kim HR, Lee M, Lee HT, Kim NM. Corporate social responsibility and employee–company identification. J Bus Ethics 2010 Feb 13;95(4):557-569. [doi: 10.1007/s10551-010-0440-2]
34. Preece J. Sociability and usability in online communities: determining and measuring success. Behav Inform Technol 2010 Nov 8;20(5):347-356. [doi: 10.1080/01449290110084683]
35. Animesh A, Pinsonneault A, Yang SB, Oh W. An odyssey into virtual worlds: exploring the impacts of technological and spatial environments on intention to purchase virtual products. MIS Q 2011;35(3):789-810. [doi: 10.2307/23042809]
36. Komiak SY, Benbasat I. The effects of personalization and familiarity on trust and adoption of recommendation agents. MIS Q 2006;30(4):941-960. [doi: 10.2307/25148760]
37. Hartono E, Holsapple CW, Kim KY, Na KS, Simpson JT. Measuring perceived security in B2C electronic commerce website usage: a respecification and validation. Decis Support Syst 2014 Jun;62:11-21. [doi: 10.1016/j.dss.2014.02.006]
38. Kim MJ, Chung N, Lee CK. The effect of perceived trust on electronic commerce: shopping online for tourism products and services in south Korea. Tour Manage 2011 Apr;32(2):256-265. [doi: 10.1016/j.tourman.2010.01.011]

39. Ben-Sira Z. Affective and instrumental components in the physician-patient relationship: an additional dimension of interaction theory. J Health Soc Behav 1980 Jun;21(2):170-180. [doi: 10.2307/2136736] [Medline: 7391531]
40. Zhou Z, Fang Y, Vogel DR, Jin XL, Zhang X. Attracted to or locked in? Predicting continuance intention in social virtual world services. J Manage Inform Syst 2014 Dec 8;29(1):273-306. [doi: 10.2753/MIS0742-1222290108]
41. Chang HH, Chen SW. Consumer perception of interface quality, security, and loyalty in electronic commerce. Inform Manage 2009 Oct;46(7):411-417. [doi: 10.1016/j.im.2009.08.002]
42. Wu T, Deng Z, Zhang D, Buchanan PR, Zha D, Wang R. Seeking and using intention of health information from doctors in social media: the effect of doctor-consumer interaction. Int J Med Inform 2018 Jul;115:106-113. [doi: 10.1016/j.ijmedinf.2018.04.009] [Medline: 29779712]
43. Goodman JK, Paolacci G. Crowdsourcing consumer research. J Consum Res 2017;44(1):196-210. [doi: 10.1093/jcr/ucx047]
44. CNNIC. 2018. The 42th China Statistical Report on Internet Development URL: https://cnnic.com.cn/IDR/ReportDownloads/201807/P020180711391069195909.pdf [accessed 2019-08-08]
45. Ringle CM, Wende S, Will A. SmartPLS. 2005. URL: http://www.smartpls.com [accessed 2019-08-08]
46. Nunnally JC, Bernstein IH, Berge JM. Psychometric Theory. New York: Mcgraw-Hill College; 1967.
47. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39-50. [doi: 10.2307/3151312]
48. Podsakoff PM, Organ DW. Self-reports in organizational research: problems and prospects. J Manag 2016 Sep 4;12(4):531-544. [doi: 10.1177/014920638601200408]
49. Lindell MK, Whitney DJ. Accounting for common method variance in cross-sectional research designs. J Appl Psychol 2001 Feb;86(1):114-121. [doi: 10.1037/0021-9010.86.1.114] [Medline: 11302223]
50. Hofstede G. Culture's Consequences: International Differences in Work-Related Values. Thousand Oaks, CA: Sage Publications; 1984.
51. Flores G. Culture and the patient-physician relationship: achieving cultural competency in health care. J Pediatr 2000 Jan;136(1):14-23. [doi: 10.1016/s0022-3476(00)90043-x] [Medline: 10636968]

Abbreviations

AVE: average variance extracted
EA: emotional attachment
EI: emotional interaction
FD: functional dependence
LYT: loyalty
PE: perceived expertise
PLN: personalization
PLS: partial least square
PS: perceived security
SBY: sociability
SN: social norm

Edited by G Eysenbach; submitted 24.04.19; peer-reviewed by A Eiser, M Reynolds, S Saeed; comments to author 10.07.19; revised version received 23.07.19; accepted 23.07.19; published 03.09.19.

Please cite as:
Wu T, Deng Z, Chen Z, Zhang D, Wu X, Wang R
Predictors of Patients' Loyalty Toward Doctors on Web-Based Health Communities: Cross-Sectional Study
J Med Internet Res 2019;21(9):e14484
URL: https://www.jmir.org/2019/9/e14484/
doi:10.2196/14484
PMID:31482855

Review

Implementation Strategies to Enhance the Implementation of eHealth Programs for Patients With Chronic Illnesses: Realist Systematic Review

Cecilie Varsi[1], PhD; Lise Solberg Nes[1,2,3], PhD; Olöf Birna Kristjansdottir[1,4], PhD; Saskia M Kelders[5,6], PhD; Una Stenberg[4], PhD; Heidi Andersen Zangi[7,8], PhD; Elin Børøsund[1], PhD; Karen Elizabeth Weiss[9], PhD; Audun Stubhaug[2,10,11], PhD; Rikke Aune Asbjørnsen[1,5,12], MSc; Marianne Westeng[1], MSc; Marte Ødegaard[13], MSc; Hilde Eide[1,14], PhD

[1]Center for Shared Decision Making and Collaborative Care Research, Division of Medicine, Oslo University Hospital, Oslo, Norway
[2]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
[3]Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, United States
[4]Norwegian National Advisory Unit on Learning and Mastery in Health, Oslo University Hospital, Oslo, Norway
[5]Center for eHealth and Wellbeing Research, Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands
[6]Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa
[7]National Advisory Unit on Rehabilitation in Rheumatology, Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
[8]Faculty of Health, VID Specialized University, Oslo, Norway
[9]Department of Anesthesiology and Pain Medicine, School of Medicine, University of Washington, Seattle, WA, United States
[10]Regional Advisory Unit on Pain, Oslo University Hospital, Oslo, Norway
[11]Department of Pain Management and Research, Division of Emergencies and Critical Care, Oslo University Hospital, Oslo, Norway
[12]Department of Research and Innovation, Vestfold Hospital Trust, Tønsberg, Norway
[13]University of Oslo Library, University of Oslo, Oslo, Norway
[14]Science Centre Health and Technology, University of South-Eastern Norway, Drammen, Norway

Corresponding Author:
Cecilie Varsi, PhD
Center for Shared Decision Making and Collaborative Care Research
Division of Medicine
Oslo University Hospital
PB 4950 Nydalen
Oslo
Norway
Phone: 47 90856339
Email: cecilie.varsi@rr-research.no

Abstract

Background: There is growing evidence of the positive effects of electronic health (eHealth) interventions for patients with chronic illness, but implementation of such interventions into practice is challenging. Implementation strategies that potentially impact implementation outcomes and implementation success have been identified. Which strategies are actually used in the implementation of eHealth interventions for patients with chronic illness and which ones are the most effective is unclear.

Objective: This systematic realist review aimed to summarize evidence from empirical studies regarding (1) which implementation strategies are used when implementing eHealth interventions for patients with chronic illnesses living at home, (2) implementation outcomes, and (3) the relationship between implementation strategies, implementation outcomes, and degree of implementation success.

Methods: A systematic literature search was performed in the electronic databases MEDLINE, Embase, PsycINFO, Scopus, Allied and Complementary Medicine Database, Cumulative Index to Nursing and Allied Health Literature, and Cochrane Library. Studies were included if they described implementation strategies used to support the integration of eHealth interventions into practice. Implementation strategies were categorized according to 9 categories defined by the Expert Recommendations for Implementing Change project: (1) engage consumers, (2) use evaluative and iterative strategies, (3) change infrastructure, (4) adapt and tailor to the context, (5) develop stakeholder interrelationships, (6) use financial strategies, (7) support clinicians, (8) provide interactive assistance, and (9) train and educate stakeholders. Implementation outcomes were extracted according to the implementation outcome framework by Proctor and colleagues: (1) acceptability, (2) adoption, (3) appropriateness, (4) cost, (5) feasibility, (6) fidelity, (7) penetration, and (8) sustainability. Implementation success was extracted according to the study authors' own evaluation of implementation success in relation to the used implementation strategies.

Results: The implementation strategies management support and engagement, internal and external facilitation, training, and audit and feedback were directly related to implementation success in several studies. No clear relationship was found between the number of implementation strategies used and implementation success.

Conclusions: This is the first review examining implementation strategies, implementation outcomes, and implementation success of studies reporting the implementation of eHealth programs for patients with chronic illnesses living at home. The review indicates that internal and external facilitation, audit and feedback, management support, and training of clinicians are of importance for eHealth implementation. The review also points to the lack of eHealth studies that report implementation strategies in a comprehensive way and highlights the need to design robust studies focusing on implementation strategies in the future.

Trial Registration: PROSPERO CRD42018085539; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=85539

(J Med Internet Res 2019;21(9):e14255) doi:10.2196/14255

KEYWORDS

chronic illness; eHealth; implementation; implementation strategies; implementation outcomes; realist review

Introduction

Electronic health (eHealth), defined as "health services and information delivered or enhanced through the Internet and related technologies" [1], has great potential for persons with chronic or long-term illnesses. For example, eHealth provides options for self-management, patient-provider communication, monitoring, and shared decision making [2-5]. A growing body of evidence indicates positive effects of eHealth services on patient health outcomes [6-9]. For example, telehealth is regarded as a safe option for delivery of self-management support [10], and internet-delivered cognitive behavioral therapy (ICBT) has shown promising results as an alternative to traditional face-to-face interventions among persons with chronic health illnesses [3]. Similarly, studies indicate that eHealth services can be effective in reducing hospital admissions for patients with chronic illnesses such as chronic obstructive pulmonary disease [11] and reducing symptoms of anxiety and depression [12] and may contribute to successful self-management of chronic pain [13]. Moreover, patients using eHealth services have reported high levels of acceptability and satisfaction [11,14], and health care providers have described clinical benefits from eHealth services [2]. Understanding more about the implementation of eHealth services for patients with chronic conditions, a large patient group with unpredictable disease trajectories and the need for coordinated long-term multidisciplinary follow-up, would be beneficial and could lead to successful implementation in other areas.

Even with a growing number of eHealth programs, many of which have shown promising results [15], the actual implementation of such programs into everyday use in clinical practice has proven to be challenging [16,17]. The implementation process can be demanding and requires significant effort to succeed [18]. The phase between the organizational decision to adopt an eHealth program and the health care providers' routine use of that program is multifaceted and complex [4,15,18]. Implementation strategies, defined as "a systematic intervention process to adopt and integrate evidence-based health innovations into usual care" [19], can aid the implementation of eHealth programs into practice [18-20]. Implementation strategies constitute the how-to component of changing health care practice [20], and a number of known implementation strategies can possibly impact implementation success [19,21,22]. When implementation is initiated in a clinical health care setting, the use of implementation strategies refers to the concrete activities taken to make patients and health care providers start and maintain use of new evidence within the clinical setting. Implementation strategies are often part of an implementation plan, which describes what will be implemented, to whom, how and when, with the implementation strategies constituting the how-to in the plan. The implementation strategies can include a wide range of activities directed toward different stakeholders (eg, involvement of health care providers and patients, training and follow-up in the delivery of the clinical intervention, leadership engagement and internal and external support) [22]. The implementation strategies can be used as standalone (discrete) strategies or as a combination of strategies (multifaceted) [23]. Even though the research on implementation strategies is still in its infancy, there is a growing recognition that implementation will not happen automatically and that use of implementation strategies can be effective, particularly as they target those intending to use the new evidence directly [23,24].

Despite existing implementation strategy taxonomies and implementation process models (ie, practical guidance in the use of implementation strategies to facilitate implementation) [25], and the fact that some organizations have developed a set of implementation strategies for use in their own implementation processes [26], there is still limited understanding regarding which strategies to use and the relative importance of these strategies when promoting use of evidence-based interventions in clinical practice [22,27]. Notably, Greenhalgh and colleagues [28,29], who recently developed and tested a framework for nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) of health and care technologies, argue that technology implementation will not succeed until the complexities of multiple, interacting domains (eg, the illness, the technology, the organization, and the implementation process) are taken into account and addressed. With the exception of a few initiatives such as the one taken by the Greenhalgh group, little emphasis has been placed on the planning of and reporting on implementation strategies related to the implementation of eHealth interventions into practice [30,31]. Research literature has summarized different aspects of eHealth implementation [18], including barriers and facilitators [32-34], frontline staff acceptance of eHealth technologies [35], patient recruitment strategies [36], and eHealth implementation in rural areas [31]. However, the empirical research literature on strategies for eHealth implementation has not yet been reviewed or summarized. Also, the relationship between implementation strategies, implementation outcomes, and implementation success is rarely adequately described.

Implementation outcomes can be measured by means of various methods (eg, qualitative, quantitative, mixed), and the success of the implementation effort can be evaluated on the basis of implementation outcomes [37]. When seeking to understand implementation outcomes, researchers have stated that the relative importance of each single outcome measurement may vary in importance depending on stakeholders and may have different consequences depending on setting [37]. This indicates that implementation success is not necessarily derived directly from the implementation outcome measurements. Therefore, assessment of implementation success in addition to implementation outcomes can, as pointed out by Proctor and colleagues [37], play an important role in understanding and assessing the success of the implementation effort.

This project sought to further research and gain knowledge in this area through a systematic realist review. The realist review approach involves identifying how and why interventions work (or fail to work) in different contexts and examines the links between context, mechanisms, and outcomes [38]. Unlike classical systematic reviews, realist reviews focus not only on if the program works but also on how, why, and for whom [38]. The approach is often described as "what works for whom under what circumstances and why." As noted by Rycroft-Malone and colleagues [39], the realist review method is especially suited when conducting reviews on implementation, due to implementation processes' complex, multifaceted nature and the limited understanding of their mechanisms of action [39]. This systematic realist review aimed to summarize evidence from empirical studies regarding (1) which implementation strategies were used when implementing eHealth interventions for patients with chronic illnesses living at home; (2) which implementation outcomes were achieved; and (3) the relationship between implementation strategies, implementation outcomes, and degree of implementation success.

Methods

Overview

A systematic realist review, by means of an aggregative approach using predefined concepts (ie, implementation strategies and implementation outcomes) [40] was considered suitable to provide an explanatory analysis focusing on which implementation strategies were used, in what circumstances, how, and leading to which implementation outcomes. In addition, as an evaluation of the reported implementation outcomes, the degree of implementation success was summarized qualitatively based on the study authors' own definition. This review focused on the implementation of eHealth programs used by patients with chronic illness in their own homes. See Table 1 for details, key terms and definitions. The protocol for this realist systematic review has been registered and published in the Prospective Register of Systematic Reviews (PROSPERO; CRD42018085539).

Table 1. Key terms and their definitions.

Term	Definition
eHealth	Health services and information delivered or enhanced through the internet and related technologies [1]. Including but not limited to: • mHealth (mobile health): health practice supported by mobile devices [41] • Telehealth: using telecommunications and virtual technology to deliver health care outside of traditional health care facilities [42] • Patient portals (secure online websites that give patients access to personal health information) [43] For inclusion in this review, the eHealth program had to have patients/clients in their own homes as the primary users, optionally with support or involvement from health care providers. In this publication, the collective term eHealth is used unless a more specific definition is considered of essence.
Implementation	Process of putting to use or integrating evidence-based interventions within a setting [44].
Implementation strategy	Systematic intervention process to adopt and integrate evidence-based health innovations into usual care [19]. The Expert Recommendations for Implementing Change project has defined and sorted implementation strategies into a taxonomy consisting of the following categories: (1) engage consumers, (2) use evaluative and iterative strategies, (3) change infrastructure, (4) adapt and tailor to the context, (5) develop stakeholder interrelationships, (6) use financial strategies, (7) support clinicians, (8) provide interactive assistance, and (9) train and educate stakeholders [22].
Implementation outcome	Effects of deliberate and purposive actions to implement new treatments, practices, and services [37]. Proctor and colleagues [37] have defined and sorted implementation outcomes into the implementation outcome framework consisting of the following terms: (1) acceptability, (2) adoption, (3) appropriateness, (4) costs, (5) feasibility, (6) fidelity, (7) penetration, and (8) sustainability.

Literature Search

A systematic literature search was performed by the librarian (MØ) in the electronic databases MEDLINE, Embase, PsycInfo and Allied and Complementary Medicine Database (Ovid), Cumulative Index to Nursing and Allied Health Literature (EBSCOhost), Scopus, and Cochrane Library. The search terms were developed by the first author (CV) and the librarian (MØ) using a combination of keywords and database-specific headings and covered the period from January 1, 2006, to October 4, 2018. The starting point for the review period was set to the year of the first issue of the journal Implementation Science (2006), since there was a pronounced focus on implementation from that point, although some researchers had been working within this field earlier. The basic search strategy (Multimedia Appendix 1) was modified for use in each database. Additional studies were detected based on references and citations in the included studies.

Criteria for Considering Studies for the Review

Inclusion criteria for studies in the review were the reporting of implementation strategies used in the implementation of eHealth programs seeking to support adults with chronic illness in their own homes. Studies were included only if they provided a description of the implementation strategies they had used. Studies were, for example, excluded if they only mentioned training had been conducted or management had been involved without any further description of the content of the training or management engagement.

The following illnesses were included: chronic disease, arthritis, chronic pain, chronic obstructive pulmonary disease, obesity, diabetes mellitus, and mental disorder. Empirical studies in English, Dutch, and Scandinavian languages published in peer-reviewed journals were included. All study designs were included. Literature reviews, meta-analyses, theoretical articles, book chapters, editorials, study protocols, dissertations, studies published in abstract form only, and duplicates were excluded. eHealth programs involving primarily children, adolescents, and family care givers or solely for health care providers were excluded.

Study Selection Process

All titles and abstracts were reviewed by the first author (CV). Irrelevant publications (eg, studies focusing on non-eHealth programs) were excluded. Next, two of the authors (CV and one of the coauthors) independently reviewed titles and abstracts using the systematic review software Covidence (Veritas Health Innovation). When the authors were in agreement, the studies were included for full-text review. When the authors were not in agreement, the first author (CV) conducted a second review and subsequently made a decision. If there was doubt, the study was selected for full-text review. Next, CV and one of the coauthors independently reviewed full-text studies separately. When the authors agreed, the studies were included. If the authors disagreed, the first author conducted a second review and subsequently made a final decision. The authors met several times during this process in order to discuss and reach agreement on the understanding of the inclusion and exclusion criteria.

Data Extraction and Evidence Appraisal

Data were extracted using a data extraction form developed by the authors for the purpose of this review relating to the study details, country of origin, design, setting, population, demographics, intervention, implementation framework, implementation strategies, implementation outcomes, and implementation success. NVivo software version 11 (QSR International) was used to organize and facilitate the extraction. The data extraction was guided by the aims of the review, focusing on (1) implementation strategies used, (2) implementation outcomes achieved, and (3) the relationship between implementation strategies, implementation outcomes, and degree of implementation success. The identified implementation strategies were sorted according to the 9 categories defined by the Expert Recommendations for Implementing Change (ERIC) project [22]. See Table 2 for specific description of implementation strategies. The identified implementation outcomes were sorted by the 8 categories in the implementation outcome framework defined by Proctor and colleagues [37]. See Table 3 for specific description of implementation outcomes. The taxonomies of ERIC and Proctor have been successfully used by other researchers [45-47] and were used in this review. Implementation success was extracted according to the study authors' own evaluation of implementation success in relation to the implementation strategies used, not based on a specific framework. The data extraction was conducted in two steps. First, implementation strategies, implementation outcomes, and implementation success were extracted separately. Next, these 3 sets of data were put together in a table to evaluate their interrelationships (eg, qualitatively assessing whether certain combinations were more common than others). The first author (CV) extracted data from all included studies. A second author (SMK) validated the data extraction of 25% (3/12) of the included studies.

Traditional quality assessment of the included studies in this review was not undertaken. The realist review methodology does not lean on the traditional study hierarchy assessment with the randomized controlled trials at the top, as it is acknowledged that multiple methods are needed to cover the entire picture of what works for whom and under which circumstances [38]. The relevance of the included studies was considered based on each study's ability to answer the research questions of the review, including that the studies had provided at least a minimum description of the content of the implementation strategies used to be incorporated. Rigor was considered related to the study authors' credibility based on the conclusions made in the included studies.

Table 2. Implementation strategies (adapted from Waltz and colleagues [22]).

Implementation strategies	Description
Engage consumers	Involving, preparing, and intervening with patients and the market to involve them and increase demand for the clinical innovation
Use evaluative and iterative strategies	Planning and conducting the implementation process, including activities such as make a plan, assess for readiness, identify barriers and facilitators, evaluate performance and progress, and provide audit and feedback
Change infrastructure	Changing external structures such as legislation models, as well as internal conditions such as facilities and equipment
Adapt and tailor to the context	Tailoring the innovation to meet local needs and tailoring the implementation strategies toward the identified barriers and facilitators
Develop stakeholder interrelationships	Involving relevant internal and external stakeholders to support and move the implementation process forward
Use financial strategies	Changing the patient billing systems, fee structures, reimbursement policies, research funding, and clinician incentives
Support clinicians	Supporting clinical staff performance
Provide interactive assistance	Supporting implementation issues
Train and educate stakeholders	Providing written and oral training

Table 3. Implementation outcomes (adapted from Proctor and colleagues [37]).

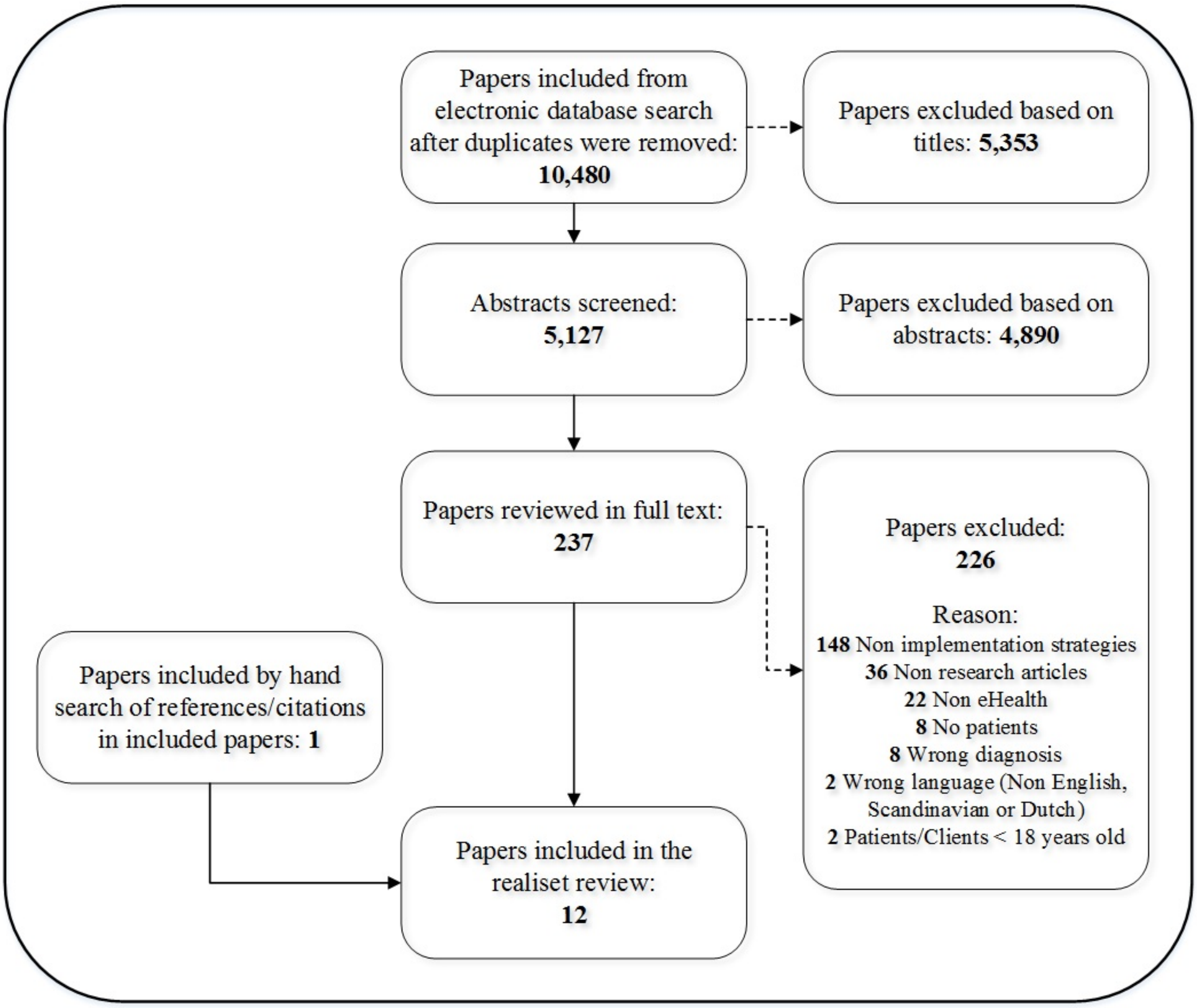
Implementation outcomes	Description
Acceptability	Perception that a given treatment, service, practice, or innovation is agreeable, palatable, or satisfactory
Adoption	Intention, initial decision, or action to try or employ an innovation or evidence-based practice
Appropriateness	Perceived fit, relevance, or compatibility of the innovation or evidence-based practice for a given practice setting, provider, or consumer and/or perceived fit of the innovation to address a particular issue or problem
Cost	Cost impact of an implementation effort (incremental or implementation cost)
Feasibility	Extent to which a new treatment or innovation can be successfully used or carried out within a given agency or setting
Fidelity	Degree to which an intervention was implemented as it was prescribed in the original protocol or intended by the program developers
Penetration	Integration of a practice within a service setting and its subsystems
Sustainability	Extent to which a newly implemented treatment is maintained or institutionalized within a service setting's ongoing, stable operations

Results

Overview of Included Studies

The search generated 10,480 unique references. From these references, 5353 were excluded based on the title alone and an additional 4890 were excluded based on the abstract. The inconsistency in terms used in the research literature on eHealth and implementation strategies led to a large number of hits on irrelevant studies. Most of these studies were therefore excluded, and 237 studies were selected for full text evaluation. Following evaluation by two independent authors (ie, the first author and one coauthor), 11 studies met all inclusion criteria and were included [48-58]. In addition, one study was included based on a manual search of references and citations in the first 11 included studies [59]. See Figure 1 for details on the study selection process.

Figure 1. Flow diagram of the study selection process.



Seven of the 12 included studies used qualitative research design [48,50,52,53,55,58,59], 2 used quantitative design in terms of surveys [51,56], and 3 used mixed-methods design [49,54,57]. Of the final 12 included studies, 2 studies were conducted in the United States [50,54], one in Canada [51], 5 in the United Kingdom [48,52,53,55,59], 2 in the Netherlands [49,56], one in Norway [58], and one in New Zealand [57]. All 12 were published in English.

Two of the 12 included studies were conducted in early phases of the implementation (ie, up to 3 months after implementation startup) [50,56]. Four studies were conducted 4 to 12 months after implementation startup, defined as middle phase [48,49,53,55]. The remaining 6 studies were conducted more than 1 year after implementation startup, defined as late phase [51,52,54,57-59], and 4 of these had multiple data collection time points [51,52,54,59].

eHealth Programs and Patient Groups

Of the 12 included studies, 5 targeted the use of online clinical monitoring programs including patient-provider communication [48,52,53,55,59]. Three studies targeted use of ICBT [51,56,58]. Two used video consultations [49,54], one studied the implementation of both video consultation and ICBT [50], and one targeted online personal health records [57].

Four of the 12 studies included patients with somatic illnesses (chronic obstructive pulmonary disease, chronic heart failure, and chronic pain) [48,53,55,56], 3 studies included patients with mental health challenges (anxiety and depression) [50,51,58], and 5 studies included patients with long-term illnesses in general [49,52,54,57,59].

Implementation Frameworks and Models

Of the total 12 studies, 8 used implementation frameworks or models to guide the analysis of implementation strategies and/or implementation outcomes. Two studies used the reach effectiveness, adoption, implementation, maintenance framework [50,54], and 2 studies used the normalization process theory [52,58]. Other frameworks/models were used by one study each: consolidated framework for implementation research [51], structurationism [49], promoting action on research implementation in health services [54] and the plan do study act cycle [55]. Finally, one study used the theoretical domains framework in combination with the technology acceptance model [56]. See Table 4 for details.

Table 4. Overview of included studies.

First author	Patient groups	eHealth	Setting	Implementation project	Implementation framework	Implementation stage	Study design	Data collection
Bailey [48]	COPD[a]	Clinical monitoring	Sheltered housing	4 tenants used telehealth for 16 weeks	N/A[b]	Middle[c]	Case study (QUAL[d])	Self-report assessment, observation, focus groups, interviews, workshops
Boonstra [49]	Long-term illnesses	Video consultation	Homecare	From a database of 11,000 regular customers in 2006, 36 used the system	Structurationism	Middle	Single case study (MIXED)	Interviews, workshops, written reports, policy plans, meeting minutes, observations, quantitative data on system use
Fortney [50]	Depression	ICBT[e], *Beating the Blues*, video consultation	Primary care safety net clinics	Implement EBP[f] in 6 federally qualified health centers	RE-AIM[g]	Early[h]	Quality improvement methods (QUAL)	Qualitative needs assessments
Hadjistavropoulos [51]	Anxiety, depression	ICBT, *Wellbeing Course*	Community mental health clinics	ICBT implementation in 7 community mental health clinics	CFIR[i]	Late[j]	Process evaluation (QUANT[k])	Online survey
Hendy [52]	Long-term illnesses	Clinical monitoring, *WSD*[l]	Primary care trusts	Case studies of 3 sites forming the WSD program	NPT[m]	Late	Comparative, longitudinal, qualitative, ethnographic case study (QUAL)	Interviews, meeting observations, document review
Hendy [59]	Long-term illnesses	Clinical monitoring, *WSD*	Health and social care organizations	Case studies representing 5 large public sector health organizations	N/A	Late	Longitudinal, ethnographic case studies (QUAL)	Observations, document review, informal discussions, interviews
Horton [53]	COPD	Clinical monitoring	Homecare	During the 6-month implementation period, only 10 users had been recruited to the scheme	N/A	Middle	Case study (QUAL)	Focus groups, field notes, meeting minutes
Lindsay [54]	PTSD[n], anxiety, depression, insomnia, chronic pain, SUD[o]	Video consultation, *Video to Home*	VA[p] Medical Center	This 2-year project included 93 patients	PARIHS[q], RE-AIM	Late	Mixed-method program evaluation (MIXED)	Interviews, quantitative data on system use
Taylor [55]	COPD, chronic HF[r]	Clinical monitoring	Community health care	4 community nursing settings involved in 7-month program of action research	PDSA[s]	Middle	Case studies and action research methodologies (QUAL)	Workshop observations, focus groups, document review, field notes

First author	Patient groups	eHealth	Setting	Implementation project	Implementation framework	Implementation stage	Study design	Data collection
Terpstra [56]	Chronic pain	ICBT, *Master Your Pain*	Mental health care institutions	13 mental health care institutions	TDF[t], TAM[u]	Early	Descriptive design (QUANT)	Evaluation questionnaire
Wells [57]	Chronic illness	Online PHR[v]	Health delivery organizations	Health care organizations that had had a PHR in place for at least 12 months	N/A	Late	Grounded theory inductive approach (MIXED)	Interviews, Web-based survey
Wilhelmsen [58]	Depression	ICBT, *MoodGYM*	General practice	3-day training package for GPs[w] on ICBT	NPT	Late	Qualitative study (QUAL)	Telephone interviews

[a] COPD: chronic obstructive pulmonary disease.

[b] N/A: not applicable.

[c] Middle: 4-12 months postimplementation startup.

[d] QUAL: qualitative.

[e] ICBT: internet-delivered cognitive behavioral therapy.

[f] EBP: evidence-based practice.

[g] RE-AIM: reach, effectiveness, adoption, implementation, maintenance framework.

[h] Early: 0-3 months postimplementation startup.

[i] CFIR: consolidated framework for implementation research.

[j] Late: >12 months postimplementation startup.

[k] QUANT: quantitative.

[l] WSD: Whole Systems Demonstrator.

[m] NPT: normalization process theory.

[n] PTSD: posttraumatic stress disorder.

[o] SUD: substance use disorder.

[p] VA: Veterans Affairs.

[q] PARIHS: promoting action on research implementation in health services.

[r] HF: heart failure.

[s] PDSA: plan, do, study, act.

[t] TDF: theoretical domains framework.

[u] TAM: technology acceptance model.

[v] PHR: patient health record.

[w] GP: general practitioner.

Implementation Strategies Reported

Overview

Nine of the 12 included studies reported the use of an overarching implementation strategy such as training [48,56,58], external facilitation [50,51,54], managerial strategies [59], action research [55], or a mixture of several discrete strategies [57]. Three studies did not describe any overarching implementation strategy, only describing the discrete strategies used [49,52,53].

When sorted according to the ERIC categories [22], 5 of the 12 studies reported implementation strategies within 7 or 8 categories [49-51,55,57], 2 reported implementation strategies within 5 or 6 categories [52,54], 2 reported implementation strategies within 3 or 4 categories [48,59], and 3 reported implementation strategies within 1 or 2 categories [53,56,58].

The category of implementation strategies most frequently reported was train and educate stakeholders (n=10), followed by change infrastructure (n=8), develop stakeholder interrelationships (n=8), use evaluative and iterative strategies (n=7), engage consumers (n=6), adapt and tailor to the context (n=5), use financial strategies (n=5), support clinicians (n=5), and finally provide interactive assistance (n=4). See Table 5 for details.

Table 5. Categories of implementation strategies [22] used in the included studies.

Studies	Engage consumers	Use evaluative and iterative strategies	Change infrastructure	Adapt and tailor to the context	Develop stakeholder interrelationships	Use financial strategies	Support clinicians	Provide interactive assistance	Train and educate stakeholders	Total categories reported	Overarching implementation strategy (study authors' description)
Bailey [48]	x		x						x	3	Training
Boonstra [49]	x	x	x	x	x	x	x			7	Not reported
Fortney [50]	x	x	x	x	x			x	x	7	External facilitation/mixed
Hadjistavropoulos [51]	x	x			x	x	x	x	x	7	External facilitation
Hendy [52]		x	x		x		x		x	5	Not reported
Hendy [59]			x		x	x		x		4	Managerial strategies
Horton [53]			x						x	2	Not reported
Lindsay [54]		x	x	x	x			x	x	6	External facilitation
Taylor [55]	x	x	x	x	x	x	x		x	8	Action research
Terpstra [56]									x	1	Training
Wells [57]	x	x		x	x	x	x		x	7	Mixed
Wilhelmsen [58]									x	1	Training
Total	6	7	8	5	8	5	5	4	10		

Engage Consumers

Six of the 12 studies reported trying to reach and engage patients as one of their implementation strategies. This included advertising about the eHealth program to patients within their own institutions and/or to the wider community by means of newsletters, webpages, television, radio, newspapers, and direct contact with patients [49-51,57]. Other implementation strategies reported used to engage patients were inclusion of patients in research activities [55] and technical setup and support for patients in their homes [48,49].

Use Evaluative and Iterative Strategies

Seven of the 12 studies reported use of different evaluative and iterative strategies, either as stand-alone strategies or in combination with other strategies. Three of the 12 studies had made an implementation plan [50,55,57], 1 study had a business plan [49], and 1 study had included the eHealth implementation in the organizational vision statement [57]. Three studies focused on readiness, barriers, and facilitators [50,51,54]. Five studies reported that they made use of multiple stakeholder teams [50-52,55,57], and 1 study reported support from local clinical champions [50]. Five of the 12 studies reviewed the implementation progress [51,52,54,55,57], and 4 of them provided audit and feedback by feeding the information about the implementation progress back to the clinicians [51,54,55,57].

Change Infrastructure

Eight of the 12 studies reported purchase or acquisition of new electronic equipment as an implementation strategy [48-50,52-55,59].

Adapt and Tailor to the Context

Four of the 12 studies reported that they had cooperated with clinical staff to ensure tailoring of the eHealth program to meet local needs and organizational capabilities [50,54,55,57]. One study had cooperated with involved stakeholders to obtain a consistent implementation plan [49].

Develop Stakeholder Interrelationships

Four of the 12 studies reported involving multiple stakeholder teams at the overall management level, including representatives of the participating organizations such as care delivery organizations, telecom firms, insurance firms, commissioners, and industry [49,51,55,57]. Onsite project teams were established in 4 of the 12 studies [49,52,55,57]. Onsite clinical champions supported and promoted adoption of the eHealth program in 5 studies [50,52,54,55,57]. Management support and endorsement were reported in 3 studies [52,57,59]. One study also had visited other clinics to discuss concerns and impart their knowledge and experience [57].

Use Financial Strategies

Five of the 12 studies reported that they had used financial strategies related to the funding of the implementation projects [49,51,52], future cost-effectiveness aspects [49], and future financial investment aspects [55]. Incentives directed toward physicians' performance indicators and monetary incentives and the use of gift card bonuses for clinicians were reported [57].

Support Clinicians

Four of the 12 studies had supported clinicians by recruiting new staff, establishing new roles, and supporting work process redesign [49,52,55,57]. Reminders to clinicians to prompt them to use the new eHealth programs were also reported [51].

Provide Interactive Assistance

Four of the 12 studies reported that they had used external researchers, consultants, or practitioners to provide external facilitation in terms of problem solving and support [50-52,54]. Training for local superusers was also reported conducted [54].

Train and Educate Stakeholders

Ten of the 12 studies reported that they had conducted training and teaching for clinicians about the eHealth programs being implemented. The education was reported as containing aspects related to the delivery of the clinical programs via eHealth [48,50-58], as well as technical aspects related to the eHealth software [48,51,57]. Six studies reported on the length of training and described a wide variety of time span, ranging from 2 to 3 hours [48,50,53] to 1 to 3 days [51,56,58].

Implementation Outcomes Reported

Overview

All the 12 included studies reported implementation outcomes, ranging from 1 to 6 in each study. The 3 most frequently reported were acceptability, penetration, and adoption. See Table 6 for details on implementation strategies used and implementation outcomes reported.

Table 6. Implementation strategies used and implementation outcomes reported in the included studies.

First author	Implementation strategies		Implementation outcomes								Implementation success	
	Categories of implementation strategies used	n	Acceptability	Adoption	Appropriateness	Cost	Feasibility	Fidelity	Penetration	Sustainability	n	Study authors' evaluation of implementation success in relation to implementation strategies used
Bailey [48]	Engage consumers, change infrastructure, train and educate stakeholders	3	+/–[a]	+/–	N/A[b]	N/A	N/A	N/A	N/A	N/A	2	Successful due to training and follow-up support
Boonstra [49]	Engage consumers, use evaluative and iterative strategies, change infrastructure, adapt and tailor to the context, use financial strategies, support clinicians, train and educate stakeholders	7	+/–	–[c]	–	N/A	–	N/A	–	N/A	5	Unsuccessful due to limited managerial agency and inconsistencies in some of the choices made during implementation phase
Fortney [50]	Engage consumers, use evaluative and iterative strategies, change infrastructure, adapt and tailor to the context, provide interactive assistance, train and educate stakeholders	7	N/A	N/A	N/A	N/A	N/A	N/A	+/–	+[d]	2	Variable success across sites
Hadjistavropoulos [51]	Engage consumers, use evaluative and iterative strategies, develop stakeholder interrelationship, use financial strategies, support clinicians, provide interactive assistance, train and educate stakeholders	7	+	–	N/A	+/–	–	N/A	–	N/A	5	Successful due to ICBT[e] program, implementation processes, and external facilitation. Could have been even better if planned in advance, all staff in the health region were informed about ICBT, and more resources were available
Hendy [52]	Use evaluative and iterative strategies, change infrastructure, develop stakeholder interrelationship, support clinicians, train and educate stakeholders	5	N/A	–	N/A	N/A	N/A	N/A	+/–	–	3	Unsuccessful despite resources deployed
Hendy [59]	Change infrastructure, develop stakeholder interrelationship, use financial strategies, provide interactive assistance	4	N/A	N/A	N/A	N/A	–	N/A	N/A	N/A	1	Unsuccessful due to lack of trust in individual managers
Horton [53]	Change infrastructure, train and educate stakeholders	2	–	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1	Unsuccessful despite training and follow-up support

First author	Implementation strategies		Implementation outcomes								Implementation success	
	Categories of implementation strategies used	n	Acceptability	Adoption	Appropriateness	Cost	Feasibility	Fidelity	Penetration	Sustainability	n	Study authors' evaluation of implementation success in relation to implementation strategies used
Lindsay [54]	Use evaluative and iterative strategies, change infrastructure, adapt and tailor to the context, develop stakeholder interrelationship, provide interactive assistance, train and educate stakeholders	6	+	N/A	N/A	+	–	–	+	+	6	Successful due to implementation facilitation strategy involving external and internal facilitators, especially clinical champions and training
Taylor [55]	Engage consumers, use evaluative and iterative strategies, change infrastructure, adapt and tailor to the context, develop stakeholder inter-relationship, use financial strategies, support clinicians, train and educate stakeholders	8	N/A	N/A	N/A	+/–	N/A	N/A	N/A	N/A	1	Mixed: 2 sites discontinued after first cycle because of competing priorities; positive experience of external facilitation by researchers and telehealth champions
Terpstra [56]	Train and educate stakeholders	1	+	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1	N/A
Wells [57]	Engage consumers, use evaluative and iterative strategies, adapt and tailor to the context, develop stakeholder interrelationship, use financial strategies, support clinicians, train and educate stakeholders	7	+	N/A	N/A	N/A	N/A	N/A	+/–	N/A	2	Successful organizations actively communicated their vision; engaged leaders at all levels; had clear governance, planning, and protocols; set targets; and celebrated achievement. The most effective strategy for patient uptake was through health professional encouragement
Wilhelmsen [58]	Train and educate stakeholders	1	+	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1	Not successful due to lack of practical training of module follow-ups in the course
Total			8	4	1	3	4	1	6	3		

[a] Mixed/neutral outcomes.

[b] Not applicable.

[c] Negative outcomes.

[d] Positive outcomes.

[e] ICBT: internet-delivered cognitive behavioral therapy.

Acceptability

Four of the 12 studies reported that health care providers had shown positive attitudes toward the eHealth program implemented [54,56-58]. One study reported low acceptability of their intervention [53]. Three studies reported mixed attitudes in that some were positive and some experienced the new eHealth program as a threat to or disturbance of their work [48,49,51].

Adoption

Four of the 12 studies reported challenges regarding the adoption of the eHealth programs into their clinical practice, describing difficulties motivating the clinicians to approach their clients with the new eHealth program [49,51-53]. Time available and time frame given were also reported to pose organizational challenges [48,52]. None of the included studies presented solely positive descriptions of the adoption of the eHealth programs.

Appropriateness

Only 1/12 included studies reported on appropriateness, stating that the technology might not always be appropriate, for example, if advanced age, poverty, or serious illnesses might amplify the clients’ vulnerability [49].

Cost

Three studies mentioned costs. One study reported no additional costs related to the eHealth implementation [51], 1 study reported travel expenditures saved [54], and 1 study described being unable to calculate costs due to lack of robust data [55].

Feasibility

Four of the 12 studies reported low feasibility for their eHealth programs [49,51,52,54], describing the innovations as an interruption to the real work and as difficult to integrate with existing patient workloads.

Fidelity

Fidelity was reported in only 1 of the 12 studies, stating that high fidelity was difficult to achieve due to providers’ need to remain flexible and the program needed to be adapted to the technology platform already present in the clinical setting [54].

Penetration

Four studies presented how many patients received an eHealth program [50,52,54,57], only one of which reported satisfaction with how many patients received the program [54]. Two studies indicated limited numbers of patients who received the eHealth program being studied, but did not provide exact figures [49,51].

Sustainability

Three of the 12 included studies reported sustainability. In two instances, the eHealth programs were sustained after the implementation efforts [50,54], while the third did not achieve sustainability [52].

Implementation Success Reported

All studies except one [56] reported on implementation success. The majority provided a direct [48,50,52,57-59] or indirect [49,53,54] description of how they defined implementation success. This spanned from concrete definitions such as “the number of people in each site using the new service” [59] to more vague descriptions such as “change in terms of telecare appropriation was realized” [49]. As the assessment of implementation success was used as a means to evaluate the reported implementation outcomes in this review, and implementation success is often derived directly from the implementation outcomes, the two aspects (ie, success and outcome) were not necessarily mutually exclusive. Four studies reported that the implementation had been successful [48,51,54,57], while 5 studies reported unsuccessful implementation [49,52,53,58,59]. Two studies reported mixed results, with implementation being successful at some of the sites and unsuccessful at the others [50,55].

Relationship Between Implementation Strategies, Implementation Outcomes, and Implementation Success

In the 12 included studies, no relationship was detected between implementation strategies [22] and implementation outcomes [37].

Regarding implementation success, the implementation strategies management support and engagement, internal and external facilitation, training, and audit and feedback were directly related to implementation success in several studies. For example, management support and engagement were highlighted as important for implementation success in 1 study [57], and lack of trust or limited managerial agency was described as a contributing factor to implementation failure in 2 other studies [49,59]. Furthermore, external facilitation was reported to be important for implementation success in 4 studies [50,51,54,55]. Internal facilitation, especially the support and engagement of clinical or implementation champions, was highlighted as important for the implementation success in 2 studies [54,55]. In addition, training and education of stakeholders were used as implementation strategies in studies reporting successful [48,54,56] as well as unsuccessful implementation [53,58].

No clear relationship was found between the number of implementation strategies used and implementation success. For example, of 3 studies using a range of implementation strategies, 1 reported implementation success [51], 1 reported implementation failure [49], and 1 reported mixed results [55]. Furthermore, of 2 studies using training and education of stakeholders as the only implementation strategy, 1 reported implementation success [56] and 1 reported implementation failure [58]. There was no relationship between reported implementation success and use of implementation frameworks.

Discussion

Summary of Evidence and Comparison With Prior Work

This systematic realist review used the categorization of implementation strategies by the ERIC taxonomy [21,22] and the implementation outcome framework by Proctor and colleagues [37] as data extraction templates. As no specific framework exists for implementation success, this was qualitatively summarized based on the study authors’ own definition. The review identified and synthesized 12 studies examining implementation strategies, implementation outcomes, and implementation success related to the implementation of eHealth programs for patients with chronic illnesses. Findings show that there has so far been little focus on reporting implementation strategies for eHealth implementation where the patient is the main user of the program. Also, there appears to be great variety in implementation terms used and considerable vagueness in the description of implementation aspects, which led the authors to have to screen a number of irrelevant full-text studies. There were also challenges in the data extraction process due to inconsistence in terminology used in the studies. Other researchers have also pointed to inconsistencies in use of terminology and definitions related to implementation [20,60,61]. Due to great heterogeneity in the included studies with regard to types of patient conditions, eHealth interventions, and phases of implementation, it was not possible to detect any relationship between these factors related to implementation strategies, implementation outcomes, and implementation success.

A wide range of implementation strategies were used in the studies included in the review. The most frequently used categories of implementation strategies were train and educate stakeholders, change infrastructure, and develop stakeholder interrelationships. Included in the latter category is involvement of champions, which has also been identified as central to implementation success by other reviews [62,63]. Several of the included studies reported training of health care personnel as a preferred implementation strategy, and this strategy was also found to be widely used by others, even though effects appear inconsistent [62,64]. Despite recent evidence pointing to tailored implementations as effective [62,65], only 4 studies in the review reported that they had tailored the eHealth intervention to meet the context where the implementation took place. Also, several frameworks for technology implementation have pointed to the importance of contextual factors as key elements to address in order to succeed, including the CeHRes (Center for eHealth Research and Disease Management) roadmap [66] and the NASSS framework [28]. The limited use of tailoring so far in the implementation context could potentially be one explanation for the limited implementation success to date.

Implementation outcomes were reported in all 12 studies included in this review, with each individual study reporting between 1 and 6 implementation outcomes. The implementation outcomes most frequently reported were acceptability and penetration. As the included studies had not aimed to report on implementation outcomes, only a few of the terms in the implementation outcome framework [37] were covered. It is thus reasonable to assume that implementation outcomes were underreported in many of the included studies. Based on this, it was not possible to detect any clear relationship between implementation strategies and implementation outcomes in the review. However, it might not be a coincidence that these 12 studies that reported implementation strategies also reported implementation outcomes. Because when people really start to think about and report implementation strategies, they will also think about reporting at least some implementation outcomes. In order to still allow for an evaluation of how successful the implementation had been when the implementation outcome framework was not suitable enough for a mechanism evaluation, implementation success was also included in this equation.

Regarding implementation success, 4 of the included 12 studies reported success, 5 reported lack of success, and 2 reported mixed results. Training and education of stakeholders showed mixed relations to implementation success, indicating that the content, duration, and facilitation of the training are important for training effectiveness. The studies offering the most training are not necessarily the most successful, indicating that other factors (eg, clinician motivation and intention to use the new eHealth program) also play an important role [4]. This review suggests that a combination of software training and training in how to use the technology in daily work may be necessary. These findings are in line with other reviews that have also highlighted training, support, and supervision as key factors in order for clinicians to start using new eHealth programs [30,35]. Due to the limited coverage provided by the implementation outcome framework, as described above, no clear relationship between implementation outcomes and implementation success could be detected in the review. For example, one of the studies showed that the implementation can be successful or experienced as successful even with negative scores on some of the implementation outcomes concepts [54]. However, in more than half of the studies in the review, there was coherence between the ratings on implementation outcomes and implementation success [49,50,52,53,55,57,59]. Due to the limited number of implementation outcome concepts covered, however, this finding must be interpreted with caution. Given a more comprehensive reporting on implementation outcomes, the coherence could have been different. The relationship between implementation outcomes and implementation success still appears a conundrum. This has also been pointed out by others [37] and should be further investigated in future studies. Although not the topic of this review, it is also worth mentioning that if the patient outcomes (eg, effect of the intervention) do not occur, positive implementation outcomes and implementation success does not have much impact.

Another important finding from the review is that several studies showed the implementation strategies related to management engagement to be directly related to implementation success. Other researchers have found leadership to be crucial in order to succeed with implementation of evidence-based practice and have also pointed to the setting in which the leader operates as being of importance [67].

The successful implementation efforts identified in this review, reaching sustainability for more than 1 year after start-up [51,54,57], were all related to use of a mixture of several implementation strategies and were also supported by internal and external facilitation. All of these studies also provided audit and feedback, one of the implementation strategies with evidence for effectiveness [62,68].

No clear relationship was found in the review between the number of implementation strategies used and implementation success. The successful implementation projects described used multifaceted strategies. However, one study used a single strategy and was still successful [56]. This shows that the quality of an implementation strategy might be more important than the quantity, which is in line with a former review concluding that multifaceted strategies are not necessarily more effective than single strategies [64].

Despite the importance of describing and sharing information about unsuccessful implementations, the continued degree of unsuccessful implementation efforts is disturbing and gives cause for concern. It is, however, possible that the lack of a systematic implementation approach and the lack of employing proposed successful implementation strategies can provide explanation for this challenge.

Finally, the results from this review also indicate that reaching sustainability is and remains a challenge despite use and focus on implementation strategies.

Implications for Research and Practice

This systematic realist review clearly demonstrates a need for more studies that report on implementation strategies, implementation outcomes, implementation success, and the relationship among these in eHealth implementation. The research on implementation strategies is still in its infancy, and more work is needed to better understand how implementation strategies can contribute to improved implementation effectiveness [23].

This review also demonstrates the need for implementation planning at a very early stage—that is, already in the design and development phase of eHealth support and intervention programs. Low feasibility of many of the eHealth programs included in this review clearly shows an urgent need to include all stakeholders in the early phases of program development. Also, implementation planning must be included from the very beginning in order to adapt interventions to context and enable implementation. As such, using frameworks for eHealth development and implementation, such as the CehRes roadmap [66] that combines aspects from human-centered design, persuasive technology, and business modeling, can help address implementation aspects already in the phase of idea generation and problem identification.

When planning and conducting eHealth implementation in clinical practice, evidence is still lacking about proposing clear advice on how implementation strategies can be used effectively when implementing eHealth programs to support patients in their own homes. This review concludes, in support of existing research, that the question of which implementation strategies are the most effective under which circumstances still remains unclear [64]. Nonetheless, this review indicates that internal and external facilitation, audit and feedback, management support, and training of clinicians are essential. Lacking more robust evidence on specific implementation strategies for eHealth implementation, general evidence on implementation strategies must be considered.

Limitations

This systematic realist review has limitations that need to be considered when interpreting the results. First, in order to get a manageable number of hits from the literature search, some limitations to the search strategy were necessary. Therefore, the search was performed on published studies only since 2006. Prior to 2006, the eHealth and implementation research fields were both in their infancy and few publications were assumed to exist. This review process showed the earliest publication included to be from 2008, supporting this assumption. Therefore, no publications were included from the period 2006 to 2008. Another restriction intended to keep the hits to a manageable number was to limit the chronic illnesses included.

Use of predefined categories for data abstraction and analysis has strengths as well as limitations. In the review, the ERIC project [22] and the implementation outcome framework [37] were used to guide the review process. There is a potential risk that aspects not covered in the two categorizations could be overlooked in the review, as different frameworks provide different lenses through which research problems can be analyzed [69]. The ERIC categories are comprehensive and posed some challenges regarding overlap between categories. Furthermore, as not all included studies had implementation aspects as their only focus, the data extraction process could have introduced potential risks of overlooking or omitting aspects of implementation strategies, implementation outcomes, and implementation success. Inconsistent use of language and terminology in the 12 included studies also made it challenging to sort and label implementation strategies and outcomes. The validation process conducted by two authors nevertheless showed no discrepancy in data extraction.

Conclusions

This is the first review examining implementation strategies, implementation outcomes, and implementation success of studies reporting on the implementation of eHealth programs for patients with chronic illnesses. Findings suggest that internal and external facilitation, management support, and training of clinicians are important factors for the success of eHealth implementation. The results also highlight the lack of eHealth studies reporting implementation strategies in a comprehensive way, pointing to the need for designing robust studies on implementation strategies in the future.

Acknowledgments

This realist systematic review was supported by the Research Council of Norway grant #256574 (principal investigator LSN).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[PDF File (Adobe PDF File)138 KB - jmir_v21i9e14255_app1.pdf]

References

1. Eysenbach G. What is eHealth? J Med Internet Res 2001 Jun;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
2. 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]
3. Mehta S, Peynenburg VA, Hadjistavropoulos HD. Internet-delivered cognitive behaviour therapy for chronic health conditions: a systematic review and meta-analysis. J Behav Med 2018 Nov 01;42(2):169-187. [doi: 10.1007/s10865-018-9984-x] [Medline: 30387008]
4. Varsi C, Ekstedt M, Gammon D, Ruland CM. Using the consolidated framework for implementation research to identify barriers and facilitators for the implementation of an internet-based patient-provider communication service in five settings: a qualitative study. J Med Internet Res 2015 Nov 18;17(11):e262 [FREE Full text] [doi: 10.2196/jmir.5091] [Medline: 26582138]
5. Borosund E, Varsi C, Clark M, Ehlers S, Andrykowski M, Sleveland H. Pilot testing an app-based stress management intervention for cancer survivors [in press]. Translat Behav Med 2019:1. [doi: 10.1093/tbm/ibz062]
6. Harerimana B, Forchuk C, O'Regan T. The use of technology for mental healthcare delivery among older adults with depressive symptoms: a systematic literature review. Int J Ment Health Nurs 2019 Jun;28(3):657-670. [doi: 10.1111/inm.12571] [Medline: 30666762]
7. Kruse CS, Beane A. Health information technology continues to show positive effect on medical outcomes: systematic review. J Med Internet Res 2018 Feb 05;20(2):e41 [FREE Full text] [doi: 10.2196/jmir.8793] [Medline: 29402759]
8. Mold F, de Lusignin S, Sheikh A, Majeed A, Wyatt JC, Quinn T, et al. Patients' online access to their electronic health records and linked online services: a systematic review in primary care. Br J Gen Pract 2015 Mar;65(632):e141-e151 [FREE Full text] [doi: 10.3399/bjgp15X683941] [Medline: 25733435]
9. 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]
10. Hanlon P, Daines L, Campbell C, McKinstry B, Weller D, Pinnock H. Telehealth interventions to support self-management of long-term conditions: a systematic metareview of diabetes, heart failure, asthma, chronic obstructive pulmonary disease, and cancer. J Med Internet Res 2017 May 17;19(5):e172 [FREE Full text] [doi: 10.2196/jmir.6688] [Medline: 28526671]
11. Brettle A, Brown T, Hardiker N, Radcliffe J, Smith C. Telehealth: the effects on clinical outcomes, cost effectiveness and the patient experience: a systematic overview of the literature. Manchester: University of Salford; 2013. URL: http://usir.salford.ac.uk/id/eprint/29392/4/Telehealth_v8_.pdf [accessed 2019-09-13]
12. Deady M, Choi I, Calvo RA, Glozier N, Christensen H, Harvey SB. eHealth interventions for the prevention of depression and anxiety in the general population: a systematic review and meta-analysis. BMC Psychiatry 2017 Aug 29;17(1):310 [FREE Full text] [doi: 10.1186/s12888-017-1473-1] [Medline: 28851342]
13. Thurnheer SE, Gravestock I, Pichierri G, Steurer J, Burgstaller JM. Benefits of mobile apps in pain management: systematic review. JMIR Mhealth Uhealth 2018 Oct 22;6(10):e11231 [FREE Full text] [doi: 10.2196/11231] [Medline: 30348633]
14. Turgoose D, Ashwick R, Murphy D. Systematic review of lessons learned from delivering tele-therapy to veterans with post-traumatic stress disorder. J Telemed Telecare 2018 Oct;24(9):575-585. [doi: 10.1177/1357633X17730443] [Medline: 28958211]
15. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]
16. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471 [FREE Full text] [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
17. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. J Med Internet Res 2018 Dec 17;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]
18. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
19. Powell BJ, McMillen JC, Proctor EK, Carpenter CR, Griffey RT, Bunger AC, et al. A compilation of strategies for implementing clinical innovations in health and mental health. Med Care Res Rev 2012 Apr;69(2):123-157 [FREE Full text] [doi: 10.1177/1077558711430690] [Medline: 22203646]

20. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. Implement Sci 2013 Dec 01;8:139 [FREE Full text] [doi: 10.1186/1748-5908-8-139] [Medline: 24289295]
21. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015 Feb 12;10:21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
22. Waltz TJ, Powell BJ, Matthieu MM, Damschroder LJ, Chinman MJ, Smith JL, et al. Use of concept mapping to characterize relationships among implementation strategies and assess their feasibility and importance: results from the Expert Recommendations for Implementing Change (ERIC) study. Implement Sci 2015 Aug 07;10:109 [FREE Full text] [doi: 10.1186/s13012-015-0295-0] [Medline: 26249843]
23. Powell BJ, Fernandez ME, Williams NJ, Aarons GA, Beidas RS, Lewis CC, et al. Enhancing the impact of implementation strategies in healthcare: a research agenda. Front Public Health 2019;7:3 [FREE Full text] [doi: 10.3389/fpubh.2019.00003] [Medline: 30723713]
24. Wensing M, Grol R. Knowledge translation in health: how implementation science could contribute more. BMC Med 2019 May 07;17(1):88 [FREE Full text] [doi: 10.1186/s12916-019-1322-9] [Medline: 31064388]
25. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015;10:53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
26. Ritchie M, Dollar K, Miller C, Oliver K, Smith J, Lindsay J. Quality Enhancement Research Initiative (QUERI) for Team-Based Behavioral Health. 2017. Using implementation facilitation to improve care in the Veterans Health Administration (version 2) URL: https://www.queri.research.va.gov/tools/implementation/Facilitation-Manual.pdf [accessed 2019-09-13]
27. Bucknall T, Rycroft-Malone J. Evidence-based practice: doing the right thing for patients. In: Bucknall T, Rycroft-Malone J, editors. Models and Frameworks for Implementing Evidence-Based Practice: Linking Evidence to Action. Oxford: Wiley-Blackwell; 2010:1-21.
28. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
29. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Analysing the role of complexity in explaining the fortunes of technology programmes: empirical application of the NASSS framework. BMC Med 2018 May 14;16(1):66 [FREE Full text] [doi: 10.1186/s12916-018-1050-6] [Medline: 29754584]
30. Drozd F, Vaskinn L, Bergsund HB, Haga SM, Slinning K, Bjørkli CA. The implementation of internet interventions for depression: a scoping review. J Med Internet Res 2016 Sep 08;18(9):e236 [FREE Full text] [doi: 10.2196/jmir.5670] [Medline: 27608548]
31. Hage E, Roo JP, van Offenbeek MA, Boonstra A. Implementation factors and their effect on e-Health service adoption in rural communities: a systematic literature review. BMC Health Serv Res 2013;13:19 [FREE Full text] [doi: 10.1186/1472-6963-13-19] [Medline: 23311452]
32. Scott KC, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
33. Mair FS, May C, O'Donnell C, Finch T, Sullivan F, Murray E. Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review. Bull World Health Organ 2012 May 1;90(5):357-364 [FREE Full text] [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
34. Saliba V, Legido-Quigley H, Hallik R, Aaviksoo A, Car J, McKee M. Telemedicine across borders: a systematic review of factors that hinder or support implementation. Int J Med Inform 2012 Dec;81(12):793-809. [doi: 10.1016/j.ijmedinf.2012.08.003] [Medline: 22975018]
35. Brewster L, Mountain G, Wessels B, Kelly C, Hawley M. Factors affecting front line staff acceptance of telehealth technologies: a mixed-method systematic review. J Adv Nurs 2014 Jan;70(1):21-33. [doi: 10.1111/jan.12196] [Medline: 23786584]
36. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
37. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
38. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review: a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy 2005 Jul;10 Suppl 1:21-34. [doi: 10.1258/1355819054308530] [Medline: 16053581]
39. Rycroft-Malone J, McCormack B, Hutchinson AM, DeCorby K, Bucknall TK, Kent B, et al. Realist synthesis: illustrating the method for implementation research. Implement Sci 2012 Apr 19;7:33 [FREE Full text] [doi: 10.1186/1748-5908-7-33] [Medline: 22515663]

40. Gough D. Meta-narrative and realist reviews: guidance, rules, publication standards and quality appraisal. BMC Med 2013 Jan 29;11:22 [FREE Full text] [doi: 10.1186/1741-7015-11-22] [Medline: 23360691]
41. World Health Organization. Global Observatory for eHealth Series. 2011. mHealth: new horizons for health through mobile technologies: based on the findings of the Second Global Survey on eHealth. Volume 3 URL: http://whqlibdoc.who.int/publications/2011/9789241564250_eng.pdf [accessed 2019-09-13]
42. World Health Organization. 2019. Telehealth URL: https://www.who.int/sustainable-development/health-sector/strategies/telehealth/en [accessed 2019-04-26] [WebCite Cache ID 77vEC0b0i]
43. What is a patient portal?. Washington: The Office of the National Coordinator for Health Information Technology; 2019. URL: https://www.healthit.gov/faq/what-patient-portal [accessed 2019-04-26] [WebCite Cache ID 77vEPCVK7]
44. Rabin BA, Brownson RC, Haire-Joshu D, Kreuter MW, Weaver NL. A glossary for dissemination and implementation research in health. J Public Health Manag Pract 2008;14(2):117-123. [doi: 10.1097/01.PHH.0000311888.06252.bb] [Medline: 18287916]
45. Huynh AK, Hamilton AB, Farmer MM, Bean-Mayberry B, Stirman SW, Moin T, et al. A pragmatic approach to guide implementation evaluation research: strategy mapping for complex interventions. Front Public Health 2018;6:134 [FREE Full text] [doi: 10.3389/fpubh.2018.00134] [Medline: 29868542]
46. Nielsen B, Slinning K, Weie Oddli H, Drozd F. Identification of implementation strategies used for the circle of security-Virginia family model intervention: concept mapping study. JMIR Res Protoc 2018 Jun 14;7(6):e10312 [FREE Full text] [doi: 10.2196/10312] [Medline: 29903703]
47. Perry CK, Damschroder LJ, Hemler JR, Woodson TT, Ono SS, Cohen DJ. Specifying and comparing implementation strategies across seven large implementation interventions: a practical application of theory. Implement Sci 2019 Dec 21;14(1):32 [FREE Full text] [doi: 10.1186/s13012-019-0876-4] [Medline: 30898133]
48. Bailey C, Cook G, Herman L, McMillan C, Rose J, Marston R, et al. Deploying telehealth with sheltered housing tenants living with COPD: a qualitative case study. Housing, Care and Support 2015 Dec 21;18(3/4):136-148. [doi: 10.1108/HCS-09-2015-0015]
49. Boonstra A, Van Offenbeek M. Towards consistent modes of e-health implementation: structurational analysis of a telecare programme's limited success. Info Systems J 2010;20:1. [doi: 10.1111/j.1365-2575.2010.00358.x]
50. Fortney JC, Pyne JM, Ward-Jones S, Bennett IM, Diehl J, Farris K, et al. Implementation of evidence-based practices for complex mood disorders in primary care safety net clinics. Fam Syst Health 2018 Dec;36(3):267-280. [doi: 10.1037/fsh0000357] [Medline: 29809039]
51. Hadjistavropoulos HD, Nugent MM, Dirkse D, Pugh N. Implementation of internet-delivered cognitive behavior therapy within community mental health clinics: a process evaluation using the consolidated framework for implementation research. BMC Psychiatry 2017 Sep 12;17(1):331 [FREE Full text] [doi: 10.1186/s12888-017-1496-7] [Medline: 28899365]
52. Hendy J, Chrysanthaki T, Barlow J, Knapp M, Rogers A, Sanders C, et al. An organisational analysis of the implementation of telecare and telehealth: the whole systems demonstrator. BMC Health Serv Res 2012;12:403 [FREE Full text] [doi: 10.1186/1472-6963-12-403] [Medline: 23153014]
53. Horton K. The use of telecare for people with chronic obstructive pulmonary disease: implications for management. J Nurs Manag 2008 Mar;16(2):173-180. [doi: 10.1111/j.1365-2834.2008.00845.x] [Medline: 18269548]
54. Lindsay JA, Hudson S, Martin L, Hogan JB, Nessim M, Graves L, et al. Implementing video to home to increase access to evidence-based psychotherapy for rural veterans. J Technol Behav Sci 2017 Nov 28;2(3-4):140-148. [doi: 10.1007/s41347-017-0032-4]
55. Taylor J, Coates E, Wessels B, Mountain G, Hawley MS. Implementing solutions to improve and expand telehealth adoption: participatory action research in four community healthcare settings. BMC Health Serv Res 2015 Dec 01;15:529 [FREE Full text] [doi: 10.1186/s12913-015-1195-3] [Medline: 26626564]
56. Terpstra JA, van der Vaart R, Spillekom-van Koulil S, van Dam A, Rosmalen JGM, Knoop H, et al. Becoming an eCoach: training therapists in online cognitive-behavioral therapy for chronic pain. Patient Educ Couns 2018 Dec;101(9):1702-1707. [doi: 10.1016/j.pec.2018.03.029] [Medline: 29661704]
57. Wells S, Rozenblum R, Park A, Dunn M, Bates DW. Organizational strategies for promoting patient and provider uptake of personal health records. J Am Med Inform Assoc 2015 Jan;22(1):213-222 [FREE Full text] [doi: 10.1136/amiajnl-2014-003055] [Medline: 25326601]
58. Wilhelmsen M, Høifødt RS, Kolstrup N, Waterloo K, Eisemann M, Chenhall R, et al. Norwegian general practitioners' perspectives on implementation of a guided web-based cognitive behavioral therapy for depression: a qualitative study. J Med Internet Res 2014;16(9):e208 [FREE Full text] [doi: 10.2196/jmir.3556] [Medline: 25208886]
59. Hendy J, Chrysanthaki T, Barlow J. Managers’ identification with and adoption of telehealthcare. Societies 2014 Aug 14;4(3):428-445. [doi: 10.3390/soc4030428]
60. Eldh AC, Almost J, DeCorby-Watson K, Gifford W, Harvey G, Hasson H, et al. Clinical interventions, implementation interventions, and the potential greyness in between: a discussion paper. BMC Health Serv Res 2017 Dec 07;17(1):16 [FREE Full text] [doi: 10.1186/s12913-016-1958-5] [Medline: 28061856]

61. Leeman J, Birken SA, Powell BJ, Rohweder C, Shea CM. Beyond “implementation strategies”: classifying the full range of strategies used in implementation science and practice. Implement Sci 2017 Dec 03;12(1):125 [FREE Full text] [doi: 10.1186/s13012-017-0657-x] [Medline: 29100551]
62. Pantoja T, Opiyo N, Lewin S, Paulsen E, Ciapponi A, Wiysonge CS, et al. Implementation strategies for health systems in low-income countries: an overview of systematic reviews. Cochrane Database Syst Rev 2017 Dec 12;9:CD011086 [FREE Full text] [doi: 10.1002/14651858.CD011086.pub2] [Medline: 28895659]
63. Shea CM, Belden CM. What is the extent of research on the characteristics, behaviors, and impacts of health information technology champions? A scoping review. BMC Med Inform Decis Mak 2016 Jan 12;16:2 [FREE Full text] [doi: 10.1186/s12911-016-0240-4] [Medline: 26754739]
64. Lau R, Stevenson F, Ong BN, Dziedzic K, Treweek S, Eldridge S, et al. Achieving change in primary care—effectiveness of strategies for improving implementation of complex interventions: systematic review of reviews. BMJ Open 2015 Dec 23;5(12):e009993 [FREE Full text] [doi: 10.1136/bmjopen-2015-009993] [Medline: 26700290]
65. Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, et al. Tailored interventions to address determinants of practice. Cochrane Database Syst Rev 2015 Apr 29(4):CD005470. [doi: 10.1002/14651858.CD005470.pub3] [Medline: 25923419]
66. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
67. Sandström B, Borglin G, Nilsson R, Willman A. Promoting the implementation of evidence-based practice: a literature review focusing on the role of nursing leadership. Worldviews Evid Based Nurs 2011 Dec;8(4):212-223. [doi: 10.1111/j.1741-6787.2011.00216.x] [Medline: 21401858]
68. Ivers NM, Grimshaw JM, Jamtvedt G, Flottorp S, O'Brien MA, French SD, et al. Growing literature, stagnant science? Systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. J Gen Intern Med 2014 Nov;29(11):1534-1541 [FREE Full text] [doi: 10.1007/s11606-014-2913-y] [Medline: 24965281]
69. Reeves S, Albert M, Kuper A, Hodges BD. Why use theories in qualitative research? BMJ 2008 Aug 07;337:a949. [doi: 10.1136/bmj.a949] [Medline: 18687730]

Abbreviations

eHealth: electronic health
CeHRes: Center for eHealth Research and Disease Management
ERIC: Expert Recommendations for Implementing Change
ICBT: internet-delivered cognitive behavioral therapy
mHealth: mobile health
NASSS: framework for nonadoption, abandonment, scale-up, spread, and sustainability
PROSPERO: Prospective Register of Systematic Reviews

Edited by G Eysenbach; submitted 26.04.19; peer-reviewed by C Shea, T E; comments to author 18.05.19; revised version received 05.06.19; accepted 18.08.19; published 27.09.19.

Please cite as:
Varsi C, Solberg Nes L, Kristjansdottir OB, Kelders SM, Stenberg U, Zangi HA, Børøsund E, Weiss KE, Stubhaug A, Asbjørnsen RA, Westeng M, Ødegaard M, Eide H
Implementation Strategies to Enhance the Implementation of eHealth Programs for Patients With Chronic Illnesses: Realist Systematic Review
J Med Internet Res 2019;21(9):e14255
URL: https://www.jmir.org/2019/9/e14255
doi:10.2196/14255
PMID:31573934

Original Paper

Predicting Dropouts From an Electronic Health Platform for Lifestyle Interventions: Analysis of Methods and Predictors

Daniel Hansen Pedersen[1], MSc; Marjan Mansourvar[2], MSc, PhD; Camilla Sortsø[1], MSc, PhD; Thomas Schmidt[2], MSc, PhD

[1]Liva Healthcare A/S, Copenhagen, Denmark

[2]Centre of Health Informatics and Technology, The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

Corresponding Author:
Daniel Hansen Pedersen, MSc
Liva Healthcare A/S
Danneskiold-Samsøes Allé 41
Copenhagen, 1434
Denmark
Phone: 45 70 20 88 22
Email: dhp@livahealthcare.com

Abstract

Background: The increasing prevalence and economic impact of chronic diseases challenge health care systems globally. Digital solutions can potentially improve efficiency and quality of care, but these initiatives struggle with nonusage attrition. Machine learning methods have been proven to predict dropouts in other settings but lack implementation in health care.

Objective: This study aimed to gain insight into the causes of attrition for patients in an electronic health (eHealth) intervention for chronic lifestyle diseases and evaluate if attrition can be predicted and consequently prevented. We aimed to build predictive models that can identify patients in a digital lifestyle intervention at high risk of dropout by analyzing several predictor variables applied in different models and to further assess the possibilities and impact of implementing such models into an eHealth platform.

Methods: Data from 2684 patients using an eHealth platform were iteratively analyzed using logistic regression, decision trees, and random forest models. The dataset was split into a 79.99% (2147/2684) training and cross-validation set and a 20.0% (537/2684) holdout test set. Trends in activity patterns were analyzed to assess engagement over time. Development and implementation were performed iteratively with health coaches.

Results: Patients in the test dataset were classified as dropouts with an 89% precision using a random forest model and 11 predictor variables. The most significant predictors were the provider of the intervention, 2 weeks inactivity, and the number of advices received from the health coach. Engagement in the platform dropped significantly leading up to the time of dropout.

Conclusions: Dropouts from eHealth lifestyle interventions can be predicted using various data mining methods. This can support health coaches in preventing attrition by receiving proactive warnings. The best performing predictive model was found to be the random forest.

(J Med Internet Res 2019;21(9):e13617) doi:10.2196/13617

KEYWORDS

eHealth; patient dropouts; adherence; law of attrition; digital health; chronic disease; data mining; logistic regression; decision trees

Introduction

Background

Chronic diseases such as diabetes, heart disease, chronic obstructive pulmonary disease, and cancer are collectively responsible for more than two-thirds of all deaths and 75% of the health care budget spending in Europe [1]. The increasing prevalence and enormous economic impact of chronic diseases are a critical threat to health care systems. This necessitates new treatments that can effectively handle more people at a lower resource-to-outcome ratio. The application of mobile computing and communication technology in health care (denoted as electronic health [eHealth]) has introduced new possibilities in terms of improving efficiency and quality of care [2]. Despite several studies showing promising results in terms of outcomes such as weight loss [3] and behavior change [4], the evidence for long-term effectiveness, and especially how to retain patients in digital interventions, remains limited [5,6,7].

In any eHealth program, adherence is a key challenge, as a substantial proportion of patients stop using the application and thus drop out of the intervention program before its completion, referred to as *nonusage attrition*, or simply a *dropout* [8]. Recently, a dropout rate of 72% was reported in an eHealth intervention for adults with type 2 diabetes [9], and more generally, dropout rates up to as high as 83% are reported [10,11]. Studies have sought to identify predictors of dropout, but a consistent set of predictors has not yet been identified [12]. Previous studies have found engagement and participation in an online forum [13], depressive mood [14], age, gender, vocational education and employment status [15], disease severity, treatment length, and chronicity [11] to be related to attrition. Prediction of dropouts has been evaluated in multiple studies, in which many often have been offset in an educational institution setting where high dropout rates are also a great concern. Survival analysis [16], logistic regression, random forest, and other machine learning algorithms [17] are commonly applied to address this problem, using demographics and other characteristics to predict dropout. The documented high attrition rates from eHealth interventions make it an attractive case to apply similar methods to predict patients at high risk of dropping out. Furthermore, the literature on data mining and predictive methods in relation to attrition in eHealth settings is very limited, suggesting a lack of implementation and integration of these methods in the eHealth domain.

Objectives

The aim of our study was to assess the variables and methods for predicting dropouts of patients with chronic diseases in a digital lifestyle intervention and review their applicability for implementation in an eHealth platform. We utilized self-reported data including patient-reported outcome measures (PROMs) submitted by chronic lifestyle disease patients in an eHealth intervention provided by the Liva Healthcare (LIVA) platform (Textbox 1). To assess the research question of *how self-reported data can be applied to address the challenge of attrition in an eHealth setting*, we evaluated the factors associated with dropout and applied logistic regression, decision trees, and random forest. We proposed how these models can be implemented to visualize the results as predictive warnings to reduce dropouts. In this way, the data are used to improve the eHealth intervention by supporting health professionals and enabling them to re-engage patients at high risk of attrition. As defined previously, we applied a broad definition of eHealth given the scope of the intervention under study that is targeting a wide range of patients. However, the challenge of attrition is relevant for most eHealth interventions for lifestyle change.

Textbox 1. Short description of the LIVA platform and intervention.

LIVA is a digital platform designed to facilitate lifestyle changes for patients with chronic diseases. The platform is used by Danish municipalities. Patients have an initial goal-setting meeting with their coach and are introduced to the LIVA app that allows setting and registering health goals (eg, steps, weight, exercise, or diet), monitoring progress, dialog with the health coach by receiving advice and sending messages, and participation in an online forum. Health coaches access the platform through an internet browser and are able to proactively advise patients on a weekly-to-monthly basis based on their patients' input in the platform. Personal data and health information are collected from the patients during the 3 to 12-month intervention program to provide the treatment service and for research purposes.

Methods

Ethics and Approvals

Only pseudonymized data for which patients had granted their consent to make them available for research purposes were used in this study. Consent was obtained explicitly in the sign-up flow before the patient's use of the service. Liva Healthcare processes the data as the data processor, using the means and purpose defined by the data controllers, that is, Danish municipalities.

Data Collection and Selection

The study was retrospective, applying data collected by Liva Healthcare from June 7, 2016, to March 21, 2018. For active users, the anchor time point for features was the date of data collection and for dropouts, we went back 4 weeks before the date of dropout (see next section). Data were extracted from a Microsoft SQL database and further processed in Alteryx. The dataset contains several unvalidated PROMs and sociodemographic information entered by the health coaches. Consequently, values for weight loss and body mass index (BMI) were filtered to remove extreme outliers and unrealistic values (weight differences of >3.5 kg/week on average for weight registrations over 30 days or more and BMI >100 kg/m^2). The dataset was cleansed to only include patients who were referred to the platform by their doctor or municipality and showed commitment to the intervention by being properly set up in an advisory, received 3 or more advices from their coach, and had been active in the platform for at least 14 days (N=2684). A baseline of 14 days was selected as patients receive weekly advice in the beginning and should therefore receive their third advice on the fourteenth day of the program. Patients with *less* than 3 advices and 14 days of participation in the program were removed from the dataset as they had either not yet shown commitment to the intervention, signed up by a mistake, or merely signed up within the last 14 days of data collection. Thus, it is not known if these patients dropped out or never meant to use the service.

Definition of Dropout

Generally, a large variation in the reporting and measurement of adherence is seen in previous literature [18]. For the objective of this study, it was relevant to look at dropouts as patients who commit to the intervention and thenceforward discontinue using the platform, consequently dropping out. We proposed a definition of dropout that aligns with Eysenbach's characterization from 2005 [8] and other operationalizations

[11]. Dropout is hence defined as *4 consecutive weeks of not performing any actions,for example,registrations or messages, in the platform*. The threshold for a dropout after 4 weeks of inactivity was defined based on the insight that less than 5.0% (117/2684) of the patients become re-engaged in the program after 4 weeks of inactivity. For 79.2% (2126/2684) of the patients, 2 weeks of inactivity equals dropout, and 84.61% (2271/2684) of them drop out after 3 weeks of inactivity. Furthermore, dropouts are limited to the active coaching period, which is a maximum of 12 months. This might be lower for some providers, for example, 3 or 6 months. Patients who are still active after 12 months will move to a *retention* phase, and they will thus not be considered as dropouts if they discontinue the intervention.

Analysis of Dataset

To gain insight into the population and understand the factors associated with nonusage attrition, we performed several analyses of the users' activity patterns by illustrating activity over time in the program for several subgroups of the population. We defined a formula for the current activity level in percentage based on these insights. Some descriptive user statistics of the population and analysis of predictors in *t* test and simple logistic regression models are provided to gain additional insights.

Data Mining and Model Evaluation

The Cross-Industry Standard Process for Data Mining framework [19] was applied as an iterative data mining approach. This allowed for several iterations of the models to be developed as the knowledge of the population increased based on a better understanding for the dataset and end users (health coaches) who were able to provide feedback during each iteration.

Inspired by studies performed in an educational setting, logistic regression, decision trees, and random forest methods were applied to classify participants in the intervention into either dropouts or nondropouts based on specific characteristics.

We tested 11 variables that were well represented in the dataset as potential predictors of dropout: Gender, age group, provider of the intervention, period of intervention, BMI at the beginning of the intervention, weight loss, number of advices received, number of messages sent, total number of weeks with inactivity, and inactivity in the last 1 or 2 consecutive weeks. For weight loss, we required registrations over at least 30 days to be included. Less than 20% of the patients who registered had provided their educational status and zip code, because of the low quantity, these variables were not applied.

The variables in the final models for logistic regression were selected based on mixed backward and forward selection using the Akaike information criterion [20]. For decision trees and random forest, variables were selected inherently by the hyperparameters. The minimum number of records allowed for a split and a terminal node was set to 50 and 25, respectively. The maximum allowed depth in the final tree was set to 10 to avoid overfitting. The trees were pruned with a complexity parameter set to 0.01 to reduce the number of branches and the relative error.

To assess the quality of the 3 different models and to compare the predictive performance, the dataset was split into an 80% training and cross-validation set and a 20% test set. Owing to the relatively small size of the dataset, the training and cross-validation were performed using stratified 10-fold cross-validation. Stratification was applied on the target variable to ensure each fold was a good representative of the overall dataset distribution to reduce the bias and variance of the models. The best performing method was then applied to the 20% holdout test set that had never been seen by the model. The quality of the models was assessed based on the area under the curve (AUC) on the receiver operating characteristic (ROC) curve, the precision, and the accuracy. The goal was to have a high precision as the false negatives were the most critical to reduce in this scenario, that is, patients at high risk for dropout not identified as a potential dropout.

Adoption and Implementation

The findings from this study have been diffused among the health coaches using the Liva Healthcare platform and prototype models have been implemented into the platform. Interviews were conducted with health coaches to adjust the models in terms of when and how warnings should be present. Feedback was continuously collected, and data were analyzed to assess dropout rate following implementation.

Results

User Statistics

The final dataset contained 2684 patients registered in the LIVA database. The population was characterized by a greater proportion of females (1943/2684, 72.39%) compared with males (741/2684, 27.6%). The majority of the population was in the age range of 40 to 59 years, and the average lifetime on the platform was 108 days. Overweight patients represent the largest treatment group, but patients might enter the program with one or more of 7 other diseases and possible comorbidities (other secondary disease). Additional characteristics of the population are provided in Table 1 (*advice received* and *messages sent* refer to the dialog between patient and health coach).

The intervention status for the patients was that 53.99% (1449/2684) had dropped out, 39.43% (1060/2684) were currently in active advisory, 3.7% (100/2684) had completed the intervention (finished intervention after >12 months), and 3% (75/2684) were in the retention phase (>12 months in program). More than 1 in 4 dropouts had occurred in the first month of the program (between day 14 and 31, n=388, 26.8% of dropouts; Table 2).

Table 1. A summary of the population from the final dataset included in the models.

Parameter (statistic)	Description
Sample size (N)	2684 patients
Number of providers (N)	18 different providers with between 13 and 581 patients ever in program
Gender (percentage distribution)	72.4% females and 27.6% males
Age (years), mean (SD)	48.6 (13.2)
Treatment groups (percentage distribution)	Overweight (85%), diabetes (17%), heart diseases (12%), chronic obstructive pulmonary disease (5%), stress (15%), cancer (1%), alcoholism (1%), smoking (6%), or another secondary disease (20%)
Days on platform (minimum, median, maximum)	14, 82, 595
Start body mass index (kg/m^2), mean (SD)	33.6 (6.0)
Advice received (minimum, median, maximum)	3, 7, 99
Messages sent (minimum, median, maximum)	0, 3, 156

Table 2. Number of dropouts over the period of intervention.

Months of intervention	Number of dropouts
1	388
2-4	633
5-8	300
9-12	128

Preliminary Analysis of Predictors

In total, a larger proportion of females (1069/1943, 55.02%) had dropped out compared with males (380/741, 51.3%), and for age groups, the highest dropout rate was found among the oldest (above 75 years; Table 3).

Characteristics for the patients in active advisory and the dropouts were assessed for the predictor variables of interest. Gender was found to be significant in a Welch 2-sample *t* test (*P*=.01). Dropouts had a slightly lower starting BMI than the active patients (*P*=.01). No major differences were seen in average age among the 2 groups (*P*=.60) nor average weight loss (*P*=.88). Large variations in the risk for dropout were found among the different providers of the intervention, varying from 7.3% to 87.0% in a simple logistic regression model.

Table 3. Percentage of dropouts distributed in age group and gender. The percentage indicates the proportion of dropouts for the patients in the specific age group and gender.

Age group (years)	Female, n (%)	Male, n (%)	Total, n (%)[a]
18-39	600 (53.72)	161 (51.60)	761 (53.18)
40-59	1040 (55.48)	395 (50.13)	1435 (54.01)
60-74	284 (52.11)	161 (50.31)	445 (51.39)
>75	19 (63.33)	24 (54.55)	43 (58.11)
Total	1943 (55.02)	741 (51.32)	2684 (53.99)

[a] Percentage of the total population of participants.

Activity Analysis

The largest proportion of active patients was in months 2 to 4 in the program and the lowest proportion was found at the beginning of the program in month 1 (Figure 1). The odds for a patient dropping out in the first month of the intervention were 4.35 times higher than for dropping out past month 8.

We analyzed if trends in the patients' activity patterns could identify attrition by analyzing patient engagement in the platform over time. Evidently, patients who drop out have a very low level of activity (defined as a registration, forum posting, or messaging the coach) in the last weeks of their time on the platform. Overall, 71.77% (1040/1449) of the dropouts decreased their activity level by more than 50% in their last 2 weeks. However, there is also an expected decrease in activity that will occur over time, and individuals will have different trend lines for patterns in activity (Figure 2). The first week in the program (week 0) was found to have a significantly higher amount of registrations than the remaining weeks, on average 23.8 registrations, and was removed from the analysis to prevent skewing the linear regression line.

Figure 1. Proportion of active patients over 4 segments of the intervention period.

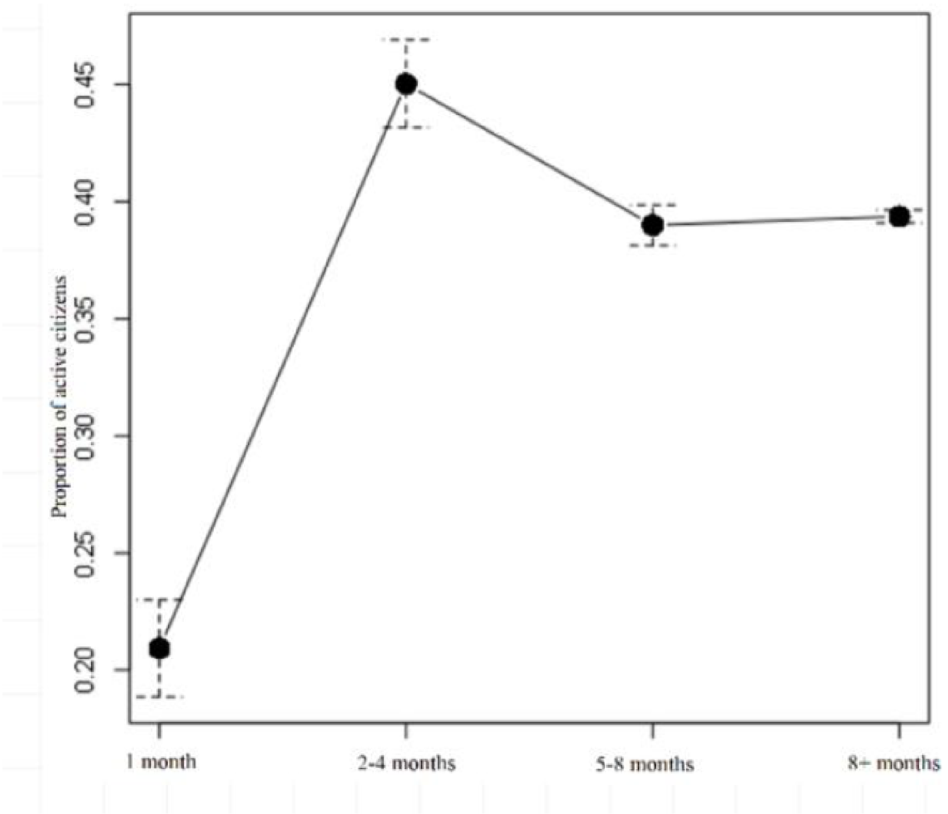
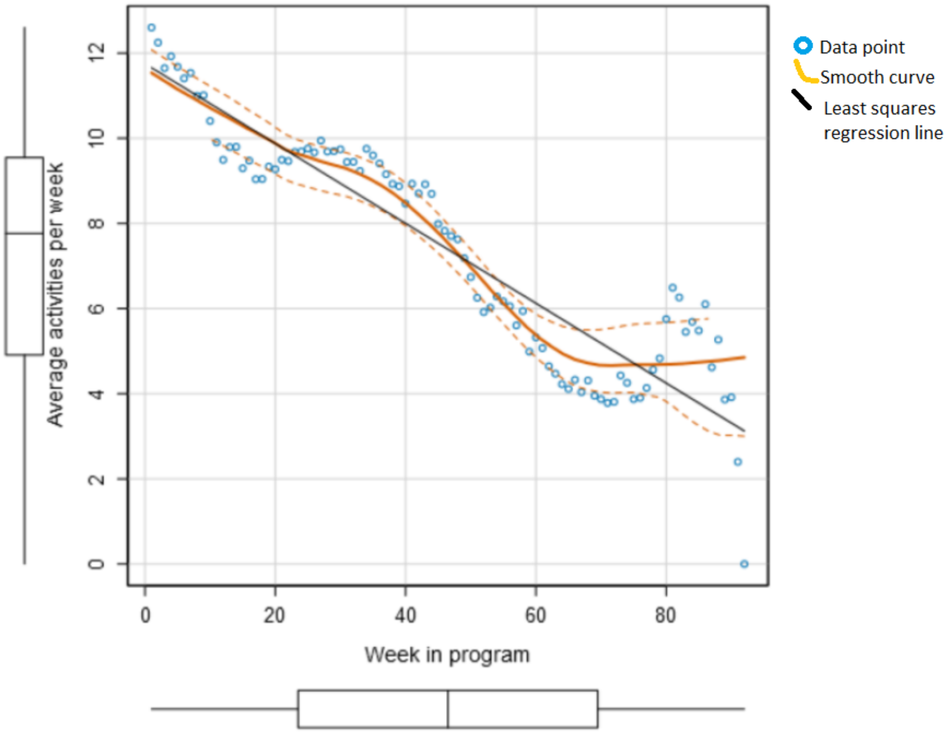


Figure 2. Average number of activities in the platform per week in the program for patients who either completed the intervention or entered retention (n=175), excluding week 0 in the program.



A variable for the patient’s current activity level defined in percentage of the baseline activity level was proposed (Equation 1). The variable accounts for (1) the patient’s average activity in the last 2 weeks, (2) the patient’s baseline activity (defined as the average activity in weeks 1-4 of the program), and (3) the regression line for the average activity levels over time (Figure 2). The average decrease in activities was found to be

0.094 per week. Patients with zero activity in the last 2 weeks will per default have a current activity level of 0%.

> Current activity level (%) = Average activity last two weeks / (Baseline activity - (0.094 × weeks in program)) (1)

Fewer women tended to be active at the same number of inactive weeks compared with males, up to 40 weeks of inactivity, at which point of time, a very small percentage were still active in the program at the time of the data collection (Figure 3). For age groups, the oldest and the youngest age groups were the least active (Figure 4). The oldest group was also found to have the highest rate of dropouts (Table 3).

Figure 3. Proportion of active patients over the total number of inactive weeks in the program defined by gender.

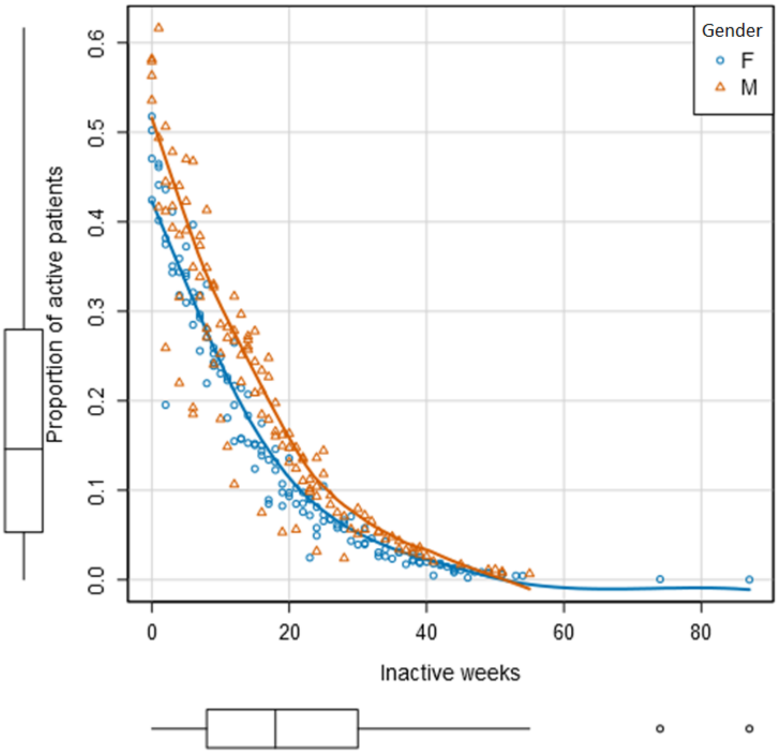


Figure 4. Proportion of active patients over the total number of inactive weeks in the program defined by age.

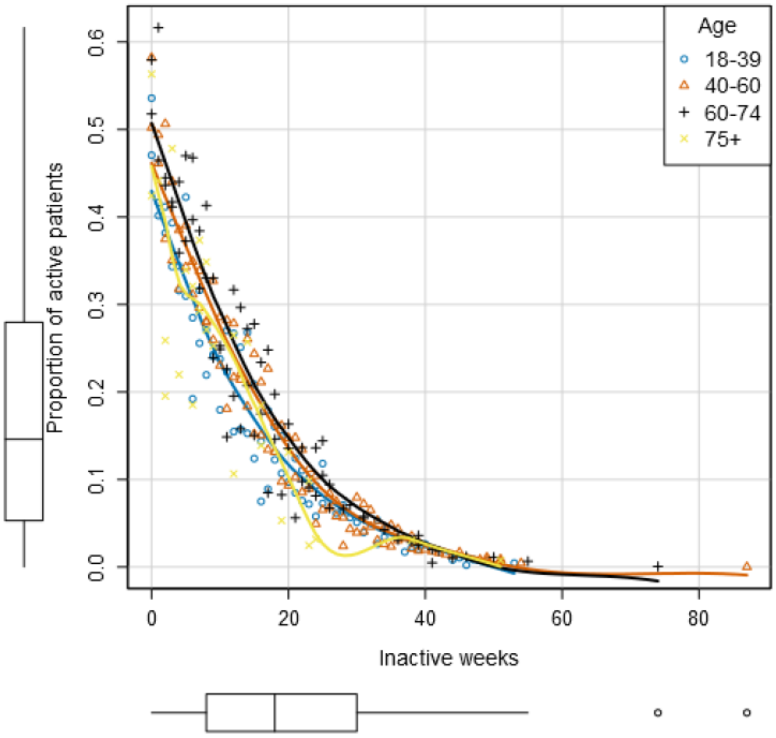
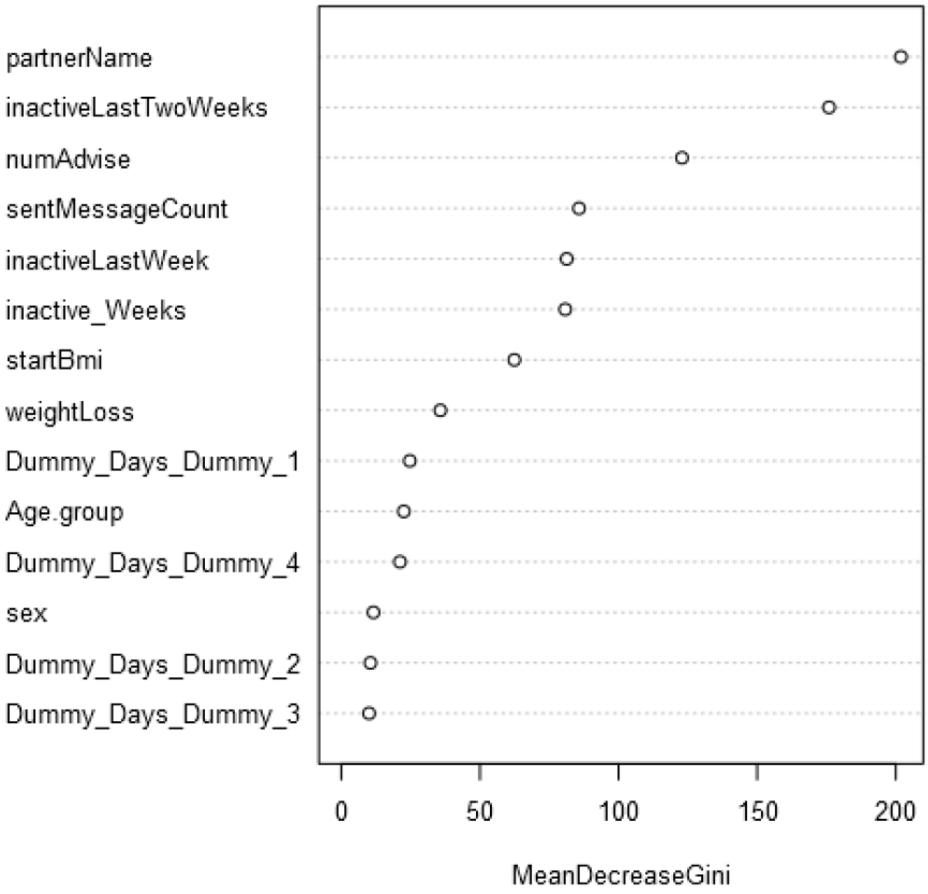


Figure 5. Variable importance plot for the 11 selected variables. Period of intervention is separated into 4 dummy variables.



Finally, introducing a variable that accounts for inactivity in the last 2 weeks of the program in any of the proposed models resulted in it becoming incomparably significant to any of the models' other variables, except for provider of intervention (partnerName; see Figure 5). Hence, inactivity in the platform should be a critical warning for the health coach. A significant decrease in activity that deviates from the overall pattern might also be a critical sign for attrition and consequently an attention mark for the health coach.

Model Selection for Dropout Prediction

The random forest achieved an AUC of 0.92 on the ROC chart and a Gini score of 0.84 on the stratified cross-validated training data, making it the best-performing model (Figure 6) compared with decision trees and logistic regression (Table 4). When applied to the holdout test data, the AUC increases by 0.01, and the model is thus not suspect to overfitting.

The precision of the random forest model was 0.89, with an overall accuracy of 0.86. This means that 89% of all dropouts were classified correctly as dropouts and 11% were mistakenly classified as active patients. This corresponds with 88.7% (253/285) dropouts in the holdout test data being classified correctly as dropouts.

The outlined models for inactivity, activity level, and dropout risk have been implemented into the LIVA platform for the health coaches to be notified of attention markers related to patients that are at high risk of dropping out. The threshold values for activity levels have been set to highlight patients at medium (current activity level below 60% of expected) and high risk (below 40%), visualized with yellow and red warnings, respectively, for the health coach (Figure 7). The random forest model for dropout will show a yellow warning as the risk increases to above 60% and red if the risk is above 75%. The thresholds were selected based on an assessment of the patient distributions in collaboration with the health coaches.

Figure 6. Receiver operating characteristic curve with area under the curve for the random forest model on the holdout test data.

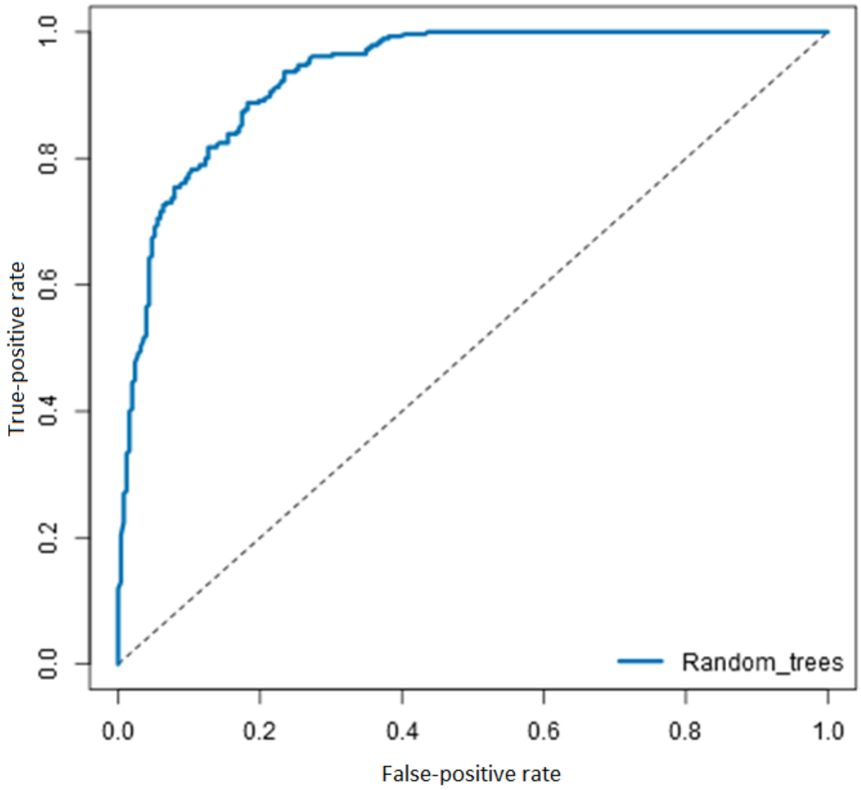
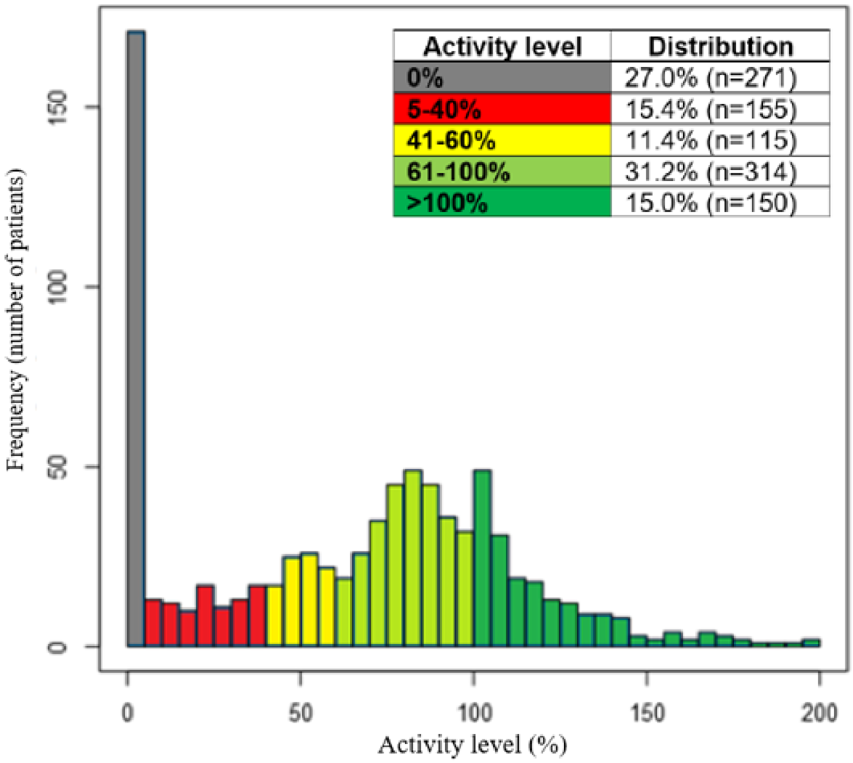


Table 4. Area under the curve (AUC) and Gini index for the receiver operating characteristic on the 3 applied best performing models.

Model	AUC	Gini
Logistic regression	0.84	0.68
Decision trees	0.82	0.64
Random forest	0.92	0.84

Figure 7. Histogram of current activity level (%; calculated using Equation 1) compared with forecasted activity based on the linear overall population trend line. Only patients with at least 6 weeks on the platform included



Discussion

Principal Findings

This study applies real-world data from chronic lifestyle disease patients enrolled in an eHealth lifestyle intervention in municipal settings in Denmark. The findings show promising results in terms of applying data mining methods for the prediction of dropouts in eHealth interventions with high precision. To summarize, the following 4 key takeaway points were made clear in this study:

1. Patients are at the highest risk of dropout at the beginning of the intervention. Most dropouts occurred in the first part of the intervention, and evidence from other studies support the finding that when participants dropout, they do so early. For instance, 65% of dropouts from a diet and physical activity short message service text message program occurred within the first 2 weeks [21].
2. Attrition is not an abrupt process but something that happens over time. We found that patients reduce their activity in the platform significantly in the weeks leading up to their dropout. Therefore, being aware of abnormal decreases in activity should be a good indicator for health coaches to initiate re-engagement.
3. Dropout is primarily related to the program provider, outline of the intervention, activity in the platform (engagement), and, to a lower degree, the demographic variables available in this study. Multiple other studies have found attrition to be related to demographic variables [22,15], but these were not available in this dataset.
4. Predicting activity level and risk of dropout can enable personalized advices and goal settings. Our findings strongly suggest that dropouts can be predicted, and personalized coaching can be supported by several parameters. However, there are some limitations to the study that will be discussed in the following section.

Limitations

The definition of dropout was based on patients not using the platform for 4 consecutive weeks. However, this might also include patients who stop using the platform because of having achieved their desired goal or goals, for example, behavior change or weight loss, or because the advisor has terminated the patient for other reasons. Termination reasons have been implemented in the platform following this study. In addition, the length of the intervention program offered by the provider was varying and unknown. This influences the likelihood of patients being incorrectly labelled as dropouts and should therefore be taken into consideration for the definition of the matter.

Dropouts in the first 14 days were excluded from this study because of large uncertainties in the reason for dropout. As this group of patients was very diverse and the amount of data were primarily limited to their sign-up registration, it provided a restricted dataset for analysis. This suggests a possibility for a future study to look into these very early dropouts.

The reliability of the findings is limited by some of the applied data (weight loss and BMI) being PROMs and most of the variables being self-reported by the patients. The data are, to some degree, validated by the health coaches, and extreme outliers are automatically marked as unrealistic by the Liva Healthcare system, but it has not been clinically verified.

The activity analysis was based on a simple linear regression despite the pattern being shaped as a higher-order polynomial. A simple approach was sufficient for the focus of this study; however, further investigations into activity patterns showed large variations in individual patterns. For future studies, time series analysis on the activity patterns of a larger population could potentially be the basis for interesting risk models that could analyze activity levels for distinct patient types.

Adoption and Implementation

The insights obtained from our work are not sufficient to maintain engaged patients on the platform. The knowledge must be put into action to have an effect on the attrition rates. The discoveries must be diffused among the health coaches using the eHealth platform but should also be integrated into the system to support them to the highest degree possible.

The finding that dropout is not an abrupt process but something that happens over time underlines the importance of the health coaches being warned of dropout risk to actively attempt a prevention of attrition. From a practical computational perspective, a random forest is a computational and expensive model, and depending on the nature of the intervention, the frequency of registrations by the patients, and the technical setup, it might not be applicable. Nevertheless, if the model is not expected to change frequently, then calculations can be performed, for example, every night, and can be used as the daily baseline for dropout risk in the advisors' overview. Otherwise, more simple models such as logistic regression might be preferred.

Previous studies have shown adherence to be closely related to the level of engagement in the platform, that is, by participating in an online forum [13]. Socioeconomic status, occupation, and educational status have shown to be related to dropout [15], but this type of data have not been available in sufficient quantities for this research. Diagnosis and condition should also be included in future models. Utilizing these data types would provide important information for the models and likely increase the accuracy and possibly make distinct patient profiles clear. Thus, this added data could be used to further enhance and individualize the models.

The provider of the intervention was found to be the most significant predictor of dropout, together with inactivity on the platform. This indicates that efforts toward preventing attrition should be targeted at providers to the same degree as patients. Provider-specific attrition models may perform even better than the generic approach proposed in this work. Further insights into the providers and their strategies are required.

Finally, to better validate the warnings that have been implemented into the platform, it requires a randomized setup or a less diverse population that is spread across multiple providers with varying programs. However, initial feedback from the health coaches is very positive, and the dropout rate for patients that have entered into the program after the date of data collection for this study is at only 19.3% (N=6402) compared with 54.0% for the population included in this study.

Perspectives

This study contributes to the literature on adherence and nonusage attrition in eHealth by analyzing activity patterns, assessing various methods and predictor variables for predicting dropout in a chronic patient lifestyle intervention, and proposing some perspectives for implementation. We expect future research and development in eHealth to apply data mining methods in the process of tailoring information to patients in a higher degree to achieve personalized interventions as the field of digital health continues to evolve. Ongoing research is currently assessing how lifestyle interventions can be tailored to the individual patient [23], and as artificial intelligence is gaining ground within health care, we expect to see interventions, treatment, and guidance being selected based on the most suited for the specific individual patient profile in the future.

Conclusions

It is possible to apply methods from data mining in the context of predicting dropouts in an eHealth setting. Stratified cross-validation shows that patients at high risk of dropout can be predicted with 89% precision using a random forest model. Computational simpler models, such as logistic regression, are applicable as well but might produce less precise predictions. The risk of dropout can be visualized as warnings for the health coaches, so they can attempt to re-engage the patient in their intervention before dropout. Initial assessment of the models implemented in an eHealth platform in use shows a decrease in dropout rate. Obtaining more rich data on educational status and socioeconomic factors in combination with a better delineation of dropouts would increase the quality of the models.

Acknowledgments

DHP analyzed the data, created and evaluated the models, and drafted the paper. MM, TS, and CS supervised the data analyses. MM, TS, and CS gave inputs on several drafts for the manuscript. All authors read, critically revised, and approved the final manuscript. Liva Healthcare provided the data and allocated resources to conduct and assist in the research and creation of this paper. The publishing of this paper was funded by the University of Southern Denmark, Health Informatics.

Conflicts of Interest

DHP and CS are employed by Liva Healthcare A/S.

References

1. Gemmill M. CiteSeerX. 2008. Research Note: Chronic Disease Management in Europe URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.619.4434&rep=rep1&type=pdf
2. Fiordelli M, Diviani N, Schulz PJ. Mapping mhealth research: a decade of evolution. J Med Internet Res 2013 May 21;15(5):e95. [doi: 10.2196/jmir.2430] [Medline: 23697600]
3. Brandt V, Brandt CJ, Glintborg D, Arendal C, Toubro S, Brandt K. Sustained weight loss during 20 months using a personalized interactive internet based dietician advice program in a general practice setting. Int J Adv Life Sci 2011;3(1-2):23-28 [FREE Full text]
4. Afshin A, Babalola D, Mclean M, Yu Z, Ma W, Chen CY, et al. Information technology and lifestyle: a systematic evaluation of internet and mobile interventions for improving diet, physical activity, obesity, tobacco, and alcohol use. J Am Heart Assoc 2016 Aug 31;5(9):pii: e003058 [FREE Full text] [doi: 10.1161/JAHA.115.003058] [Medline: 27581172]
5. Coiera E. Telehealth and mobile health. In: Guide to Health Informatics. Third Edition. Florida, United States: CRC Press; 2015:343-369.
6. Mistry H. Systematic review of studies of the cost-effectiveness of telemedicine and telecare. Changes in the economic evidence over twenty years. J Telemed Telecare 2012 Jan;18(1):1-6. [doi: 10.1258/jtt.2011.110505] [Medline: 22101609]
7. 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]
8. 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]
9. Lie SS, Karlsen B, Oord ER, Graue M, Oftedal B. Dropout from an ehealth intervention for adults with type 2 diabetes: a qualitative study. J Med Internet Res 2017 May 30;19(5):e187 [FREE Full text] [doi: 10.2196/jmir.7479] [Medline: 28559223]
10. Melville KM, Casey LM, Kavanagh DJ. Dropout from internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]
11. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
12. Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic review of the literature. Obes Rev 2011 Nov;12(11):912-934. [doi: 10.1111/j.1467-789X.2011.00915.x] [Medline: 21815990]
13. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. J Med Internet Res 2010 Dec 17;12(4):e71 [FREE Full text] [doi: 10.2196/jmir.1338] [Medline: 21169160]
14. von Brachel R, Hötzel K, Hirschfeld G, Rieger E, Schmidt U, Kosfelder J, et al. Internet-based motivation program for women with eating disorders: eating disorder pathology and depressive mood predict dropout. J Med Internet Res 2014 Mar 31;16(3):e92 [FREE Full text] [doi: 10.2196/jmir.3104] [Medline: 24686856]
15. Kannisto KA, Korhonen J, Adams CE, Koivunen MH, Vahlberg T, Välimäki MA. Factors associated with dropout during recruitment and follow-up periods of a mhealth-based randomized controlled trial for mobile.net to encourage treatment adherence for people with serious mental health problems. J Med Internet Res 2017 Feb 21;19(2):e46 [FREE Full text] [doi: 10.2196/jmir.6417] [Medline: 28223262]
16. Ameri S, Fard MJ, Chinnam RB, Reddy CK. Survival Analysis Based Framework for Early Prediction of Student Dropouts. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. 2016 Presented at: CIKM'16; October 24-28, 2016; Indianapolis, Indiana, USA p. 903-912. [doi: 10.1145/2983323.2983351]
17. Rovira S, Puertas E, Igual L. Data-driven system to predict academic grades and dropout. PLoS One 2017;12(2):e0171207 [FREE Full text] [doi: 10.1371/journal.pone.0171207] [Medline: 28196078]
18. 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 5;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
19. Shearer C. The CRISP-DM model: the new blueprint for data mining. J Data Warehous 2000;5(4):13-22 [FREE Full text]
20. Akaike H. A new look at the statistical model identification. IEEE T Automat Contr 1974 Dec;19(6):716-723. [doi: 10.1109/TAC.1974.1100705]
21. Coa K, Patrick H. Baseline motivation type as a predictor of dropout in a healthy eating text messaging program. JMIR Mhealth Uhealth 2016 Sep 29;4(3):e114 [FREE Full text] [doi: 10.2196/mhealth.5992] [Medline: 27688034]
22. Grutzmacher SK, Munger AL, Speirs KE, Vafai Y, Hilberg E, Duru EB, et al. Predicting attrition in a text-based nutrition education program: survival analysis of Text2BHealthy. JMIR Mhealth Uhealth 2019 Jan 21;7(1):e9967 [FREE Full text] [doi: 10.2196/mhealth.9967] [Medline: 30664489]
23. Elissen AM, Hertroijs FD, Schaper CN, Vrijhoef JH, Ruwaard D. Profiling patients' healthcare needs to support integrated, person-centered models for long-term disease management (profile): research design. Int J Integr Care 2016 Apr 29;16(2):1 [FREE Full text] [doi: 10.5334/ijic.2208] [Medline: 27616957]

Abbreviations

AUC: area under the curve
BMI: body mass index
eHealth: electronic health
PROMs: patient-reported outcome measures
ROC: receiver operating characteristic

Edited by G Eysenbach; submitted 05.02.19; peer-reviewed by H de Vries, N Miyoshi, S Kamalakannan; comments to author 29.04.19; revised version received 24.06.19; accepted 07.07.19; published 04.09.19.

Please cite as:
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;21(9):e13617
URL: https://www.jmir.org/2019/9/e13617/
doi:10.2196/13617
PMID:31486409

Original Paper

Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis

Amit Baumel[1], PhD; Frederick Muench[2], PhD; Stav Edan[1]; John M Kane[3], MD

[1]Department of Community Mental Health, University of Haifa, Haifa, Israel

[2]The Partnership for Drug-Free Kids, New York, NY, United States

[3]Psychiatry Research, The Zucker Hillside Hospital, Northwell Health, Glen Oaks, NY, United States

Corresponding Author:
Amit Baumel, PhD
Department of Community Mental Health
University of Haifa
Abba Khoushy Ave 199
Haifa, 3498838
Israel
Phone: 972 048240111
Email: abaumel@univ.haifa.ac.il

Abstract

Background: Understanding patterns of real-world usage of mental health apps is key to maximizing their potential to increase public self-management of care. Although developer-led studies have published results on the use of mental health apps in real-world settings, no study yet has systematically examined usage patterns of a large sample of mental health apps relying on independently collected data.

Objective: Our aim is to present real-world objective data on user engagement with popular mental health apps.

Methods: A systematic engine search was conducted using Google Play to identify Android apps with 10,000 installs or more targeting anxiety, depression, or emotional well-being. Coding of apps included primary incorporated techniques and mental health focus. Behavioral data on real-world usage were obtained from a panel that provides aggregated nonpersonal information on user engagement with mobile apps.

Results: In total, 93 apps met the inclusion criteria (installs: median 100,000, IQR 90,000). The median percentage of daily active users (open rate) was 4.0% (IQR 4.7%) with a difference between trackers (median 6.3%, IQR 10.2%) and peer-support apps (median 17.0%) versus breathing exercise apps (median 1.6%, IQR 1.6%; all $z \geq 3.42$, all $P<.001$). Among active users, daily minutes of use were significantly higher for mindfulness/meditation (median 21.47, IQR 15.00) and peer support (median 35.08, n=2) apps than for apps incorporating other techniques (tracker, breathing exercise, psychoeducation: medians range 3.53-8.32; all $z \geq 2.11$, all $P<.05$). The medians of app 15-day and 30-day retention rates were 3.9% (IQR 10.3%) and 3.3% (IQR 6.2%), respectively. On day 30, peer support (median 8.9%, n=2), mindfulness/meditation (median 4.7%, IQR 6.2%), and tracker apps (median 6.1%, IQR 20.4%) had significantly higher retention rates than breathing exercise apps (median 0.0%, IQR 0.0%; all $z \geq 2.18$, all $P \leq .04$). The pattern of daily use presented a descriptive peak toward the evening for apps incorporating most techniques (tracker, psychoeducation, and peer support) except mindfulness/meditation, which exhibited two peaks (morning and night).

Conclusions: Although the number of app installs and daily active minutes of use may seem high, only a small portion of users actually used the apps for a long period of time. More studies using different datasets are needed to understand this phenomenon and the ways in which users self-manage their condition in real-world settings.

(J Med Internet Res 2019;21(9):e14567) doi:10.2196/14567

KEYWORDS

user engagement; usage; adherence; retention; mental health; depression; anxiety; mHealth

Introduction

The wide dissemination of mobile phone devices and the leap in the development and distribution of mobile health (mHealth) apps have altered the ways in which scholars conceptualize care management in the behavioral health domain. The conversation has shifted from patients and providers to individuals who can now engage in self-care around the clock outside of traditional health care settings (eg, [1-3]). Approximately 77% of the US adult population, and more than 89% of those younger than 50 years, now own a mobile phone [4,5] where they can store and use computerized apps. This widespread use has established a market for mHealth apps. Accordingly, a 2015 World Health Organization survey identified approximately 15,000 mobile apps for health care, with at least 29% designed for mental health [6].

The use of unguided apps has the potential to increase access to care in a scalable manner by reducing the costs associated with service uptake [7,8]. However, the impact of digital interventions is limited by their ability to engage users in therapeutic activities and to support user adherence to the therapeutic process [9,10]. Digital interventions require individuals to engage with self-care outside of traditional settings; therefore, individuals' engagement must compete with other events in their daily lives and endure fluctuating motivation to be involved in effortful behavior [11]. As a result, user engagement with mobile apps and websites across the behavior change spectrum is low in the absence of human support [12-14]. Furthermore, various studies have suggested that most users of unguided Web-based programs exit websites before the full completion of the offered program [9,10,15,16]. For example, Christensen and colleagues [17] reported that less than 1% of users completed all modules in MoodGym, an open-access website for depression. In a systematic review of published articles reporting real-world user engagement with unguided programs for depression, anxiety, or mood enhancement, Fleming and colleagues [18] reported that 7% to 42% of users of Web- and app-based programs engaged in moderate use (completing between 40% and 60% of modular fixed-length programs or continuing to use the app after 4 weeks). For example, the developers of the PTSD Coach mobile app reported a usage decline over time, with 41.6% continuing to use the app 1 month after installation and 19.4% after 6 months [19]. Among Happify mobile app users, 3.5% completed a 6-week assessment. However, the authors noted that these users might have completed assessments without engaging in other content [20] (see [18] for a review).

Understanding patterns of real-world usage of e-mental health apps outside of empirical trials is key to maximizing the potential of apps to increase the public self-management of care. Utilization in real-world settings may differ from that in study settings for several reasons. First, empirical study settings include enrollment and assessment procedures that are not part of real-world utilization of the app, as trials largely emphasize internal validity over real-world generalizability [13]. Ebert and Baumeister [21] claim, for example, that within randomized trials "the securing of commitment represents an adherence-promoting element in self-help interventions." It is reasonable to assume that the human contact provided by research coordinators, provision of ongoing assessments, and reimbursement to incentivize the completion of assessments—none of which are available in real-world use—impact engagement patterns with the interventions. Second, from an external validity perspective, recruitment challenges in trials are often addressed by increasing the reach to potential participants through the expansion of participating venues and the refinement of social media strategies [13]. In this way, researchers unintentionally recruit people who are much more likely to adhere to e-mental health technologies than people in the general population who download and try available programs "in the wild." Such assumptions are supported by a systematic review of internet interventions for anxiety and depression, which found that the rates of attrition in randomized controlled trials were lower than the reported dropout rates from open-access websites [22].

Overall, there is a need to understand how the general population engages with the most popular unguided mobile apps targeting anxiety, depression, or emotional well-being, and whether there is a difference in how individuals engage with these apps depending on the mental health focus or incorporated techniques. Although some developer-led studies have published results on the use of individual mental health apps deployed in real-world settings, to the best of our knowledge, no study has examined a large sample of mental health apps relying on independently collected data. This investigation is feasible by leveraging the big data commonly generated and stored by digital platforms that record user traffic in the wild [23,24]. Leveraging such data, this examination provides benchmarks of app usage in the real world, where the general public is expected to benefit from their engagement with unguided programs. This information could shed light on specific engagement problems and opportunities for new intervention development and may offer a resource for researchers and developers who want to study and compare their app performance with similar apps.

For this study, a panel provided objective aggregated nonpersonal data on user engagement with mobile apps to analyze patterns of mental health app usage. The three primary aims were to (1) describe common usage patterns of popular unguided apps based on available metrics, (2) identify patterns of user retention over the first 30 days after app installation, and (3) explore whether these patterns differ based on the app's mental health focus and primary incorporated techniques.

Methods

Search Strategy

The search strategy aimed at identifying the most-installed unguided apps targeting depression, anxiety-related problems, or mental health. We used keywords related to depression and anxiety because of the high prevalence of these conditions [25,26]. We also included mental health apps that focused on happiness or the enhancement of mental health (ie, mindfulness meditations) because our previous work identified them as highly popular mental health tools [27,28]. We conducted a systematic engine search of the Google Play Store in November 2018 using the following terms: "depression" OR "mood" OR "anxiety" OR "panic attack" OR "phobia" OR "social phobia" OR "PTSD" OR "posttraumatic stress disorder" OR "stress reduction" OR "worry relief" OR "OCD" OR "obsessive compulsive disorder" OR "mental health" OR "emotional well-being" OR "happiness." One researcher documented all the apps emerging from the first 100 search results of each keyword, removed duplicates, and sorted them alphabetically. We also included a manual search of apps presented on MindTools.io [27] and PsyberGuide [29].

App Screening and Inclusion Criteria

Determining Apps' Number of Installs Threshold

To avoid including apps without a representative number of users, and to determine a minimum threshold for inclusion, we assessed the install categories presented by Google Play based on the number of app installs (eg, 10,000, 50,000 installs). Table 1 presents a preliminary analysis of the number of identified apps in each install category and the aggregated minimum number of app installs and corresponding percentages. Included apps had at least 5000 installs after removing any nonrelevant apps based on their title (ie, apps that were clearly not targeted at emotional well-being such as Heart Rate Monitor & Pulse Checker, 7 Minute Workout, 30 Day Fitness Challenge). Adding all the apps in the 5000 installs category would have resulted in a less than 0.5% increase in the total sample of users. Therefore, we determined an inclusion threshold of 10,000 app installs. Table 1 also shows that a small number of apps within the higher install categories were responsible for the most app installs. To make sure that including a large portion of apps with a relatively smaller number of installs (eg, <10,000 app installs) would not bias the results, we also examined whether there was a difference in the pattern of results based on the number of app installs. This will be further explained in the data analysis section.

Table 1. Analysis of install categories based on the number of apps in each category.

Install category	Apps identified, n	Minimum identified app installs within this category[a], n	Cumulative frequency of app installs based on category threshold[b], n	Added percentage of installs to the overall sample[c], %
≥10,000,000	2	20,000,000	20,000,000	100.00
5,000,000-9,999,999	6	30,000,000	50,000,000	60.00
1,000,000-4,999,999	21	21,000,000	71,000,000	29.58
500,000-999,999	23	11,500,000	82,500,000	13.94
100,000-499,999	69	6,900,000	89,400,000	7.72
50,000-99,999	33	1,650,000	91,050,000	1.81
10,000-49,999	103	1,030,000	92,080,000	1.12
5000-9999	66	330,000	92,410,000	0.36

[a] The number of apps multiplied by the minimum number of installs based on the install category.

[b] The accumulated number of app installs in all install categories above and including the current install category.

[c] The added percentage of installs to the total sample if the current install category is added to the analysis; it represents the percentage of the total number of app installs within this category divided by the accumulated number of app installs based on the current category threshold.

Inclusion and Exclusion Criteria

To be included in this review, apps had to:

1. Be in English;
2. Have at least 10,000 installs documented on Google Play;
3. Focus on mental illness, mental health, or emotional well-being not specifically related to another medical condition (for example, we excluded apps specifically focused on stress reduction due to a physical medical issue such as heart attack); and
4. Incorporate recognized techniques aimed at promoting self-management of mental health problems such as coping with negative symptoms (eg, feeling nervous, loss of energy), achieving positive results (eg, feeling better), or symptom management (eg, mood tracking). We excluded apps focused on the incorporation of sham techniques (see Multimedia Appendix 1 for a definition of sham techniques).

We excluded apps that:

1. Required payment for installation or provided a free trial only for a limited amount of time because it would be expected to bias program usage (free to install apps that included in-app purchases were not excluded);
2. Were therapist-based (eg, telepsychiatry) because the study was focused on unguided interventions; and
3. Were not meant to be used for more than a few times (eg, tests, one-time exposure technique) or were merely magazines.

Two independent reviewers screened the apps based on the inclusion and exclusion criteria. All disagreements were discussed with a third author with reference to the apps until consensus was reached.

Coding

Two independent reviewers coded the apps' incorporated techniques based on the following categories: mindfulness/meditation, tracker (including diary or journal), psychoeducation, peer support, and breathing exercise (not exercised as part of a meditation program). These categories were based on previous work done on the therapeutic components of mental health apps [27,30], drawing on the thematic analysis method suggested by Braun and Clarke [31]. The categories were designed to represent nonoverlapping components of potential therapeutic engagement (see Multimedia Appendix 2 for definitions of categories). Although our goal was to identify how specific techniques related to patterns of app use, our metrics did not enable us to differentiate between various techniques incorporated within the same app (ie, we could not tell which parts in the app the users were using). Therefore, we also added a coding of "primary technique" in cases where the app mostly incorporated one technique that was deemed to be the main reason for the app's use (eg, mindfulness/meditation). It is important to note that this limitation did not enable us to include app features that might influence user engagement but were not identified as a primary incorporated technique. Similarly, it was not feasible to target specific theoretical modalities, such as cognitive behavioral therapy. Because nearly all apps included some components of cognitive behavioral therapy, these were impossible to dismantle given our data.

An app's mental health focus was determined in the following manner: first, the app's description had to explicitly state that it targeted people with [mental health focus] and, second, most of the techniques used within the app had to have been built to help users cope with or manage their symptoms directly related to the mental health focus. We grouped apps based on several mental health foci. Under "mental health problems," we included apps that were focused on supporting people coping with depression, anxiety-related disorders, and emotional difficulties. We also subcoded the app with the terms (a) anxiety-related disorders or (b) depression if the app specifically targeted only one of these aims. (During our coding process, we did not identify another theme for the remaining apps.) Under "happiness," we included apps that focused on nurturing happiness or general positivity (eg, exercising gratitude, happiness assessment, suggestions for activities nurturing positive feelings), rather than the management of mental health states or problems.

During our coding process, we found a greater ambiguity around the description of apps with a primary incorporated technique of mindfulness/meditation, which leaned more toward enhancing emotional well-being (ie, helping users achieve a positive sense of experience and good mental health), but also aimed at stress reduction. Therefore, we grouped mindfulness and meditation apps separately and did not attribute either of the two mental health foci to them. For this reason, and to enable a proper comparison between categories, we present the mindfulness/meditation category in both the mental health focus and technique outcomes, despite being the same results.

A Cohen kappa interrater agreement of .92 was obtained for coding the variables of interest (incorporated technique, primary technique, and mental health focus). All disagreements were discussed with a third author with reference to the apps until consensus was reached.

Behavioral Data on User Engagement in the Real World

Information on user traffic was obtained from SimilarWeb's Pro panel data [32]. The panel provides aggregated nonpersonal information on user engagement with websites and mobile apps all over the world to enable Web and mobile app traffic research and analytics. The panel is based on several sources of anonymized usage data, such as data obtained from consenting users of mobile apps (ie, products). A dedicated product team at SimilarWeb is responsible for building and partnering with hundreds of high-value consumer products that make up the panel. According to SimilarWeb, the products are used across diverse audiences, without cluttering the user with advertisements. While benefiting from the products, users contribute to the panel because they enable the documentation of their online or mobile app usage activities seamlessly and anonymously [32]. The data are not used by SimilarWeb or provided to any third parties for the purposes of marketing, advertising, or targeting of individual subjects. The data-gathering procedures comply with data privacy laws, including the way data are collected, anonymized, stored, secured, and used. These procedures are updated regularly based on evolving data privacy legislation and requirements, such as the European Union's General Data Protection Regulation [33].

Our examination of data validity was tested and presented in a previous study [28]. An Oath researcher [34] (RW) examined 30 randomly selected mobile apps with data on SimilarWeb and usage data in Oath's independent records. The researcher examined the correlation between the average number of user sessions per day in the two datasets, finding a very strong Spearman correlation (N=30, r=.77, P<.001). In our study, we also examined the Spearman correlation between app install categories presented on Google Play (eg, 10,000, 50,000) and the number of downloads documented on SimilarWeb, and found a very strong correlation (N=93, r=.81, P<.001). These findings suggest a sufficient convergent validity, which is recommended to be above .70 [35].

The study was approved by University of Haifa Institutional Review Board, Haifa, Israel. The measures were set to include data gathered over a 12-month period from August 1, 2017, to July 31, 2018. For each app, available metrics on the panel included app open rate (the average percentage of daily active users out of the total sample of people who currently had the app installed), average number of sessions in a day per daily active user, and average daily minutes of use per daily active user. User 30-day retention included the percentage of users who opened the app each day between day 1 and day 30 out of the number of users who installed and opened the app on day 0. Usage patterns by time were available only for apps with a very large number of users. It was represented by two metrics—average percentage of use per hour (24 hours; eg, 7:00 am, 8:00 am) and per day (7 days; eg, Sunday, Monday)—both calculated based on total app usage.

Data Analysis

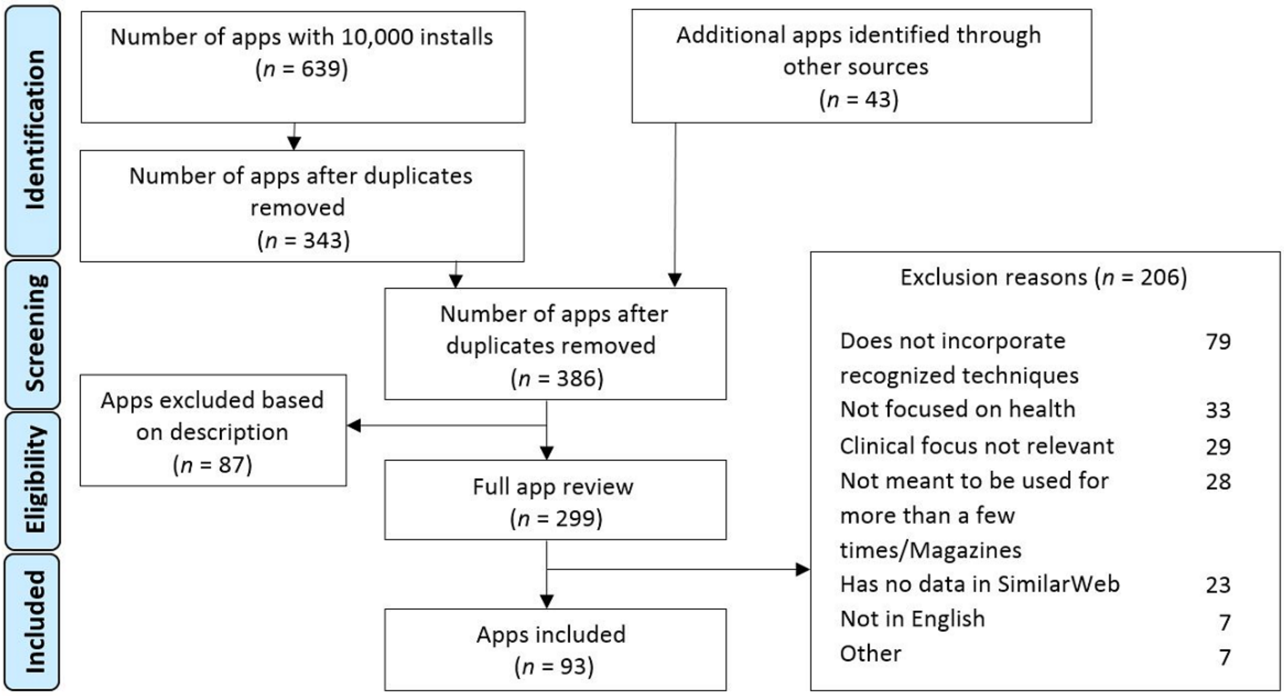
We did not assume a normal distribution of the metrics; therefore, medians and interquartile ranges (IQRs) were used as descriptive statistic measures. In cases in which a category included a small number of apps (n≤5), we used range instead of IQR. To examine differences in usage metrics between apps with different mental health foci or techniques, a Kruskal-Wallis one-way analysis of variance (ANOVA) was performed, followed by Mann-Whitney *U* tests to identify the source of the difference. To examine dependencies in the distribution of categorical values in relevant cases, we used chi-square tests. Most app installs came from a small number of apps with a large number of installs (see Table 1), so we conducted a sensitivity analysis to examine whether including apps with a smaller number of installs would bias the results. Mann-Whitney *U* tests were conducted to compare the distributions of the usage patterns for the top 5 installed apps and the remaining apps from each category presented in the results section (and that included more than five apps). We picked the top 5 apps based on their install category in Google Play. In cases in which several apps "competed" for the fifth place in the same install category, the app with the higher number of downloads (as documented in the SimilarWeb user panel) was chosen.

Results

Screening

Figure 1 presents the app inclusion flow diagram. The engine search and manual searches produced a total of 386 apps with 10,000 installs or more. Through the first screening process, 299 apps were identified and accessed for a detailed evaluation, and 93 apps were finally included in this study analysis (see Multimedia Appendix 3 for a complete list of included apps).

Figure 1. App inclusion flow diagram.



Description of Apps

The mental health focus of 59 (63%) apps was a mental health problem. Of these, 19 focused specifically on anxiety-related disorders and 4 focused specifically on depression. In addition, 8 (9%) apps focused on happiness, and 26 (28%) apps focused on the enhancement of emotional well-being through mindfulness/meditation. The distribution of apps based on incorporated techniques is presented in Table 2. Overall, 60 of 93 (65%) apps had a primary incorporated technique, and 33 (36%) apps had two or more incorporated techniques, none of which were primary. Mindfulness/meditation was the most frequent technique as the primary technique of the app (26/93, 28%), followed by use of a tracker (22/93, 24%). Psychoeducation (35/93, 38%) was the most frequent salient technique to be used not as the primary technique, followed by use of a tracker (28/93, 30%).

Table 2. Distribution of incorporated techniques in the app sample (N=93).

Incorporated technique	Primary technique, n (%)	Cotechnique[a], n (%)	Total, n (%)
Mindfulness/meditation	26 (28)	14 (15)	40 (43)
Tracker	22 (24)	28 (30)	50 (54)
Breathing exercise	7 (8)	20 (22)	27 (29)
Psychoeducation	3 (3)	35 (38)	38 (41)
Peer support	2 (2)	7 (8)	9 (1)

[a] The technique is saliently presented in the app but is not considered a primary technique.

App Usage by Daily Active Users

All apps had complete metrics on app usage by daily active users. Medians and IQRs of daily app usage are presented in Table 3 based on the app's mental health focus and in Table 4 based on the app's incorporated techniques. As shown in Table 3, the median app open rate was 4.0% (IQR 4.7%), with medians of 3.28 (IQR 2.53) daily sessions and 13.03 (IQR 14.27) minutes of app use per active user. Daily active usage of mindfulness/meditation apps (median 21.47, IQR 15.00) was found to be significantly higher than the usage of apps for mental health problems (median 10.02, IQR 10.60; z=4.64, P<.001) or for happiness (median 7.77, IQR 6.90; z=3.82, P<.001). No other significant difference in app usage was found between mental health foci, including between anxiety- and depression-related apps. As seen in Table 4, the number of app minutes of use was significantly higher for mindfulness/meditation (median 21.47, IQR 15.00) and peer support (median 35.08, n=2) than for other techniques (all z ≥2.11, all P<.05). In addition, tracker (median 6.3%, IQR 10.2%) and peer support (median 17.0%, n=2) apps had significantly higher open rates than breathing exercise apps (median 1.6%, IQR 1.6%; all z ≥3.42, all P<.001). No significant differences in usage patterns were found for apps without a primary strategy that incorporated more than one technique.

Table 3. App usage based on app mental health focus (N=93).

Mental health focus	Apps, n	Installation category, median (IQR)	Open rate (%), median (IQR)	Daily number of sessions per active users, median (IQR)	Daily minutes of use per active user, median (IQR)[a]
All apps	93	100,000 (90,000)	4.0 (4.7)	3.28 (2.53)	13.03 (14.27)
Mental health problems	59	50,000 (90,000)	4.0 (5.1)	3.77 (3.15)	10.02 (10.60)*
Anxiety	19	10,000 (40,000)	2.6 (2.5)	3.58 (3.49)	08.17 (09.42)
Depression	4	100,000 (50,000-100,000[b])	4.8 (3.0-6.8[b])	5.22 (3.97-6.55[b])	06.97 (02.05-15.12[b])
Happiness	8	100,000 (50,000)	3.7 (5.3)	3.50 (4.18)	7.77 (6.90)*
Mindfulness/meditation[c]	26	100,000 (650,000)	4.1 (3.3)	2.96 (1.66)	21.47 (15.00)**

[a] Categories with different number of asterisks (*, **) within a column are significantly different (P<.05) based on our analytical approach, which included Kruskal-Wallis one-way ANOVA at the variable level, followed by Mann-Whitney U tests.

[b] Due to a small number of included apps, brackets in this cell reflect the range (minimum-maximum value) and not the IQR.

[c] Mindfulness/meditation is presented as a separate mental health focus because all apps in this category were not attributed to another focus as they focus on enhancement of well-being as well as stress reduction.

Table 4. App usage based on app incorporated technique (N=93).

Incorporated technique	Apps, n	Installation category, median (IQR)	Open rate (%), median (IQR)[a]	Sessions per active user, median (IQR)	Daily minutes of use per active user, median (IQR)[a]
Primary technique					
Mindfulness/meditation	26	100,000 (650,000)	4.1 (3.3)	2.96 (1.66)	21.47 (15.00)*
Tracker	22	50,000 (90,000)	6.3 (10.2)*	4.58 (4.47)	07.27 (08.83)**
Breathing exercise[b]	7	10,000 (40,000)	1.6 (1.6)**	2.19 (1.23)	08.32 (19.02)**
Psychoeducation	3	10,000 (10,000-100,000[b])	3.0 (2.5-3.3[c])	4.16 (2.57-4.80[c])	03.53 (02.07-19.23[c])**
Peer support[d]	2	300,000 (N/A[e])	17.0 (N/A)*	8.67 (N/A)	35.08 (N/A)*
Number of primary techniques					
2 techniques	17	50,000 (90,000)	4.0 (5.6%)	3.18 (1.40)	07.83 (11.93)
≥3 techniques[f]	16	100,000 (50,000)	3.2 (3.1%)	4.06 (3.91)	12.88 (07.13)

[a] Categories with different number of asterisks (*, **) within a column are significantly different (*P*<.05) based on our analytical approach, which included Kruskal-Wallis one-way ANOVA at the variable level, followed by Mann-Whitney *U* tests.

[b] Not including mindfulness/meditation.

[c] Due to the small number of included apps, brackets in this cell reflect the range (minimum-maximum value) and not the IQR.

[d] Due to the small number of included apps, IQR or range could not be calculated (marked with N/A).

[e] N/A: not applicable.

[f] Includes two apps that use a chatbot (Wysa, Woebot), which did not have a different pattern of results emerging for a certain direction.

User 30-Day Retention

Fifty-nine apps (63%) had data on user retention. Chi-square tests for independence revealed no difference between apps with or without user retention data in the distribution of mental health foci (χ^2_2=2.1, P=.36) and primary incorporated techniques (χ^2_4=3.8, P=.44). Figure 2 presents user 30-day retention by the app's mental health focus; Figure 3 presents user 30-day retention by the app's incorporated technique. In both figures, there is a sharp decline of more than 80% in app open rates between day 1 and day 10, whereas the differences between day 15 and day 30 are smaller and represent a decline of approximately 20% in app open rates. Figure 2 reveals that, relative to users who opened the app on day 0, the median app open rate was as follows: 69.4% (IQR 27.8%) of users opened it on day 1, 3.9% (IQR 10.3%) of users opened it on day 15, and 3.3% (IQR 6.2%) of users opened it on day 30. Kruskal-Wallis one-way ANOVAs revealed no significant differences in app open rates on day 30 based on mental health focus (H_2=1.88, P=.39) and a significant difference in app open rates on day 30 based on incorporated technique (H_5=11.31, P=.046). Mann-Whitney *U* tests revealed that on day 30 peer support (median 8.9%), mindfulness/meditation (median 4.7%, IQR 6.2%), and tracker/diary apps (median 6.1%, IQR 20.4%) had significantly higher retention rates than breathing exercise apps (median 0.0%, IQR 0.0%; all z ≥2.18, all P ≤.04). This pattern of difference is also descriptively apparent in 15-day retention, in which the median retention for breathing exercise apps was 0.0% (IQR 0.0%), whereas the range of medians for peer support, mindfulness/meditation, and tracker/diary apps was from 4.9% (IQR 7.1%) to 11.9% (IQR 0.7%).

Figure 2. App 30-day retention by mental health focus. The percentages reflect the number of users who opened the app from day 1 to day 30 out of the number of users who installed and opened the app on day 0.

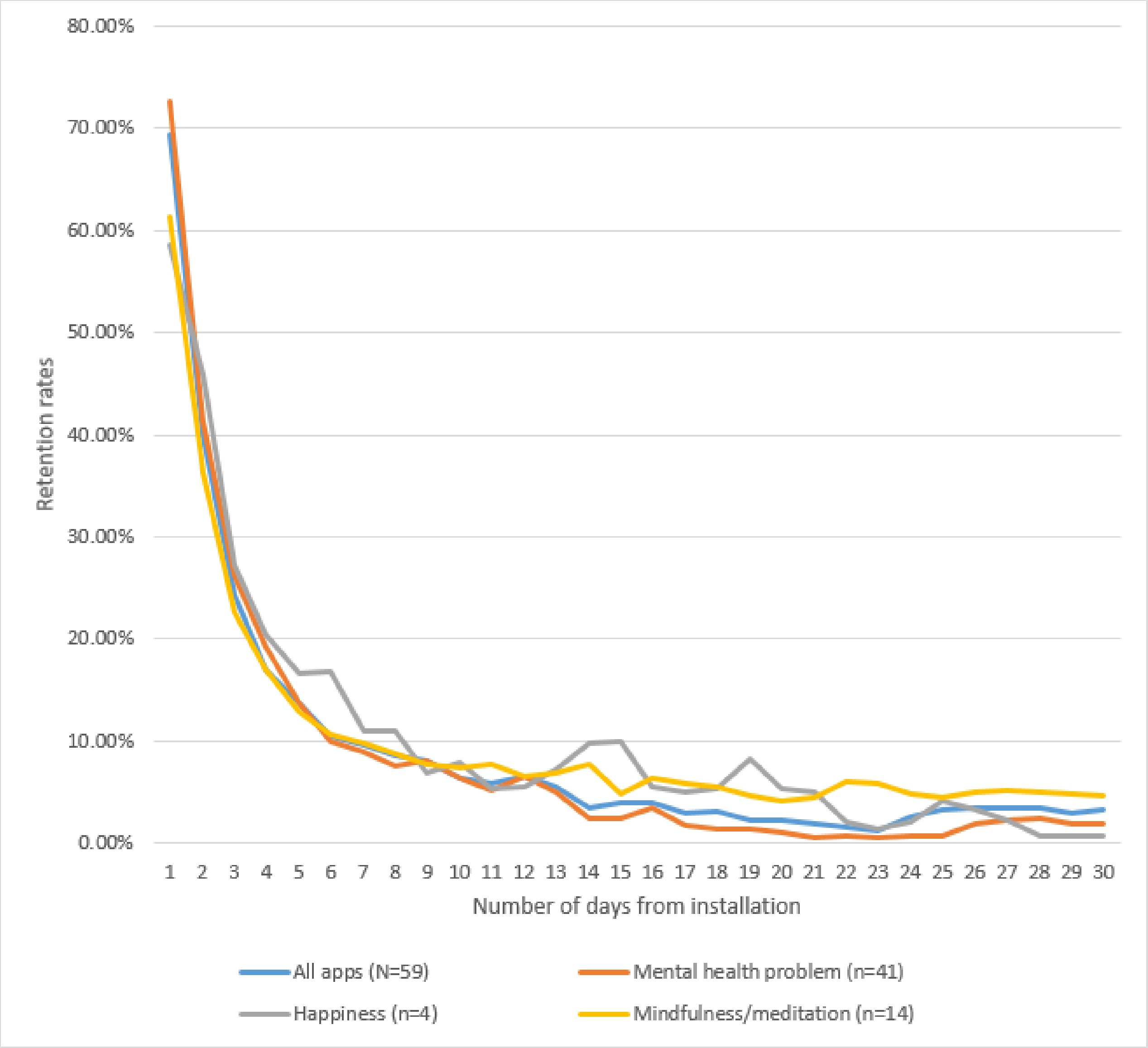
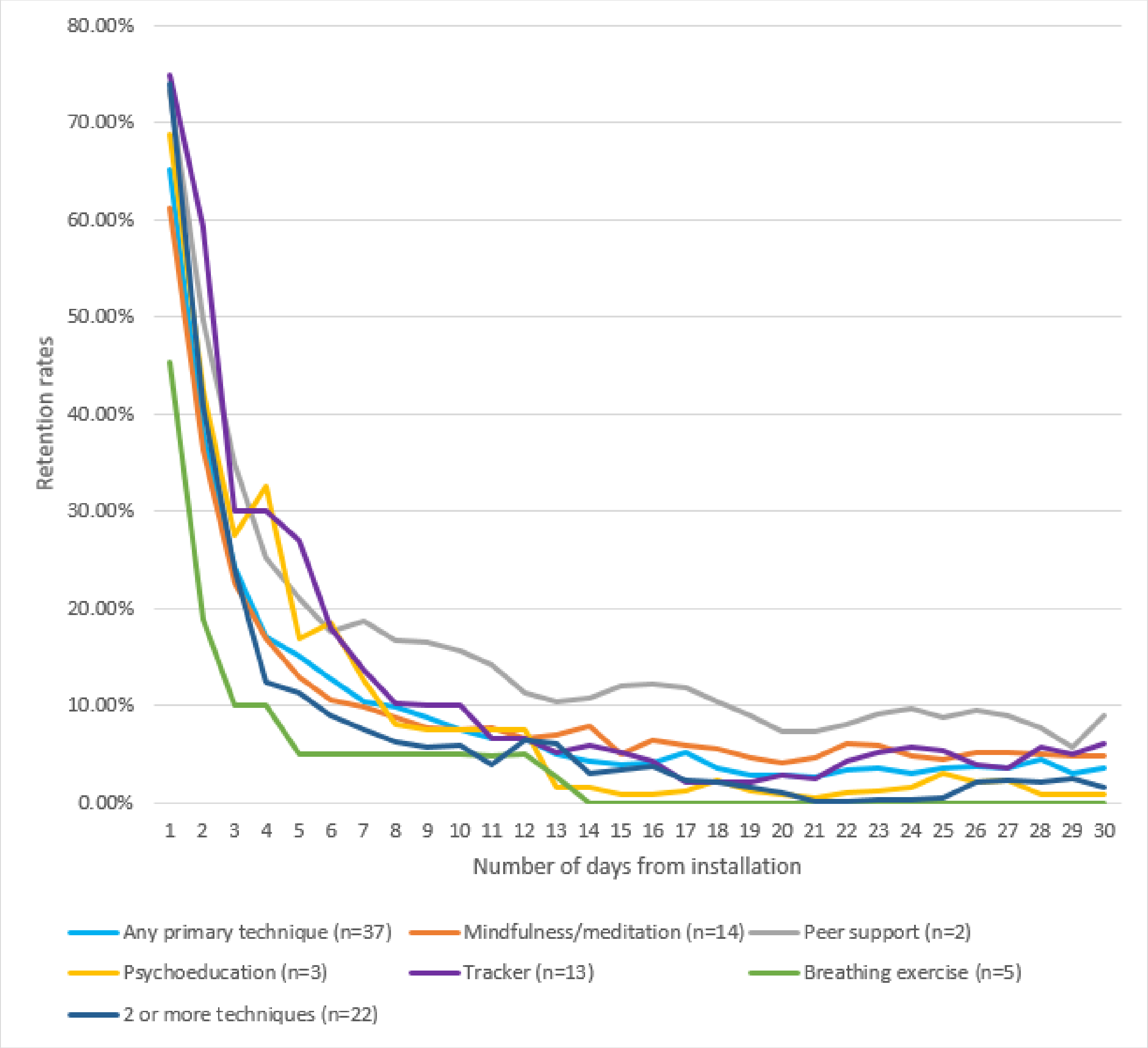


Figure 3. App 30-day retention by primary incorporated technique. The percentages reflect the number of users who opened the app from day 1 to day 30 out of the number of users who installed and opened the app on day 0.



Usage Pattern by Hours and Days

Sixteen apps had data on hourly and daily app usage. Figure 4 presents the hourly usage patterns of apps and Figure 5 presents the daily usage patterns of apps. The number of apps with available data was small; therefore, we only present categories with data on more than three apps. Furthermore, we have not conducted statistical testing to compare program usage among the different categories. For hourly usage, the results pointed to a peak in app usage in the evening (8:00 pm) for apps targeting mental health problems. The results also showed that mindfulness/meditation apps had two usage peaks: one in the morning (7 am-9 am) and the other in the late evening (10 pm-midnight). In terms of daily usage, the results showed a peak in app usage on Thursday for mindfulness/meditation apps.

Sensitivity Analysis

We conducted a series of Mann-Whitney *U* tests to examine the difference in app open rate, number of sessions, daily minutes of use, and 30-day retention among the top 5 installed apps and the remaining apps per mental health focus and incorporated technique. We found a significant difference in the open rate of mental health apps favoring the top 5 installed apps (z=1.68, $P \leq .05$; top 5 installed apps: median 9.0%, IQR 6.9%; remaining apps: n=54, median 4.0%, IQR 4.7%). Among these five apps, one incorporated online peer support and three incorporated mood trackers. No other differences were found. A series of Mann-Whitney *U* tests was also conducted to examine whether app usage (app open rates, daily number of sessions, daily minutes of use) in each app category (mental health focus, incorporated technique) differed between apps with or without in-app purchases and no significant differences were found (all P>.05).

Figure 4. Hourly usage pattern. Usage is presented by hour out of the total app usage; therefore, the sum of percentages within each category is 100%. Note: a subset of apps for which that data were available is included; "All apps" includes both categories and one app targeting happiness.

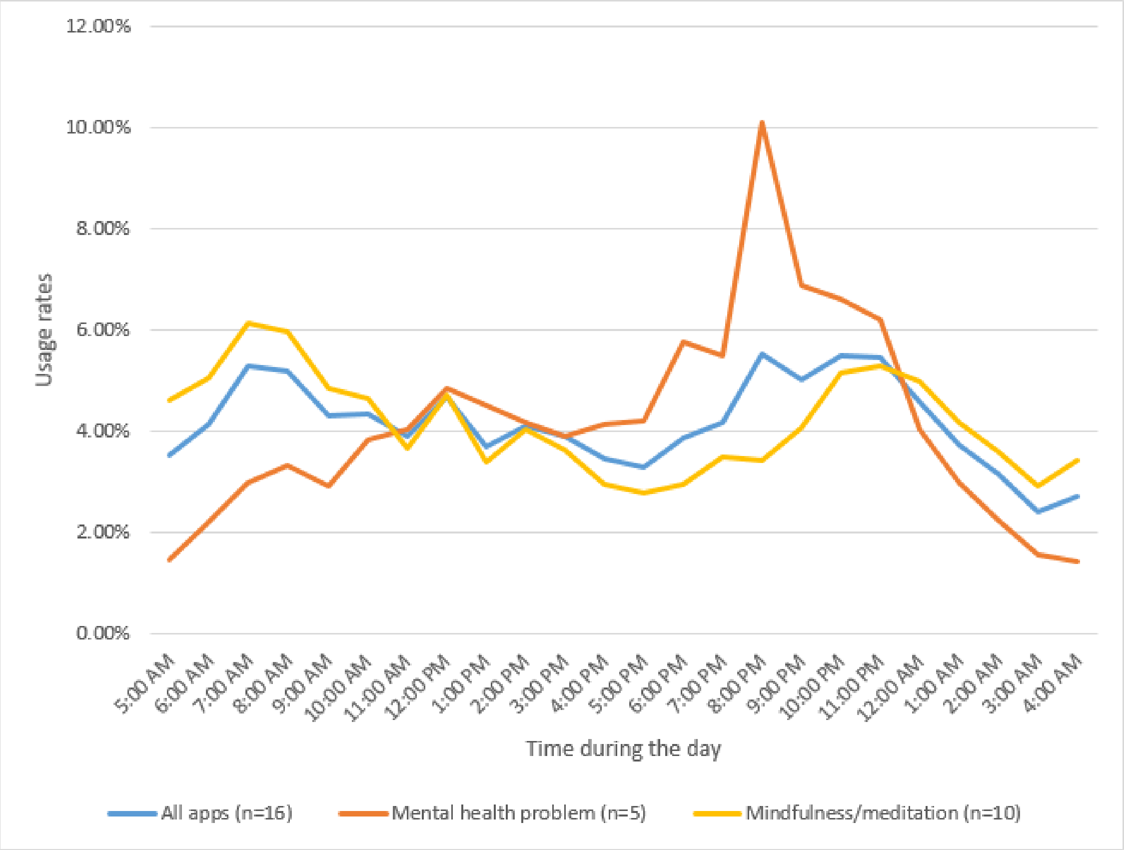


Figure 5. Daily usage pattern. Percentage of app usage is presented by day out of the total app usage; therefore, the sum of percentages within each category is 100%. Note: a subset of apps for which that data were available is included; "All apps" includes both categories and one app targeting happiness.

Discussion

Principal Findings

This is the first study to report the usage and retention metrics of a large number of frequently installed, unguided mental health apps as recorded “in the wild” and independent of developer-led data. Based on Google Play Store data (using keyword search terms), there were over 90 million mental health app installs documented by the end of 2018 (ie, reach). Although our findings revealed that daily active users use apps for a significant amount of time during the day (daily usage median of 13.03 minutes), most people with the app installed on their device do not open it in any given day (median open rate of 4.0%). Furthermore, general user retention is poor, with a median 15-day retention of 3.9% and 30-day retention of 3.3%. These findings reflect the lower ranges of real-world retention rates reported in developer-led studies [17-20,22].

Our results also indicate that there are significant differences in app usage and user retention that are associated with the app’s incorporated techniques. Daily minutes of use were significantly higher for mindfulness/meditation (median 21.47) and peer support (median 35.08) apps than for apps incorporating other techniques. Daily open rates were significantly lower for breathing exercise apps (median 1.6%) than for apps incorporating the two techniques with the highest open rates (tracker: median 6.3%; peer support: median 17.0%). User 30-day retention was significantly lower for breathing exercise apps (median 0.0%) than for all other incorporated techniques (mindfulness/meditation: 4.7%; trackers: 6.1%; peer support: 8.9%), except for psychoeducation, which exhibited a pattern similar to the breathing exercise apps at 30-day retention. These patterns could be explained using the notion of *effective engagement* described by Yardley and colleagues [36], wherein there is “sufficient engagement with the intervention to achieve intended outcomes.” From this perspective, it might be that once people acquire the desired skills (breathing exercise) or knowledge (psychoeducation) they no longer use the app, thus affecting the pattern of retention over a longer period. By contrast, mindfulness/meditation apps often include guided meditations designed for repeated use over longer periods of time, while not fostering learning or direct skill acquisition.

Our findings on user retention highlight the low engagement with these apps. Although this warrants a re-evaluation of current engagement and retention strategies, it does not necessarily suggest that these apps are only helpful for a small number of users. First, we do not have data implying that users engage only with one app in the self-management of their states or conditions. However, it is difficult to assume that users are knowledgeable about the different apps available, which apps to use, and when to use them. Although there are some recommender websites [27,29,37] and approaches to help users identify the right apps [38-41], a therapeutic framework that provides guidance to users about how to use the right app at the right time could be useful. For example, in their novel study of IntelliCare—a suite of 13 apps and one Hub app accompanied by 8 weeks of coaching to encourage participants to try the apps recommended to them through the Hub app—Mohr and colleagues [42] found that 95% of participants eventually downloaded five or more of the IntelliCare apps as part of their therapeutic process. In another study, patients with schizophrenia spectrum disorders received 6 months of treatment that included health technology coaching around the use of three digital tools that were offered to patients based on their needs; 96% of patients rated the program as beneficial [43]. Future studies are needed to examine the feasibility of executing a scalable framework of care in which users receive the right app recommendation at the right time as part of a self-management routine.

Second, user retention patterns might also indicate the low burden associated with app installation (ie, the simplicity of opening the Google Play Store and clicking the app download and installation buttons), which implies that user context, motivation, and ability to engage [44] with these apps were not tested before app installation. The poor active user rates found in our analysis (median open rates of 4%) suggest that the number of app installs available in app stores do not provide a proper estimation of the proportion of users who actually self-manage their state by using the app. These issues further justify a previous call for the development of models to conceptualize the relationships between user state, need, ability, and motivation to engage with early interventions in the digital public space [8]. Although we need to significantly improve our ability to engage users who have made initial attempts at help-seeking, taking a public health engagement approach that is also focused on sustainability represents an important step forward in scaling effective care.

Finally, we identified that the two apps that incorporated peer support as a primary technique had relatively high engagement and retention rates. In our previous work, we defined a program’s *relatability* as “a good representation of a human factor that is easily relatable within the therapeutic context/process” [38]. Relational factors have also been previously acknowledged to nurture a therapeutic alliance with users [45-47], and have demonstrated to be a quality aspect that predicts user engagement with mobile health interventions [28]. Future studies are needed to determine whether technology has a special advantage as an infrastructure that connects between users and results in better engagement rates.

Limitations

This study has several limitations that should be considered. First, because we used an anonymous user panel, we did not have data about how different users use the apps and which parts of the apps were more engaging. The absence of such data means that some apps might have been more engaging due to the characteristics of their users, a phenomenon suggested previously by Ernsting and colleagues [48]. In addition, due to this limitation we were only able to focus on primary incorporated techniques within the apps and not on the way different design features (not deemed to be a primary technique) may have impacted the results. Subsequently, because we were leaning on off-the-shelf programs available to the public, we could not manipulate the programs themselves to account for aspects which lacked variability in our data, such as the impact of theoretical modalities on usage. That is, although our study advantage is that it enables us to present benchmarks of real-world use independent to trial settings, one advantage of direct experiments is the ability to control participant identity and manipulate intervention modalities and features to identify the group of active components leading to the best outcome (eg, [49]). Such experiments could be also helpful in determining causal relationships between intervention modalities and user behaviors, based on the context of use.

Second, some techniques such as peer support were only incorporated by a small number of highly installed apps (median installation category of 300,000). However, our results did not indicate a significant difference in any incorporated technique in terms of app installs, which suggests that these apps usage patterns go beyond an app's popularity.

Third, because we were led by the available metrics on the platform, we could not examine retention rates after the first 30 days. The retention slope presented a slower decline in app open rates between day 15 and 30 and, based on previous reports, it would be reasonable to assume that there is a continuous usage decline over time (eg, [19,50]), but more studies are needed to determine the magnitude of the decline.

Finally, this study was only based on Android users. Current estimates suggest that the Android market share is approximately 88% of mobile phone users globally [51] and approximately 42.7% of mobile phone users in the United States [52]. Although these data suggest that a sufficient portion of users use the Android operating system, it would be beneficial to validate these results with datasets from the Apple market.

Conclusions

The use of digital platforms that record user traffic "in the wild" enables us to examine patterns of app usage outside of study settings and to assess real-world public engagement. Although we found daily active minutes of use to be relatively high, only a small portion of users actually used popular apps regularly. More studies leveraging different datasets are needed to understand these phenomena. On a broader level, findings point to the importance of the ways we measure, report, and address aspects of user engagement in the real world. It would be helpful to track the context of users who eventually use apps, hopefully through the use of digital footprints, while also tracking the use of multiple apps and websites across times. Obviously, aspects that relate to security and privacy of data have to be addressed. In addition, new studies are needed to better conceptualize our understanding of users' contexts and the ways they search for and engage with beneficial services outside of traditional health care settings.

Acknowledgments

This study was supported by the Donald & Barbara Zucker Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definition of sham techniques.
[PDF File (Adobe PDF File)75 KB - jmir_v21i9e14567_app1.pdf]

Multimedia Appendix 2
Definition of coded techniques.
[PDF File (Adobe PDF File)62 KB - jmir_v21i9e14567_app2.pdf]

Multimedia Appendix 3
List of included apps.
[PDF File (Adobe PDF File)72 KB - jmir_v21i9e14567_app3.pdf]

References

1. Krishna S, Boren S, Balas E. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
2. 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]
3. 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;24(5):321-332 [FREE Full text] [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
4. Pew Research Center. 2019 Jun 12. Mobile fact sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-01-07] [WebCite Cache ID 75F2i2NKz]

5. WalkerSands Communications. 2019. How technology is expanding the scope of online commerce beyond retail URL: https://www.walkersands.com/resources/the-future-of-retail-2018/ [accessed 2019-01-07] [WebCite Cache ID 75F2nwClg]
6. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 7;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
7. Kazdin A. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. Behav Res Ther 2017 Jan;88:7-18. [doi: 10.1016/j.brat.2016.06.004] [Medline: 28110678]
8. Baumel A, Baker J, Birnbaum ML, Christensen H, De Choudhury M, Mohr DC, et al. Summary of key issues raised in the Technology for Early Awareness of Addiction and Mental Illness (TEAAM-I) meeting. Psychiatr Serv 2018 May 01;69(5):590-592. [doi: 10.1176/appi.ps.201700270] [Medline: 29334875]
9. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
10. Christensen H, Mackinnon A. The law of attrition revisited. J Med Internet Res 2006;8(3):e20. [Medline: 17032636]
11. Baumeister R, Vohs K. Self-regulation, ego depletion, and motivation. Social Pers Psych Compass 2007 Nov;1(1):115-128. [doi: 10.1111/j.1751-9004.2007.00001.x]
12. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
13. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research…and three things we can do about them. Psychiatr Serv 2017 May 01;68(5):427-429. [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
14. Day J, Sanders M. Do prents benefit from help when completing a self-guided parenting program online? A randomized controlled trial comparing Triple P Online with and without telephone support. Behav Ther 2018;49(6):1020-1038.
15. Glasgow RE. eHealth evaluation and dissemination research. Am J Prev Med 2007 May;32(5 Suppl):S119-S126. [doi: 10.1016/j.amepre.2007.01.023] [Medline: 17466816]
16. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. J Med Internet Res 2005;7(1):e7 [FREE Full text] [doi: 10.2196/jmir.7.1.e7] [Medline: 15829479]
17. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J Med Internet Res 2004 Dec 22;6(4):e46 [FREE Full text] [doi: 10.2196/jmir.6.4.e46] [Medline: 15631970]
18. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. J Med Internet Res 2018 Jun 06;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]
19. Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the wild: using novel data to examine the reach, use, and impact of PTSD Coach. JMIR Ment Health 2015;2(1):e7 [FREE Full text] [doi: 10.2196/mental.3935] [Medline: 26543913]
20. Carpenter J, Crutchley P, Zilca RD, Schwartz HA, Smith LK, Cobb AM, et al. Seeing the "big" picture: big data methods for exploring relationships between usage, language, and outcome in internet intervention data. J Med Internet Res 2016 Aug 31;18(8):e241 [FREE Full text] [doi: 10.2196/jmir.5725] [Medline: 27580524]
21. Ebert D, Baumeister H. Internet-based self-help interventions for depression in routine care. JAMA Psychiatry 2017 Aug 01;74(8):852-853. [doi: 10.1001/jamapsychiatry.2017.1394] [Medline: 28636717]
22. Christensen H, Griffiths K, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13. [doi: 10.2196/jmir.1194] [Medline: 19403466]
23. Moller A, Merchant G, Conroy D, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. J Behav Med 2017 Feb;40(1):85-98 [FREE Full text] [doi: 10.1007/s10865-016-9818-7] [Medline: 28058516]
24. Baumel A, Yom-Tov E. Predicting user adherence to behavioral eHealth interventions in the real world: examining which aspects of intervention design matter most. Transl Behav Med 2018 Sep 08;8(5):793-798. [doi: 10.1093/tbm/ibx037] [Medline: 29471424]
25. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, WHO World Mental Health Survey Consortium. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. JAMA 2004 Jun 02;291(21):2581-2590. [doi: 10.1001/jama.291.21.2581] [Medline: 15173149]
26. World Health Organization. The World Health Report 2001. Mental Health: New Understanding, New Hope. Geneva: World Health Organization; 2001.
27. MindTools.io. URL: https://mindtools.io/ [accessed 2018-06-24] [WebCite Cache ID 70Pg0KTXJ]
28. Baumel A, Kane J. Examining predictors of real-world user engagement with self-guided eHealth interventions: analysis of mobile apps and websites using a novel dataset. J Med Internet Res 2018 Dec 14;20(12):e11491 [FREE Full text] [doi: 10.2196/11491] [Medline: 30552077]
29. PsyberGuide. URL: http://psyberguide.org/ [accessed 2015-04-10] [WebCite Cache ID 6Xh99eQmZ]

30. Baumel A, Birnbaum ML, Sucala M. A systematic review and taxonomy of published quality criteria related to the evaluation of user-facing eHealth programs. J Med Syst 2017;41(8):128. [doi: 10.1007/s10916-017-0776-6] [Medline: 28735372]
31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
32. SimilarWeb. 2018. URL: https://www.similarweb.com/ [accessed 2018-06-18] [WebCite Cache ID 70GZ6eAt5]
33. European Commission. Data protection URL: https://ec.europa.eu/info/law/law-topic/data-protection_en [accessed 2018-06-18] [WebCite Cache ID 70GY1yLYp]
34. Oath. URL: https://www.oath.com/ [accessed 2018-10-14] [WebCite Cache ID 73A1knb5w]
35. Carlson KD, Herdman AO. Understanding the impact of convergent validity on research results. Organ Res Methods 2010 Dec 30;15(1):17-32. [doi: 10.1177/1094428110392383]
36. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
37. ADAA Mental Health Apps rates. URL: https://www.adaa.org/finding-help/mobile-apps [accessed 2016-11-01] [WebCite Cache ID 6lhtKNvkc]
38. Baumel A, Faber K, Mathur N, Kane JM, Muench F. Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. J Med Internet Res 2017 Mar 21;19(3):e82 [FREE Full text] [doi: 10.2196/jmir.7270] [Medline: 28325712]
39. 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]
40. 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]
41. Torous JB, Chan SR, Gipson SY, Kim JW, Nguyen T, Luo J, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. Psychiatr Serv 2018 May 01;69(5):498-500. [doi: 10.1176/appi.ps.201700423] [Medline: 29446337]
42. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare:an eclectic, skills-based app suite for the treatment of depression and anxiety. J Med Internet Res 2017 Jan 05;19(1):e10 [FREE Full text] [doi: 10.2196/jmir.6645] [Medline: 28057609]
43. Baumel A, Correll CU, Hauser M, Brunette M, Rotondi A, Ben-Zeev D, et al. Health technology intervention after hospitalization for schizophrenia: service utilization and user satisfaction. Psychiatr Serv 2016 Jun 1;67(9):1035-1038. [doi: 10.1176/appi.ps.201500317] [Medline: 27247171]
44. Fogg B. A behavior model for persuasive design. 2009 Presented at: Persuasive '09 4th international Conference on Persuasive Technology; April 26-29, 2009; Claremont, CA. [doi: 10.1145/1541948.1541999]
45. Cavanagh K, Millings A. (Inter)personal computing: the role of the therapeutic relationship in e-mental health. J Contemp Psychother 2013 Jul 17;43(4):197-206. [doi: 10.1007/s10879-013-9242-z]
46. Holter MT, Johansen A, Brendryen H. How a fully automated eHealth program simulates three therapeutic processes: a case study. J Med Internet Res 2016 Jun 28;18(6):e176 [FREE Full text] [doi: 10.2196/jmir.5415] [Medline: 27354373]
47. Barazzone N, Cavanagh K, Richards DA. Computerized cognitive behavioural therapy and the therapeutic alliance: a qualitative enquiry. Br J Clin Psychol 2012 Nov;51(4):396-417. [doi: 10.1111/j.2044-8260.2012.02035.x] [Medline: 23078210]
48. Ernsting C, Dombrowski S, 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]
49. Collins L, Murphy S, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
50. Cheung K, Ling W, Karr C, Weingardt K, Schueller S, Mohr D. Evaluation of a recommender app for apps for the treatment of depression and anxiety: an analysis of longitudinal user engagement. J Am Med Inform Assoc 2018 Aug 01;25(8):955-962 [FREE Full text] [doi: 10.1093/jamia/ocy023] [Medline: 29659857]
51. Statistica. Global mobile OS market share in sales to end users from 1st quarter 2009 to 2nd quarter 2018 URL: https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/ [accessed 2019-01-06] [WebCite Cache ID 75DpPpzz6]
52. Statcounter GlobalStats. Mobile operating system market share United States of America URL: http://gs.statcounter.com/os-market-share/mobile/united-states-of-america [accessed 2019-01-06] [WebCite Cache ID 75Dr29YOS]

Abbreviations

ANOVA: analysis of variance
IQR: interquartile range

Edited by G Eysenbach; submitted 02.05.19; peer-reviewed by A Rotondi, B Chaudhry, L Vizer, MS Aslam, L Bonet; comments to author 31.05.19; revised version received 01.07.19; accepted 19.07.19; published 25.09.19.

Please cite as:
Baumel A, Muench F, Edan S, Kane JM
Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis
J Med Internet Res 2019;21(9):e14567
URL: http://www.jmir.org/2019/9/e14567/
doi:10.2196/14567
PMID:31573916

Original Paper

What Players of Virtual Reality Exercise Games Want: Thematic Analysis of Web-Based Reviews

Nuša Faric[1*], BSc, MSc; Henry W W Potts[2*], BA, MSc, PhD; Adrian Hon[3*], BA, MA; Lee Smith[4*], BSc, MSc, PhD; Katie Newby[5*], BSc, MSc, PhD; Andrew Steptoe[2*], BA, MA, DPhil, DSC; Abi Fisher[1*], BSc, PhD

[1] Behavioural Science and Health, University College London, London, United Kingdom
[2] University College London, London, United Kingdom
[3] Six to Start, LSD Accountants, London, United Kingdom
[4] The Cambridge Centre for Sport and Exercise Sciences, Anglia Ruskin University, Cambridge, United Kingdom
[5] Coventry Univesity, Coventry, United Kingdom
[*] all authors contributed equally

Corresponding Author:
Nuša Faric, BSc, MSc
Behavioural Science and Health
University College London
1-19 Torrington Place
London, WC1E 6BT
United Kingdom
Phone: 44 020776794466
Email: nfaric@gmail.com

Abstract

Background: Physical activity (PA) is associated with a variety of physical and psychosocial health benefits, but levels of moderate-to-vigorous intensity PA remain low worldwide. Virtual reality (VR) gaming systems involving movement (VR exergames) could be used to engage people in more PA.

Objective: This study aimed to synthesize public reviews of popular VR exergames to identify common features that players liked or disliked to inform future VR exergame design.

Methods: We conducted a thematic analysis of 498 reviews of the 29 most popular exergames sold in the top 3 VR marketplaces: Steam (Valve Corporation), Viveport (Valve Corporation), and Oculus (Oculus VR). We categorized reviews as positive and negative as they appeared in the marketplaces and identified the most common themes using an inductive thematic analysis.

Results: The reviews were often mixed, reporting a wide variety of expectations, preferences, and gaming experiences. Players preferred highly realistic games (eg, closely simulated real-world sport), games that were intuitive (in terms of body movement and controls), and games that provided gradual increases in skill acquisition. Players reported feeling that they reached a high level of exertion when playing and that the immersion distracted them from the intensity of the exercise. Some preferred features included music and social aspects of the games, with multiplayer options to include friends or receive help from experienced players. There were 3 main themes in negative reviews. The first concerned bugs that rendered games frustrating. Second, the quality of graphics had a particularly strong impact on perceived enjoyment. Finally, reviewers disliked when games had overly complex controls and display functions that evoked motion sickness.

Conclusions: Exergames prove to be a stimulating avenue for players to engage in PA and distract themselves from the negative perceptions of performing exercise. The common negative aspects of VR exergames should be addressed for increased uptake and continued engagement.

(J Med Internet Res 2019;21(9):e13833) doi:10.2196/13833

KEYWORDS

virtual reality; exercise; video games; sedentary lifestyle; weight loss; behavior; obesity; sports

Introduction

Background

There is a large body of evidence showing that regular and sustained participation in physical activity (PA) aids in the prevention and management of noncommunicable disease [1-5]. Age-specific PA recommendations have been formulated, for example, the United Kingdom (UK) government recommends at least 150 min per week of moderate-to-vigorous PA for adults and 60 min per day for children and adolescents aged 5 to 18 years [6]. However, global population levels of PA are insufficient [6-8]. The UK has particularly low levels of PA compared with other European countries [9]. Objective measures of PA in 4507 adults aged above 16 years showed that only 6% of males and 4% of females met the recommended guidelines for PA [10]. Furthermore, 50% of males and 58% of females were classified as *inactive* [10]. Levels of inactivity are particularly high among socioeconomically deprived groups [11,12]. The US Department of Health and Human Services reported that less than 5% of American adults engage in 30 min of PA every day and only one-third reach the recommended PA levels in a week [13]. An increase in sedentary lifestyle and insufficient PA has also been observed in other countries such as Australia [14], India [15], and China [16]. Interventions to increase levels of PA in the population are urgently required.

Gaming is a recreational activity regularly performed by 2 to 3 billion people worldwide across socioeconomically deprived groups and by males and females [17,18]. The popularity and nature of gaming makes it a potential avenue for a population health intervention [19]. A market research report by Limelight Networks in 2018 found that gamers around the world spent an average 6 hours in playing each week [20]. Games requiring bodily movement (exergames) offer a potential to increase PA and reduce sitting time or replace sedentary screen time [21]. Early generation exergames such as Dance Dance Revolution released in 1998 and Nintendo’s Wii Fit released in 2007, (which sold over 22 million copies worldwide) and, more recently, Pokémon Go (downloaded over 800 million times) are examples of exergames that have had huge commercial success [22,23]. Empirical studies of previous generation exergames have shown that they can achieve PA of at least moderate intensity and can increase energy expenditure up to 300% above resting levels [23]. In addition, a meta-analysis of 35 studies in young people found that they achieved a similar level of activity in field-based sports and enhanced reported enjoyment of, and intrinsic motivation for, PA [24]. Exergaming has also been positively associated with psychological well-being [25,26]. Randomized controlled trials in children [27], preadolescents [28], adolescents [29,30], and adults [31] found that exergames supported weight loss and increased fitness. However, trials have not been able to confirm exactly which psychological or social factors might lead to long-term engagement [21,25,26].

Virtual reality (VR) has the potential to enhance the impact of an exergaming intervention. The most prominent feature of VR is high-quality virtual realism or immersion, coupled with body tracking, where the person feels fully present or located in the virtual world [32,33]. In exergames, presence is not always required, but in VR, full presence is required and, as such, psychological processes may potentially differ and even influence the performance or affective states of players [33]. There is evidence from laboratory studies that immersive VR exergaming is more engaging than standard exercise and may distract participants from exertion. For example, one study involving 88 university staff and students found that although the heart rate was higher in VR than a standard exercise condition, participants reported feeling less tired and had higher ratings of enjoyment in the VR condition [34]. VR is moving from a niche to a more affordable product, suggesting, for the first time, potential for the use of VR as a population health intervention. Ownership is increasing: Statista reported worldwide sales of VR headsets totaling 13.5 million units in 2017, and the sales are predicted to increase drastically by 2020 because of decreases in price [35].

Objectives

VR exergames are a unique platform for PA intervention, and not much has been reported about the experiences of players themselves. This study aimed to investigate the factors that were reported to help or hinder engagement with VR exergames as reported by the players themselves in Web-based VR marketplaces or platforms. This review was performed to learn more about their actual experiences and synthesize results to be mindful when building a new VR exergame for PA. In line with intervention development frameworks, gathering the views of target users is crucial [32]. Some feedback can be obtained only from people who are using the technology in question, even though they may not be representative of all future users. Markets are changing rapidly and are future oriented, and the insights that we gather from current users may benefit the development of future VR PA interventions. The systems need to be well designed to be as effective as possible and preferably with an involvement with the target population [36].

New data collection could have been conducted with players of VR exergames, but we instead opted to take advantage of Web-based reviews that have been spontaneously submitted. Using material submitted Web-based has become increasingly common in research, and we drew on examples in app research [37]. This approach is quick and cost-effective. Web-based reviews may not directly answer researchers’ questions, and those submitting reviews may not be representative. However, Web-based reviews may have greater validity than responses given in a research context.

Methods

Main gaming databases were searched in July 2018. Steam, Viveport, and Oculus were selected as the 3 main marketplaces for VR games and gaming experiences. Steam is the most popular hub in the world for purchasing and downloading desktop and VR games [38]. At present, Steam supports HTC Vive, Oculus Rift, and Windows Mixed Reality headsets [39,40]. Viveport and Oculus support their respective VR headsets. All 3 marketplaces encourage users to share their reviews.

Virtual Reality Exergame Search

This study included games that were specifically marketed as being fitness related. Games that related to PA or sport but did not involve movement to play (eg, football management) were excluded. Each marketplace had slightly different game selection categories and or filter options, and these were applied to obtain a selection of VR exergames. The search strategies were as follows:

1. Steam: In the “Virtual Reality” category, “sort by Top Sellers” was selected and manually searched for fitness-related games.
2. We also visited the “Virtual Reality” category and selected “sort by What’s Being Experienced” (ie, hours played) and manually searched for any fitness-related games that were missed in the previous search.
3. Viveport: We searched for apps in the “Health & Fitness or Sports” genre, filtering by “Popularity”.
4. Oculus: We searched the fitness-related games in the categories “Top-selling experiences for Rift” and “Top Free experiences for Rift.”

Types of User Reviews and Sampling

For each of the sampled VR games, 2 researchers (NF and AF) extracted the most recent 10 positive and 10 negative reviews. The reviews were extracted into an Excel sheet.

Steam and Oculus reviews require users to post either a positive or negative review. Viveport allows reviewers to post reviews using a starring system, ranging from 1 to 5. For all 3 marketplaces, we selected the first 10 reviews (chronologically most recent). This posed an issue of whether to include reviews with 3 stars (which would constitute as *neutral* and not necessarily a *positive* or *negative* review). We decided to include these because Steam and Oculus platforms, in fact, included reviews that would fall under a *neutral* category if such options existed. For example, in Steam and Oculus, some reviewers provided a breakdown and a list of *pros*, *neutrals*, and *cons* within one review, whereas others directly reported how many stars or categories they would assign if such options were available: “I could not give a positive review and there’s no neutral ones”, and “I’m a bit torn on reviewing this [before listing both pros and cons and arguing it was a neutral review].” It was merely the fact that the review could not be posted unless they have chosen either a positive or negative category, but in either of the categories, they have reported what would otherwise constitute a neutral review.

Data Analysis

A thematic analysis was used as it provides a flexible approach to analyzing qualitative data [41]. We chose the thematic analysis over other diverse qualitative approaches because it helped us locate themes and patterns or *thematize meanings* within the body of data. Themes were organized in detail in a coherent manner [41]. The theoretical position of the thematic analysis in our study is concerned with a particular context and as such fell under a *contextualist* method or *critical realism*. It reports on the reality and experiences of the VR exergame users but at the same time unravels the way in which individuals make meaning of their experiences.

Results

Overview

The number of reviews extracted from each database is shown in Table 1. A total of 36 games were identified, but 7 of these had no associated reviews, leaving 29 games for inclusion. These produced a total of 498 reviews (260 positive, 238 negative, and 2 neutral reviews, ie, 3 stars for Viveport only). Table 2 gives details of the included games.

The VR exergames included in the analysis covered a wide variety of sport simulators (eg, Racket Fury) and super hero (eg, Echo Arena) or arcade-style games in VR (eg, Fruit Ninja). Of 498 reviews, 85 (17%) had both positive and negative comments within one review. This was apparent from the content of the reviews, and reviewers often separated feedback into defined paragraphs with *pros*, *neutrals*, and *cons* as subtitles in one review, as mentioned earlier. Approximately 70% of reviewers commented on the technological aspects of the games that included the settings and options in the games, the bugs, display functions such as the glitches and freezes, the overall flow of the game, the use of the controllers, and other potential issues (eg, the game requiring users to jump or spin while being potentially hazardous, eg, cables and space requirement).

Each review varied in length from 5 to 880 words. In terms of the context of this study, the word *workout* was mentioned 112 times in all reviews; *sweat* or *sweaty*, 61 times; *fitness*, 14 times; and *physical activity*, twice.

We identified 14 main themes for both positive and negative reviews, which we illustrate along with supporting quotes and summarize in Table 3.

Table 1. Selection of virtual reality exergames based on the marketplace.

Marketplace, game types	Number of games identified
Steam	
Top sellers	12
What’s being experienced (hours played)	6
Viveport	
Health and fitness or sports genre (ordered by popularity)	10 (7 had no reviews)
Oculus	
Top selling experiences for Rift	7
Top free experiences for Rift	1

Table 2. Final virtual reality exergame selection and description of each virtual reality exergame.

VR[a] exergame	Brief description of the VR exergame by marketplace (descriptions taken from or adopted from the marketplaces)
Steam	
1—Beat Saber	The goal is to make players almost dance while cutting the cubes flying toward them and avoiding obstacles using 2 sabers. Each cut is strongly supported by great sound and visual effects to emphasize the rhythm
2—Audioshield+C9	The players need to block the beats by punching the orbs flying toward them. Audioshield puts players at the point of impact for every hit in the music. Works with any song
3—Sprint Vector	Run, jump, climb, drift, and fly at extreme velocity as you race up to 8 players and battle obstacles in this frenetic VR adrenaline platformer
4—Hover Skate VR	Physics-based hoverboard simulator for VR with over 250 skateboard tricks to learn and master
5—VR Super Sports	8 sports games combined into one
6—Climbey	A VR-only climbing game! The goal is to climb to the finish as fast as you can, avoiding obstacles and trying not to fall along the way
7—Soundboxing	A VR music video kickboxing game where you create the beats! You can either punch to the music and compete to beat the best, or record your own challenge for others to beat—all set to your favorite music in a cool outdoor environment
8—BoxVR[b]	Rhythm-based boxing-inspired workouts. The only VR workout app with regularly updated workouts by professional fitness instructors
9—The Thrill of the Fight—VR Boxing	A room-scale VR boxing game. Face off in the virtual ring where you'll jab, dodge, and sweat your way to the top of the boxing world. Grab your gloves, step through the ropes, and become a champion
10—Racket: Nx	Racquetball meets breakout inside a giant pinball machine! Racket: Nx is a new kind of game that challenges the limits of what you can do with a racket and ball in VR
11—Racket Fury: Table Tennis VR[b]	Face 16 opponents in 4 distinctive cups and step into the most intense table tennis experience! Take on challenging rivals in the single-player campaign to climb the ladder of the Crown Galaxy! Get involved in heated exchanges with players from all over the world who, just like you, are hungry to win
12—Holopoint	An archery game. Fight your way through waves of responsive targets, samurai, and highly dangerous ninjas—all while drawing, nocking, and shooting your arrows as quickly as possible
13—Fruit Ninja VR	Step inside the Fruit Ninja universe and experience a slice of VR like never before
14—Knockout League—Arcade VR Boxing	A single-player arcade-style boxing game built from the ground up for VR. Dodge flaming uppercuts, block sweeping tentacle attacks, and pummel your opponents using 1:1 tracking of your head and hands
15—Eleven: Table Tennis	The ultimate table tennis simulator. Play opponents in Web-based multiplayer or practice against the advanced AI[c]. With physics designed to be as real as ever achieved in a Table Tennis simulator, you will forget you are in VR
16—To the top	Play opponents in Web-based multiplayer or practice against the advanced AI. With physics designed to be as real as ever achieved in a Table Tennis simulator, you will forget that you are in VR
17—Fancy Skiing VR	The game's structure is based on real skiing that players use the handles just as ski poles to get power and lean left or right to control the direction, which brings players a strong immersion
18—Sparc	A unique physical sport only possible in VR in which players compete in fast-paced, full-body VR gameplay and connect in a Web-based community
Viveport	
19—Arcade Saga	Allows you to take the role of a newly sentient being and battle for survival against AI Overlords determined to destroy you and your freethinking. Through 3 different futuristic sports—Fracture, Smash, and Bowshot—you'll harness your reactions and strategic thinking to shoot and dodge viruses, break firewall bricks, and spin data-balls past the Overlords and their minions. You will need to use your reflexes, wits, and changing strategies to survive
20—VR Cricket	A cricket game for VR with 2 bowling options, that is, fast and spin. Game also features the real pitch and 2× dimensions of the real pitch. Game features a stadium, sound effects, and animations

VR[a] exergame	Brief description of the VR exergame by marketplace (descriptions taken from or adopted from the marketplaces)
21—Knockout League[b]	A single-player arcade style boxing game built from the ground up for VR. Its intuitive gameplay involves moving and dodging with your body and 1:1 punch movement that lets you attack how you want without relying on buttons for main gameplay. Train up with various boxing drills to take on a crazy cast of characters with different fighting styles and gameplay to become the champion of the Knockout League
Oculus	
22—Eleven: Table Tennis VR	The ultimate table tennis simulator. Play opponents in Web-based multiplayer or practice against the advanced AI. With physics designed to be as real as ever achieved in a Table Tennis simulator, you will forget you are in VR
23—The Climb	Scale huge heights and feel the exhilaration of extreme free solo climbing. Explore and enjoy the view or compete for the fastest times on leaderboards with touch or gamepad controls
24—BoxVR[b]	See game 8
25—Knockout League[b]	See game 21
26—Fruit Ninja VR	Fruit slicing game in VR. Step inside the Fruit Ninja universe and experience a slice of VR like never before
27—Racket Fury: Table Tennis VR[b]	See game 11
28—Streetball VR	Arcade-style basketball when you step into the street one on one with the top finest in basketball action and play in different streets with beautiful backgrounds in VR
29—Echo Arena	Echo VR is your gateway to Echo Arena and Echo Combat. You play as a futuristic battle-ready robot armed with an array of weapons and abilities as your team of 3 competes against the opposition in a zero-gravity clash of robotic glory as you glide, boost, and punch your way to scoring goals in a breath-taking virtual arena

[a] VR: virtual reality.

[b] Where game repeats but in a different marketplace.

[c] AI: artificial intelligence.

Table 3. Main themes for both positive and negative virtual reality exergame reviews.

Themes	Description
Real-world feeling	Body movement akin to real-world sport or activities made games more appealing
Intuitive controls	Intuitive controls made games more playable and less frustrating, which was especially true for the less experienced players
Immersion	Immersion was important for sense of presence and overall enjoyment of the game
Music	Music greatly enhanced the overall experience of the games
Matching skill level	Gradual step-by-step acquisition of skill and skill-matched intensity levels highly impacted on playability
Social aspects	Sociable aspects of the games were favored, but sometimes multiplayer games left players feeling lonely or stuck
Exertion and matching real-world exercise levels	VR exergaming provides a high level of exertion, equivalent to real-world exercise
Appearance	The feeling of how you look when you play impacted on player's enjoyment of the game
Graphics[a]	Good graphics and display allowed players to play for longer periods of time
Bugs[a]	Bugs frustrated the players, but when they were fixed quickly, people were more likely to play on
Add-ons enhance gaming experience	Simple add-ons and tweaks had the potential to enhance the overall game experience
Motion sickness[a]	Motion sickness made it difficult for players to engage
Developers	Responsive and present developers gained more players and praise
Price	Price influenced the overall perception of the VR exergames

[a] Main themes of the negative reviews, although also present in positive reviews

Body Movement Akin to Real-World Sport or Activities Made Games More Appealing

People purchased and engaged with VR exergames for different reasons. Some people used VR exergames merely as entertainment, whereas others used them to build a skill, practice sport, and replace real-world sport or exercise. Sometimes VR exergaming did not match the feeling of performing sport or PA in the real world, and this had to do with controllers, mechanics (ie, the overall gameplay, sometimes interchangeably used as physics), physics (ie, simulation or game effects appearing to be realistic), or display quality and graphics of the game. These themes are interconnected. Reviewers often disagreed about the same game:

> *I'm a martial artist, I think this game is the closest to what we could get fighting in VR at the moment. If you would like a good work out, try playing ten rounds daily. It cannot replace actual boxing training but it will raise your confidence and fitness.*
>
> *With all the good reviews, this must just not work with my system. The hit detection was horrible, couldn't block a thing, and it was constantly sticking me on top of my opponent.*

Matching the physical skill levels meant that some exergames offered different modes of playing. Transferring the same physics to the VR exergames as those experienced in real-life sports or exercises was quite important for most players:

> *The most important aspect of table tennis is the physics and [name of the game] certainly ticks that box. It provides one of the best VR table tennis physics I have experienced to date. I have played over 240 hours of VR table tennis and play regularly in real life so I have a good idea of how the game of table tennis plays. [name of the game] translates so well that I can play all the shots that I play in real life with the correct amount of spin, power and accuracy. It is so accurate in fact that I can use it to improve my real game.*
>
> *It's designed for people who have never skated (it seems) whereby people that have will find it nonintuitive. Not sure if there's a middle ground but I hope you find it (or add in an option for "skater mode" ["skater mode" would be a difficulty and skill level for more experienced skaters]).*

Intuitive Controls Made Games More Playable and Less Frustrating, Which Was Especially True for the Less-Experienced Players

Intuitive controllers affected playability and overall satisfaction with the game:

> *Braking requires hitting both side buttons, which is weird and uncomfortable (I've always hated those side buttons). All the buttons are annoying to remember and I can't count the number of accidental screenshots I've taken.*
>
> *The controls are incredibly simple, too: throw punches by pushing your hands out in front of your while holding the controllers; block by raising your hands in front of your face. There's no need for button presses, as [name of game] is entirely based on movement.*

Immersion Was Important for the Sense of Presence and Overall Enjoyment of the Game

Intuitive controllers had an effect on immersion, but immersion also depended on other factors such as the mechanics of the game. The overall *flow* of the exergame was often related to immersion but was interrupted by poorly designed controllers or bugs:

> *Backstory is nonexistent. If you're going for this style of game at least have a setup. Does wonders for immersion and gameplay. Trust me... [the name of an example game] made the game 10× better for me.*
>
> *Especially when you complete a level and you're at the peak. I usually sit there for 10 minutes taking in the view. The world is full of life, you can see boats and planes off in the distance. I also use wrist weights while playing, it enhances the immersion and its exercise.*
>
> *The ball doesn't follow your hand and instead leaves several seconds after you throw. This and a few other issues related to lag. Unfortunately it kills the immersion completely.*

Immersion Was Important for Distraction From Thinking About Physical Activity

One of the key features of VR exergaming is immersion, and immersion was reported as helping the players perform PA:

> *It's a great workout and time flies while you tell to yourself: just one more song!*
>
> *I'm a nerd. I hate sports and physical activity with a passion[...] I played until my body just couldn't go on any more. As I'm writing this, I'm a sweaty mess and my whole body is hurting. The cross platform play is awesome, players are super nice to each other and the gameplay mechanics just work and the inputs are responsive [...] Now I need to go rest...*
>
> *I also use wrist weights while playing, it enhances the immersion and its exercise.*

Music Greatly Enhanced the Overall Experience of the Games

Reviewers explicitly mentioned music. Most reviews referred to music directly in terms of enhancing the experience of playing satisfaction, immersion, replayability, and presence:

> *Music is excellent for the game's style. I love this soundtrack.*
>
> *Not a big fan of the music so would appreciate some audio controls.*
>
> *Even though the songs are far from the kind of music that I listen to, I cannot help but walk around having the melodies in my head constantly. Some songs I play because I like the music, others I play because*

> *the level is just really fun to play. And some because both of those reasons.*

Sometimes music hampered the enjoyment of the game entirely:

> *[name of the game] is a music game where the gameplay doesn't match with the music. Not worth the price, especially now that [name of a competitor game] is out.*

Gradual Step-by-Step Acquisition of Skill and Skill-Matched Intensity Levels Highly Impacted on Playability

Some gamers spent long periods of time learning the mechanics of the game and commented on the technical support present in the game when they were stuck:

> *Great progression of levels keep the game interesting and require building new skills.*
>
> *Good workout but overall, it jumps to ridiculously hard after 5 levels. I have been working on it for the last 2 days and I can't get to round 10 because you get hit twice and it starts you over. The levels don't even seem much different from 5-9.*

Sociable Aspects of the Games Were Favored, But Sometimes Multiplayer Games Left Players Feeling Lonely or Stuck

Fun was also dependent on whether players could play with other people such as friends and family but when players had no one to play against, which defeated the gameplay:

> *You & friends can even just hangout and chat & play in the skate park and not do races at all. Love the freedom.*
>
> *This could be the best VR game so far. I love the game play. But the matchmaking wait is so bad - you have to wait and watch for 2.5 minutes between every game.*
>
> *I live in Tokyo...my son is working in New Jersey...we met up in [name of the game] for over 2 hours and had a blast. As if he was next to me...we could actually grab hold each other's hands to climb up the ledge.*

Virtual Reality Exergaming Provides a High Level of Exertion, Equivalent to Real-World Exercise

As already illustrated in the previous examples, the players commented on the level of exertion and workout quality. Additionally, some reviewers felt worried about their headsets getting sweaty and sought advice about maintenance and cleaning:

> *This game made me realise I am really out of shape.*
>
> *This game is an amazing workout. It's great for working the legs and your core (some arms too, but not nearly as intense). Right now I used 2 games [within the game] for exercise.*
>
> *I was sweating on my head mounted display and just had to quit. Not only that, but I was tripping over my wired headset. If you have a wireless headset, then this game is perfect. If you don't, you'll have to setup your wires or just be very cognizant of this problem.*

The Feeling of How You Look When You Play Impacted on Players' Enjoyment of the Game

Some players wanted to look as cool to someone else when playing the game as they do inside the game itself. This was especially true for those games that offered superhero qualities or exceptional skills:

> *You shoot down holograms and dodge the projectiles. It is one of the few games where you actually look cool in the real world while you play it as you contort your body to dodge and shoot.*
>
> *My wife thought I looked dumb shooting invisible arrows straight up at the ceiling while kneeling and swearing. I didn't. I felt like I looked awesome. You will be ducking, kneeling, spinning, nocking, shooting, nocking, shooting until you fall over.*

Good Graphics and Display Allowed Players to Play for Longer Periods of Time

Clear, solid graphics made players like the exergames more, but sometimes too much detail evoked discomfort:

> *Graphics and models are well designed. Good amount of detail, but not so much that things are ever a pixelated aliased mess at a distance.*
>
> *The graphics are terrible for a game this simple. Everything is shimmering and my eyes started to hurt soon after I started playing and made me uncomfortable.*

Bugs Frustrated the Players, But When They Were Fixed Quickly, People Were More Likely to Play On

Bug complaints were the most prevalent mechanics complaint across all reviews and for all games. VR technologies, hardware, and software issues were all mentioned. Some games were experienced as smooth, whereas others as unplayable:

> *Amount of bugs is recent versions is insane. Glitches, freezes control issues. Happen exclusively in this game. I wish there was a way to revert everything to the previous version.*
>
> *It's kind of frustrating to play through the song over and over again and not get a chance to really practice the hard part before you have to restart the song again.*
>
> *Within minutes of use, I ran into several bugs and had to restart the game several times before I could even try it. So my first 10-15 minutes are just troubleshooting.. fun..*
>
> *All the bugs are possible to work around and this kind of game lives on community engagement anyway.*
>
> *Annoying bugs. Love the games but I hate the "Your head is intersecting with an object. Please move it" screen. It just randomly pops out when I fall making it impossible to grab on to a grip to the one beneath*

> *it [...] also, It keeps making me jump out of the middle of nowhere. So annoying.*

Simple Add-Ons and Tweaks Had the Potential to Enhance the Overall Game Experience

Sometimes reviews on mechanics had to do with simple add-ons or tweaks that helped players track fitness-related behaviors and PA:

> *Hours logged do not necessarily reflect all time spent in the game.*
>
> *Need only to add more boxers and rings. Please add a calorie counter.*

Motion Sickness Made It Difficult for Players to Engage

Motion sickness was reported as problematic when playing certain VR exergames:

> *I am sometimes susceptible to VR sim sickness from artificial locomotion, depending on the game. [the name of the game] didn't give me any problems.*
>
> *I've got hundreds of hours in VR at this point, and no game has ever really made me sick the way [name of the game] has. I put on the comfort options in the game, but I still can't run more than one (maybe two) challenges before I have to leave the game.*

Responsive and Present Developers Gained More Players and Praise

Many reviews have tips and feedback for developers from praise to critique. Sometimes the developers were addressed directly and feedback was given to them about their response, behavior, or their presence Web-based, but sometimes the feedback was given to the developers more generally in terms of how they could improve the game mechanics or game controls:

> *Hats of[f] to the developers for the immersive experience and environments! Job well done! The touch controllers and style of locomotion are a great pairing.*
>
> *I have been unable to launch the game for a few months now. I am extremely disappointed with the services in effort to help my problem and think that the services to help problems such as mine need to be looked at further by the developers of the game.*

Price Influenced the Overall Perception of the Virtual Reality Exergames

Reviewers commended or complained about the price. Some reviews described a greater degree of tolerance for glitches and bugs if the game price was low. Sometimes a low price meant that reviewers gave it a more balanced rating. The opposite was also true: reviewers expected more when the price was high:

> *Not nearly enough content in the game to justify the current price.*
>
> *I'll probably buy it again after they work out some of the bugs, or if the price comes down significantly.*
>
> *For this small price, I give it a go!*

Discussion

Principal Findings

This study sought to learn from users' reviews of VR exergames from the 3 most popular marketplaces: Steam, Viveport, and Oculus. A thematic analysis revealed that VR can provide an immersive experience where the user feels fully present in the virtual environment and can provide a distraction and alternative to conventional PA while still providing exercise at an intensity equivalent to real-world PA. However, poor game design can frustrate players and lead to nonuse.

Players preferred games that are realistic (eg, closely simulated real-world sport); games that are intuitive (in terms of the body movement and physics of the game, game mechanics, and controls); games that provide gradual step-by-step increases in skill acquisition; games that either provide support (eg, tutorial) or social support; and games that provide access to accessories or player options. Concordant with a 2011 review of the literature on exergames, which emphasized how exergames can help physical education courses and improve a student's physical and social outcomes, our study also suggested exergames to be enjoyable technology that increase energy expenditure during play, motivate players to become more physically active, and promote social interaction [42]. We were not able to measure or discern a player's cognitive performance before and after playing VR exergames, but the aforementioned review also suggested exergaming (but not exclusively VR exergaming) to have the potential for positive cognitive effects such as problem solving, estimation, pattern recognition, memory, and improved academic performance [42]. Music was a popular choice to increase fun, immersion, and enjoyment of VR exergames. Players did not discuss weight loss but discussed fitness levels achieved as part of playing the VR exergames. Reviewers consistently reported that VR exergaming was providing a high level of exertion equivalent to real-world exercise and that the immersion and enjoyment was a welcome distraction from the exertion. Discussions centered on playability and enjoyment of the VR exergames and the experience of playing them.

A number of players did not feel that they were getting value for money. The overall experience with the VR exergame was tied to the price of the game. Some games were judged as worth the price, and some games were acceptable and more enjoyable because they were offered at a discounted rate. More was expected from games with higher price points.

There were 3 themes in negative reviews: the first was around bugs that rendered games frustrating and sometimes unplayable. The quality of graphics had a particularly strong impact on perceived enjoyment. Reviewers disliked when games had overly complex controls and display functions that evoke motion sickness, which is a potential side effect in VR displays [43,44].

These negative effects clash with VR's potential to improve mood, visual-spatial skills, coordination, motivation, and energy expenditure [34,42], and as such, careful consideration for the design features of the future exergames is paramount. What is particularly important is that the caloric expenditure, the end goal of the engagement with VR exergames, may be hampered by technological glitches and factors that can most likely only be modified by the designers of the games. Therefore, the first important factor is the design and consideration of the potential negative effects when designing a new VR exergame and taking user's feedback on board. In doing this, the designers are more likely to create a game that will flow and then also promote the state of *flow*, a state of absorption in the activity [45], which is a key factor in the attractiveness of exergames. Other important factors include the balance of skill and the perceived challenge, clear goals and rules, fair feedback or guidance, a sense of personal control over the activity or game, loss of self-doubt or feelings of self-concern (loss of consciousness), transformation of time (losing sense of time and feeling absorbed), and autotelic experience (an activity undertaken for its own sake) [46]. The balance between making the game appealing to a wide variety of people when taking into account individual differences (eg, age and disability status) may mean accepting that a great variety of choices are required, even within the same exergame to keep the game challenging, fun, and novel [46]. Recently, attempts have been made to maximize individualization using interactive deep reinforcement learning, which allows the game to adapt to user preferences [47].

Limitations and Future Research

This study has a sufficiently large sample size [48,49]. However, the user-submitted reviews may be biased because they could be from people who felt more strongly or passionately about the VR exergames or who had something very particular to say. There were no demographic data about reviewers or their experience of playing VR games, and as such, we cannot draw any conclusions about the characteristics, motivations, and overall VR experiences of reviewers.

Nor could we examine whether comments varied systematically by any individual differences between players. A study on affective and attentional states when running in a VR environment concluded that the effects of performing PA in a VR environment depend on individual difference factors (especially individual differences in presence and immersive tendencies) [50].

This study was an inductive thematic analysis with no a priori hypothesis or theory. It is difficult to draw conclusions with regard to whether these findings are transferable to all cultures or ages and if these findings were an optimal reflection of the VR exergaming community. However, this relatively novel approach to thematic analysis allowed us to capture the views of a large number of users, and it also complements traditional qualitative interviewing approaches.

Conclusions

This thematic analysis indicates that VR exergames have potential as a public health intervention and gives specific advice on what a game should do (including the importance of price and building a relationship between developers and the gaming community—issues often overlooked in a research or public health context). More research is necessary to determine which factors matter most across a wide range of exergame genres in the VR gaming space. However, we believe that users and developers will benefit from our findings. However, further exploration of factors influencing long term engagement with VR exergames is warranted.

Acknowledgments

The vEngage study [51] was funded by the Medical Research Council (grant number MR/R015430/1). We would like to thank our adolescent steering committee who have helped shape the original idea for the study.

Conflicts of Interest

This research is sponsored by the Medical Research Council industry partnership grant and leads to the development of a VR game licensed by Six to Start. There is no legal, financial, or commercial conflict with our industry partner company, Six to Start.

References

1. ukactive Research Institute. 2014. Turning the Tide of Inactivity URL: http://researchinstitute.ukactive.com/downloads/managed/Turning_the_tide_of_inactivity.pdf [accessed 2018-12-03] [WebCite Cache ID 76SGrNRZq]
2. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity--a systematic review of longitudinal studies. BMC Public Health 2013 Sep 8;13:813 [FREE Full text] [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]
3. Ekelund U, Ward HA, Norat T, Luan J, May AM, Weiderpass E, et al. Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European prospective investigation into cancer and nutrition study (EPIC). Am J Clin Nutr 2015 Mar;101(3):613-621 [FREE Full text] [doi: 10.3945/ajcn.114.100065] [Medline: 25733647]
4. Mueller N, Rojas-Rueda D, Cole-Hunter T, de Nazelle A, Dons E, Gerike R, et al. Health impact assessment of active transportation: a systematic review. Prev Med 2015 Jul;76:103-114. [doi: 10.1016/j.ypmed.2015.04.010] [Medline: 25900805]
5. Sallis JF, Floyd MF, Rodríguez DA, Saelens BE. Role of built environments in physical activity, obesity, and cardiovascular disease. Circulation 2012 Feb 7;125(5):729-737 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.110.969022] [Medline: 22311885]
6. Department of Health and Social Care. GOV.UK. 2011. UK Physical Activity Guidelines URL: https://www.gov.uk/government/publications/uk-physical-activity-guidelines [accessed 2018-12-07] [WebCite Cache ID 76SHn58vn]

7. World Health Organisation. Prevalence of Insufficient Physical Activity URL: https://www.who.int/gho/ncd/risk_factors/physical_activity_text/en/ [accessed 2018-12-07] [WebCite Cache ID 76SHufrAr]
8. Thomson K. British Nutrition Foundation. 2011. UK Physical Activity Guidelines URL: https://tinyurl.com/y6f3m5pd [accessed 2019-01-14] [WebCite Cache ID 76SHzY9bq]
9. Allender S, Foster C, Scarborough P, Rayner M. The burden of physical activity-related ill health in the UK. J Epidemiol Community Health 2007 Apr;61(4):344-348 [FREE Full text] [doi: 10.1136/jech.2006.050807] [Medline: 17372296]
10. NHS Digital. 2009. Health Survey for England - 2008: Physical Activity and Fitness URL: https://tinyurl.com/y3bgagp8 [accessed 2019-01-15] [WebCite Cache ID 76SI9jwOY]
11. Townsend P. Deprivation. J Soc Policy 2009 Jan 20;16(2):125-146 [FREE Full text] [doi: 10.1017/S0047279400020341]
12. Kamphuis CB, van Lenthe FJ, Giskes K, Huisman M, Brug J, Mackenbach JP. Socioeconomic status, environmental and individual factors, and sports participation. Med Sci Sports Exerc 2008 Jan;40(1):71-81. [doi: 10.1249/mss.0b013e318158e467] [Medline: 18182936]
13. US Department of Health and Human Services. 2017. Facts & Statistics: Physical Activity URL: https://www.hhs.gov/fitness/resource-center/facts-and-statistics/index.html [accessed 2019-02-25] [WebCite Cache ID 76SINMgyC]
14. Chau J, Chey T, Burks-Young S, Engelen L, Bauman A. Trends in prevalence of leisure time physical activity and inactivity: results from Australian national health surveys 1989 to 2011. Aust N Z J Public Health 2017 Dec;41(6):617-624. [doi: 10.1111/1753-6405.12699] [Medline: 28749561]
15. Anjana RM, Pradeepa R, Das AK, Deepa M, Bhansali A, Joshi SR, ICMR– INDIAB Collaborative Study Group. Physical activity and inactivity patterns in India - results from the ICMR-INDIAB study (phase-1) [ICMR-INDIAB-5]. Int J Behav Nutr Phys Act 2014 Feb 26;11(1):26 [FREE Full text] [doi: 10.1186/1479-5868-11-26] [Medline: 24571915]
16. National data. 2018. Simple Query URL: http://data.stats.gov.cn/easyquery.htm?cn=C01 [accessed 2019-02-25] [WebCite Cache ID 76SOCkXPc]
17. Ukie. 2018. UK Video Games Fact Sheet October 2018 URL: https://tinyurl.com/yygzmbab [accessed 2019-02-25] [WebCite Cache ID 76SOH1i64]
18. Renevey R. Ukie. 2014. IAB Games Research Shows Games Dominate URL: https://ukie.org.uk/news/2014/09/iab-games-research-shows-games-dominate [accessed 2019-02-26] [WebCite Cache ID 76TbXqwdq]
19. Goodman W, McFerran E, Purves R, Redpath I, Beeken RJ. The untapped potential of the gaming community: narrative review. JMIR Serious Games 2018 Sep 25;6(3):e10161 [FREE Full text] [doi: 10.2196/10161] [Medline: 30274962]
20. Limelight Networks. 2018. The State of Online Gaming 2018 URL: https://www.limelight.com/resources/white-paper/state-of-online-gaming-2018/ [accessed 2019-02-26] [WebCite Cache ID 76TcxcH4M]
21. Krause JM, Benavidez EA. Potential influences of exergaming on self-efficacy for physical activity and sport. J Phys Health Educ Recreat Dance 2014 Mar 27;85(4):15-20. [doi: 10.1080/07303084.2014.884428]
22. Nintendo. 2018. Top Selling Title Sales Units URL: https://www.nintendo.co.jp/ir/en/finance/software/3ds.html [accessed 2019-03-22] [WebCite Cache ID 76TeTDtAv]
23. Harris I. Pocket Gamer. 2018. Pokemon Go Captures 800 Million Downloads URL: https://www.pocketgamer.biz/news/68209/pokemon-go-captures-800-million-downloads/ [accessed 2019-01-22] [WebCite Cache ID 76Tg6R0wJ]
24. Gao Z, Chen S, Pasco D, Pope Z. A meta-analysis of active video games on health outcomes among children and adolescents. Obes Rev 2015 Sep;16(9):783-794. [doi: 10.1111/obr.12287] [Medline: 25943852]
25. Zeng N, Gao Z. Exergaming and obesity in youth: current perspectives. Int J Gen Med 2016;9:275-284 [FREE Full text] [doi: 10.2147/IJGM.S99025] [Medline: 27536158]
26. Sun H. Impact of exergames on physical activity and motivation in elementary school students: a follow-up study. J Sport Health Sci 2013 Sep;2(3):138-145. [doi: 10.1016/j.jshs.2013.02.003]
27. Lanningham-Foster L, Jensen TB, Foster RC, Redmond AB, Walker BA, Heinz D, et al. Energy expenditure of sedentary screen time compared with active screen time for children. Pediatrics 2006 Dec;118(6):e1831-e1835. [doi: 10.1542/peds.2006-1087] [Medline: 17142504]
28. Gao Y, Gerling KM, Mandryk RL, Stanley KG. Decreasing Sedentary Behaviours in Pre-Adolescents Using Casual Exergames at School. In: Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play. 2014 Presented at: CHI PLAY'14; October 19-21, 2014; Toronto, Ontario, Canada p. 97-106 URL: https://dl.acm.org/citation.cfm?id=2658693 [doi: 10.1145/2658537.2658693]
29. Staiano AE, Abraham AA, Calvert SL. Adolescent exergame play for weight loss and psychosocial improvement: a controlled physical activity intervention. Obesity (Silver Spring) 2013 Mar;21(3):598-601 [FREE Full text] [doi: 10.1002/oby.20282] [Medline: 23592669]
30. Staiano AE, Marker AM, Beyl RA, Hsia DS, Katzmarzyk PT, Newton RL. A randomized controlled trial of dance exergaming for exercise training in overweight and obese adolescent girls. Pediatr Obes 2017 Dec;12(2):120-128 [FREE Full text] [doi: 10.1111/ijpo.12117] [Medline: 26918815]
31. Huang HC, Wong MK, Lu J, Huang WF, Teng CI. Can using exergames improve physical fitness? A 12-week randomized controlled trial. Comput Hum Behav 2017 May;70:310-316. [doi: 10.1016/j.chb.2016.12.086]
32. Slater M, Wilbur S. A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. Presence-Teleop Virt 1997 Dec;6(6):603-616. [doi: 10.1162/pres.1997.6.6.603]

33. Witmer BG, Singer MJ. Measuring presence in virtual environments: a presence questionnaire. Presence-Teleop Virt 1998 Jun;7(3):225-240. [doi: 10.1162/105474698565686]
34. Plante TG, Aldridge A, Bogden R, Hanelin C. Might virtual reality promote the mood benefits of exercise? Comput Hum Behav 2003 Jul;19(4):495-509. [doi: 10.1016/S0747-5632(02)00074-2]
35. Liu S. Statista. 2018. Virtual Reality (VR) - Statistics & Facts URL: https://www.statista.com/topics/2532/virtual-reality-vr/ [accessed 2019-02-26] [WebCite Cache ID 76TjvDfKu]
36. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, de Bourdeaudhuij I. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. J Med Internet Res 2016 Apr 29;18(4):e94 [FREE Full text] [doi: 10.2196/jmir.4444] [Medline: 27129447]
37. Stawarz K, Cox AL, Blandford A. Don't Forget Your Pill!: Designing Effective Medication Reminder Apps That Support Users' Daily Routines. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI'14; April 26-May 1, 2014; Toronto, ON, Canada p. 2269-2278 URL: http://chi2014.acm.org/ [doi: 10.1145/2556288.2557079]
38. Edwards C. Bloomberg. 2013. Valve Lines Up Console Partners in Challenge to Microsoft, Sony URL: https://tinyurl.com/y4ne3ay2 [WebCite Cache ID 76TkJOGMA]
39. STEAM. 2018. Windows Mixed Reality for SteamVR URL: https://store.steampowered.com/app/719950/Windows_Mixed_Reality_for_SteamVR/ [accessed 2018-12-17] [WebCite Cache ID 76TkL86yo]
40. STEAM. 2018. Steam Support URL: https://help.steampowered.com/en/ [accessed 2019-02-26] [WebCite Cache ID 76Tl5DHxa]
41. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
42. Staiano AE, Calvert SL. Exergames for physical education courses: physical, social, and cognitive benefits. Child Dev Perspect 2011 Jun;5(2):93-98 [FREE Full text] [doi: 10.1111/j.1750-8606.2011.00162.x] [Medline: 22563349]
43. Gavgani AM, Walker FR, Hodgson DM, Nalivaiko E. A comparative study of cybersickness during exposure to virtual reality and 'classic' motion sickness: are they different? J Appl Physiol (1985) 2018 Oct 4:- (forthcoming). [doi: 10.1152/japplphysiol.00338.2018] [Medline: 30284516]
44. Munafo J, Diedrick M, Stoffregen TA. The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. Exp Brain Res 2017 Dec;235(3):889-901. [doi: 10.1007/s00221-016-4846-7] [Medline: 27915367]
45. Csikszentmihalyi M. Beyond Boredom and Anxiety: Experiencing Flow in Work and Play. San Francisco, California: Jossey-Bass Inc; 1975.
46. Sinclair J, Hingston P, Masek M. Considerations for the Design of Exergames. In: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia. 2007 Presented at: GRAPHITE'07; December 1-4, 2007; Perth, Australia p. 289-295. [doi: 10.1145/1321261.1321313]
47. Dobrovsky A, Borghoff UM, Hofmann M. Improving adaptive gameplay in serious games through interactive deep reinforcement learning. In: Klempous R, Nikodem J, Baranyi PZ, editors. Cognitive Infocommunications, Theory and Applications. New York City: Springer International Publishing; 2018:411-432.
48. Fugard AJ, Potts HW. Supporting thinking on sample sizes for thematic analyses: a quantitative tool. Int J Soc Res Methodol 2015 Feb 10;18(6):669-684. [doi: 10.1080/13645579.2015.1005453]
49. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
50. Neumann DL, Moffitt RL, Thomas PR, Loveday K, Watling DP, Lombard CL, et al. A systematic review of the application of interactive virtual reality to sport. Virtual Real 2017 Jul 19;22(3):183-198. [doi: 10.1007/s10055-017-0320-5]
51. vEngage Study. 2018. URL: https://vengagestudy.com [accessed 2019-07-15]

Abbreviations

PA: physical activity
VR: virtual reality

Edited by G Eysenbach; submitted 26.02.19; peer-reviewed by L Kelly, S Yang; comments to author 10.04.19; revised version received 17.04.19; accepted 18.04.19; published 16.09.19.

Please cite as:
Faric N, Potts HWW, Hon A, Smith L, Newby K, Steptoe A, Fisher A
What Players of Virtual Reality Exercise Games Want: Thematic Analysis of Web-Based Reviews
J Med Internet Res 2019;21(9):e13833
URL: https://www.jmir.org/2019/9/e13833/
doi:10.2196/13833
PMID:31538951

Original Paper

Health Care Professionals' Perspectives on the Secondary Use of Health Records to Improve Quality and Safety of Care in England: Qualitative Study

Ana Luísa Neves[1,2*], MSc, MD, PhD; Dilkushi Poovendran[1*], MSc, PhD; Lisa Freise[1], BSc, MSc; Saira Ghafur[3], MBChB, MRCP (Resp), MSc; Kelsey Flott[1], BA, MSc, PhD; Ara Darzi[1], PC, KBE, FRS, FMedSci, HonFREng; Erik K Mayer[1], MBBS, BSc, MRCS, PhD, FRCS (Urol)

[1]Patient Safety Translational Research Centre, Institute of Global Health Innovation, Imperial College London, London, United Kingdom

[2]Center for Health Technology and Services Research / Department of Community Medicine, Health Information and Decision, Faculty of Medicine, University of Porto, Porto, Portugal

[3]Centre for Health Policy, Institute of Global Health Innovation, Imperial College London, London, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Ana Luísa Neves, MSc, MD, PhD
Patient Safety Translational Research Centre
Institute of Global Health Innovation
Imperial College London
St Mary's Campus
Queen Elizabeth Queen Mother Wing
London, W2 1NY
United Kingdom
Phone: 44 (0)20 7589 5111
Fax: 44 (0)20 7589 5111
Email: ana.luisa.neves14@ic.ac.uk

Abstract

Background: Health care professionals (HCPs) are often patients' first point of contact in what concerns the communication of the purposes, benefits, and risks of sharing electronic health records (EHRs) for nondirect care purposes. Their engagement is fundamental to ensure patients' buy-in and a successful implementation of health care data sharing schemes. However, their views on this subject are seldom evaluated.

Objective: This study aimed to explore HCPs' perspectives on the secondary uses of health care data in England. Specifically, we aimed to assess their knowledge on its purposes and the main concerns about data sharing processes.

Methods: A total of 30 interviews were conducted between March 27, 2017, and April 7, 2017, using a Web-based interview platform and following a topic guide with open-ended questions. The participants represented a variety of geographic locations across England (London, West Midlands, East of England, North East England, and Yorkshire and the Humber), covering both primary and secondary care services. The transcripts were compiled verbatim and systematically reviewed by 2 independent reviewers using the framework analysis method to identify emerging themes.

Results: HCPs were knowledgeable about the possible secondary uses of data and highlighted its importance for patient profiling and tailored care, research, quality assurance, public health, and service delivery planning purposes. Main concerns toward data sharing included data accuracy, patients' willingness to share their records, challenges on obtaining free and informed consent, data security, lack of adequacy or understanding of current policies, and potential patient exposure and exploitation.

Conclusions: These results suggest a high level of HCPs' understanding about the purposes of data sharing for secondary purposes; however, some concerns still remain. A better understanding of HCPs' knowledge and concerns could inform national communication policies and improve tailoring to maximize efficiency and improve patients' buy-in.

(J Med Internet Res 2019;21(9):e14135) doi:10.2196/14135

KEYWORDS

electronic health records; information technology; health policy; safety culture

Introduction

Background

The use of electronic health records (EHRs) for secondary purposes, such as health research, public health surveillance, quality improvement, and safety monitoring, is essential for improving patient care [1]. In recent years, the exponential growth in EHR adoption, together with developments in health care informatics and data mining tools, has generated a range of health discoveries that ultimately improve the quality and safety of health care delivery [2-4]. Furthermore, health care data sharing for secondary uses, and particularly for research purposes, seems to be supported by the wider public [5-7]. At the same time, evidence suggests that harm due to inappropriate or insufficient use of available data is a real problem with potential serious consequences such as increased mortality and financial burden [8]. Therefore, while preserving patients' privacy, health care systems must find solutions to leverage health care data to deliver safer and better care.

Governmental initiatives worldwide have been advocating the use of EHR for secondary purposes. In the United States, the Health Information Technology for Economic and Clinical Health Act (2009) sets the adoption and meaningful use of EHR as a key national goal [9]. According to this Act, as a condition for clinicians and hospitals to receive incentive payments, they needed to *meaningfully use* certified EHRs. This refers to the use of EHR data to inform quality improvements, monitor safety, and drive efficiency. It can also refer to much broader societal health efforts such as reducing health disparities, engaging patients and families in their health, and improving care coordination and data security [10]. In Europe, the European Institute for Innovation through Health Data was launched with the aim to develop and promote best practices in the governance, quality, semantic interoperability, and uses of EHR data, including, importantly, its reuse for research purposes [11].

In the United Kingdom, the National Health Service (NHS) Care.data also aimed to securely link information from general practitioners' (GPs) records with data from secondary care, with the ultimate goal of providing a better understanding of patients' holistic needs [12]. However, the progress of this program stalled for a variety of reasons, including general concerns about data security, lack of information about benefits of data sharing, and complications around opt-out procedures, which eventually lead to the closure of the program in 2016 [13]. This experience showed that although governmental endorsement is key, it is not per se enough to ensure a successful implementation. Lessons learned include a greater awareness of the importance of adequately engaging with patients, the wider public, and, importantly, health care professionals (HCPs) in the debate about health care data sharing to learn their experiences and manage governmental programs accordingly [14].

Being in the frontline of health care delivery, HCP are often patients' first and closest point of contact in what concerns the communication of the purposes, benefits, and risks of sharing health care data for secondary uses. The provider-patient relationship in health care is a paramount example of a *trust relationship*; the nature of this trust-based relationship is the key reason why their support and engagement is fundamental to ensure patients' buy-in [15]. Provider-based communication with patients about the importance of using health care data for secondary purposes is critical: it enhances patients' awareness of the benefits of sharing their health information, increases patients' buy-in, and can ultimately have a positive impact on the availability of data for these uses. Altogether, these reasons justify the vital role of HCPs on the successful implementation of programs aiming to implement, or enhance, health care data sharing for secondary purposes. However, to date, most research in the United Kingdom focused on patients' perspectives [6,16], and the perceptions of HCPs on data sharing for these purposes were seldom evaluated.

Objectives

This study aimed to assess HCPs' perspectives on the secondary uses of health care data and, specifically, explore their knowledge on its purposes and the main concerns about data sharing processes.

Methods

Overview of the Methods Used

To meet the study's aims and objectives, a qualitative descriptive approach was adopted. In-depth interviews were chosen because of their ability to capture rich, descriptive data about individual perception, attitudes, and behaviors [17]. A multidisciplinary team including medical doctors (ALN, EKM, SG, and AD), a neuroscientist (DP), a health service researcher (KF), and a cognitive scientist (LF) with previous experience in qualitative research performed this study.

Recruitment

HCPs from a variety of roles (ie, doctor, nurse, and allied HCP) were invited to participate by email, using a combination of recruitment approaches, whereas primary and secondary care doctors were invited through UK's largest professional network of doctors (Doctors.net) membership, and HCPs from other roles were recruited from medeConnect Healthcare Insight's contact network database. Purposive sampling by health care role was used, and participants from mixed roles were included. Participants were excluded if they were previously involved in clinical research (ie, having treated or managed patients who had participated in clinical trial or being a principal or site investigator for clinical trials). Informed consent was obtained for each participant when replying to a screener survey.

Data Collection

A total of 30 Web-based in-depth interviews were conducted between March 27, 2017, and April 7, 2017. In-depth interviews allow the interviewer to explore and examine a given topic or experience in detail and are therefore an effective method for interpretative inquiry [18].

A Web-based bulletin system provided by medeConnect Healthcare Insight was used. A topic guide with open-ended questions was used to cover all the relevant topics during the online interviews, particularly their knowledge on the purposes of secondary uses of health care data, as well as their main concerns about data sharing processes (Multimedia Appendix 1). Questions were launched over the course of several days so that participants had time to reflect and answer each question.

As participants typed their own answers in the online bulletin system, transcription was automatic. The transcripts were compiled verbatim, following the required reliability and validity procedures for qualitative studies, and were not returned to participants for comments and/or corrections. The online interviews were hosted on secure servers belonging to the software providers for the interview software.

All interviews were conducted in English, and the interviewer had no established relationship with participants before study commencement. As online interviews were used, only the participants and the interviewer were present. Participants had a minimal knowledge of the characteristics of the interviewer and of the research team; thus, the potential for bias and assumptions was kept to the minimum. No repeat interviews were conducted.

Data Analysis

A total of 2 independent researchers systematically reviewed the transcripts, using the framework analysis method, which includes 5 main stages: familiarization, identifying a thematic framework, indexing, charting, and mapping and interpretation [19]. The charting stage is applied as a principle for developing the coding framework through a process of abstraction to ensure that coding elements that might have been missed with an a priori approach are adequately captured [19]. The defining feature of this method is the organization of qualitative data as a matrix output: rows (ie, cases and interviewees), columns (ie, themes), and cells of summarized data, thus providing a structure that systematically reduces qualitative data to analyze it by theme [20].

At every stage of the data analysis process, the coding framework was kept deductive and inductive, allowing the ongoing inclusion of emergent themes. Themes were supported by quotations derived from the interviews. Data saturation was reached after 22 interviews. As participants did not provide consent for further contact, it was not possible to ask them to provide feedback on the findings. The findings will be shared with patient partners (Research Partners Group), who will be involved in the codevelopment of a dissemination strategy and in summarizing the research findings into lay summaries and reports. The Consolidated Criteria for Reporting Qualitative studies were used to ensure the study meets the recommended standards of qualitative data reporting (Multimedia Appendix 2).

Results

Participants’ Characteristics

The 30 HCPs who completed the interview represented a variety of geographic locations across England (London, n=12; West Midlands, n=6; East of England, n=4; North East England, n=4; and Yorkshire and the Humber, n=2), covering both primary (n=11) and secondary care services (n=17). Of the 30 participants, 2 did not provide geographic location and care setting information. A full description of the participants by professional role is shown in Table 1.

Participants identified a wide range of individuals and organizations with whom they share EHR data (Table 2). Respondents acknowledged to share health care data with a variety of individuals and organizations for secondary purposes, including administrative and finance departments, occupational health and public health services (ie, Public Health England), health care trusts and commissioners, and teams involved in evaluation and research (Table 2).

Table 1. Description of participants by professional role.

Role	Value, n (%)
Primary care	
GP[a]	5 (17)
GP receptionist	3 (10)
Practice manager	3 (10)
Secondary care	
Allied health professional	9 (30)
Consultant	5 (17)
Pharmacist	1 (3)
Specialist nurse	2 (7)
Unknown	2 (7)

[a] GP: general practitioner.

Table 2. Health care professionals' perceptions on which individuals and organizations have access to health care data.

Parties with access to health care data	Examples
Health care professionals (clinical staff)	
Primary care	General practitioners, other local staff within primary care teams, and dentists
Secondary care	Hospital staff and community care
Tertiary care	Hospital staff and voluntary sector organizations
Secondary uses	
Occupational and public health	Occupational health departments; Public Health England
Trusts and commissioners	Trusts, Clinical Commissioning Groups, and National Health Service England
Evaluation and research	Audit teams and monitoring systems; research teams
Other	Judiciary system (courts and police) and schools
Supporting services	Biochemistry services, pathology services, suppliers and equipment companies, administrators, finance departments, and information technology

Understanding of the Purposes of Sharing Health Care Data for Secondary Uses

The level of content and details varied greatly between interviews; however, all participants were able to identify at least one secondary purpose for health care data sharing. Thematic analysis of the patients' narratives revealed 5 emerging themes, including (1) patient profiling, (2) research and evidence-based practice, (3) quality assurance, (4) public health purposes, and (5) health care delivery planning (Textbox 1).

Use of health care data to profile patients and improve tailored care appeared as a relevant purpose for several participants. HCPs highlight that analyzing health care data can provide useful insights into the context of preventive medicine, identify high-risk groups, and inform the design and implementation of tailored preventive measures (Textbox 1).

A few participants mentioned the importance of sharing health care data with researchers to generate evidence to support, guide, and improve health care delivery (Textbox 1). Participants also acknowledged quality assurance as a major purpose of sharing health care data for purposes beyond individual care. Pragmatic examples included the use of health care data to quantify and monitor various aspects of quality of care (ie, effectiveness and timeliness of service delivery) and to evaluate the compliance with previously set quality standards (Textbox 1).

Sharing health care data for public health purposes was also mentioned as a means to provide insights on disease surveillance and outbreaks. In this context, 2 HCPs recognized that health care data can be particularly useful to identify geographic trends and thus expose underserved or excluded areas or groups of individuals (Textbox 1).

Finally, HCPs emphasized that sharing EHR with trusts and commissioners can support planning and optimization of health care delivery (Textbox 1). Participants mention that these data can inform resource allocation and service provision and contribute to the improvement of current pathways for patient care and to the development of new ones.

Concerns Regarding Electronic Health Record Data Sharing Beyond Individual Care

Most participants declared feeling comfortable with data sharing policies and their implementation. One participant highlighted the influence of having personally experienced the benefits of data sharing and how this positively impacted his perceptions:

Textbox 1. Understanding of purposes of secondary uses of health care data and thematic analysis of the patients' narratives revealing 5 emerging themes.

Theme 1: patient profiling

- I would imagine [data sharing] is useful for creating key profiles for injury types, mechanisms, recovery and therefore contributing to strategies for prevention of such injuries to others. [ID 8]

Theme 2: research and evidence-based practice

- Being an allied health professional, many if not all our treatments are guided by research which is essential for effective management. For this reason it would be important to share patient data. [ID 18]
- Well planned effective research supports new methods of treatment. [ID 11]

Theme 3: quality assurance

- I think it is important to share information on service performance to ensure that services are functioning effectively in treatment delivery and identifying areas of good or bad treatment delivery. [ID 11]
- […] previously [collected] data contribute to seeing if quality standards were being met in GP setting, such as chronic disease monitoring. [ID 28]
- Sharing patient information beyond patient individual care for research or statistical purposes, [allows] to check if care services provided to patient are good and if patient receive them at the right time. [ID 16]
- […] a lot of other information is used for audit and service improvement it uses patient identifiable data but is always anonymised before going anywhere else. [ID 5]

Theme 4: Public health purposes

- I feel I have some understanding when patients' data is used for individual and non-individual care with other organisations such as area data when looking at epidemics. [ID 13]
- Data sharing can help to understand practice variations around the country or in different areas of the organization. [ID 9]
- […] you can also then use the data to compare how your area is doing compared to the national average. [ID 12]

Theme 5: Health care delivery planning

- [data sharing could] allow resource allocation and service provision from health care providers. [ID 28]
- I definitely think there is a place for patient data sharing when planning future services because it could improve services provided to patients. Additionally, it could allow more effective strategies to be put into place. [ID 18]
- Use of personal information allows for the amendment of and planning of new pathways for patient care. [ID 1]
- For example: Patient A might have followed a distinct pathway for a medical condition that was found to be too unwieldy and investigations might have taken place too late in the pathway. By sharing specific information about patients then the pathways can be reviewed and amended and improved for others in the future. [ID 1]

Nevertheless, some concerns were also mentioned. The themes that emerged formed 6 identifiable but sometimes interwoven concerns: (1) data inaccuracy, (2) patient unwillingness to share their data, (3) challenges to obtain free and informed consent, (4) data security, (5) lack of adequacy/need for clear policies, and (6) potential patient exposure to distress/exploitation (Textbox 2). Regarding data inaccuracy, a few participants mentioned being concerned about its implications on both diagnostic characterization and evaluation of Quality and Outcomes Framework targets (Textbox 2). Some participants also highlighted being concerned that patients might not be able to understand current communication policies around data sharing for secondary purposes and therefore perceive it as a negative initiative and thus decide to opt-out (Textbox 2). Respondents indicated that obtaining patients' free and informed consent to share their data may pose particular challenges, as patients may feel external pressure to share their information and therefore not take this decision as freely as would be desirable. One participant also mentioned being concerned about the adequacy and completeness of the information provided to patients when asking for their consent to share data, particularly when consent is obtained verbally (Textbox 2). Several participants reported having concerns on the security of data transmission and, importantly, about the possibility of data being inappropriately abusively assessed by third parties during this process (Textbox 2).

Participants reported concerns not only about the robustness of regulations and policies but also about the ways these are communicated to the public, which may create a negative impact on patients' buy-in (Textbox 2).

Altogether, the themes mentioned have the potential to result in deleterious consequences for the patient and thus to contribute to patient's exposure to distress/possible exploitation (Textbox 2).

Textbox 2. Concerns regarding health care data sharing beyond individual care and thematic analysis of the patients' narratives revealing 6 emerging themes.

Theme 1: data inaccuracy

- I […] have concerns about data accuracy and implications of diagnostic categorization, and particular concerns regarding data recording and its implications on QOF DATA and targets. [ID 22]

Theme 2: patients' unwillingness to share data

- I am concerned patients […] [can] feel that sharing data is a negative thing and a breach of confidentiality.

Theme 3: quality assurance

- I am happy to share where I have consent, though I am not entirely convinced that the patient understands the different scenarios. [ID 1]
- I have concerns that there is pressure for me to ask patients and their carers to agree for their information to be shared for research purposes. I believe that research […] is essential to try to find curative treatments however I believe that the patient has the right to say no. [ID 3]
- Concerns about formally getting consent from the patients especially regarding the equipment company, it is always verbal so we have no evidence of what was said and what the patient said. Is it truly informed, have we given the correct information? [ID 12]

Theme 4: data security

- The only concern I have is regarding the security of the information that is shared electronically. Are IT processes to safeguard data robust, considering occasional reports in local media of leaks of personal data from private companies? [ID 6]
- I always remain concerned about how safe this data remains during transmission especially. [ID 9]
- I am worried hackers and people not delivering care to an individual, should not be able to access this sensitive information about someone's health. [ID 28]

Theme 5: lack of adequacy/understanding of current policies

- I would hope that the sharing of data for secondary use is strictly regulated, confidential and, where necessary, anonymous. [ID 14]
- I am afraid that policies are inadequate regarding data security. [ID 22]
- I am cynical about Government's ability, and track record, on protecting data. [ID 27]
- I am concerned that the policies can sometimes be too detailed for staff/patients to fully understand. [ID 10]

Theme 6: patient exposure to distress/possible exploitation

- Not specifically, but one is always mindful of the possibility of abuse of access. [ID 20]
- The concerns are often about the patient being vulnerable and giving out information that could lead to them becoming distressed/possible exploitation. [ID 29]

Discussion

Principal Findings

In this study, we explored HCPs' knowledge of secondary uses of health care data and their major concerns about data sharing processes. Participants showed a comprehensive knowledge of the purposes of sharing health care data for purposes beyond individual care, including for uses such as patient profiling, research, quality assurance activities, and public health purposes and to support health care delivery and planning.

Although the majority of the participants mentioned feeling comfortable with data sharing policies and their implementation, they also acknowledged concerns about data sharing, including data accuracy, patients' willingness to share their records, challenges on obtaining free and informed consent, data security, and potential patient exposure and exploitation.

Comparison With Previous Work

Understanding of Purposes of Sharing Health Care Data for Secondary Uses

The use of patient-level data for population segmentation, tailored care, and understanding patterns in patient needs has been recognized in the literature as one of the biggest promises of big data in health care [21]. Computing and analytics frameworks allow aggregation and integration of big data and the identification of more accurate, stratified disease risk profiles, which can then be targeted with care models and intervention programs tailored to specific population segments' needs [3].

Similarly, several studies have acknowledged the importance of sharing EHR data for research purposes. In the last decade, the United Kingdom has witnessed a surge in the secondary use of health care data to derive population-based insights, namely, concerning mental health status [22] and prescription patterns [23]. In the United States, pilot studies have already started predicting readmissions and estimating the risk of complications in newborns [21]. Once impossible because of the effort required to collect data, these studies have now become possible by reusing data originally collected for the purposes of individual care, thus maximizing the potential for the use of EHR data.

Previous literature also highlighted the importance of sharing health care data in the context of quality assurance activities [24], which was similarly acknowledged by the participants of this study. It is important to note, however, that these approaches have potential limitations, as EHR-derived measures can undercount practice performance [25].

The use of health care data was also previously suggested to be a powerful tool to support public health activities and facilitate timely and efficient surveillance of both noncommunicable and communicable diseases [4,26].

As acknowledged by our participants, sharing EHR data can also inform strategic planning and service delivery. As previously discussed, strategic planning can be informed by stratification by patients' risk and also by specific geographic needs and discrepancies. In line with our findings, Rumsfeld et al also suggest that geocoded health care data as a source for big data analytics might improve targeting of community and health resources for patients [2].

Challenges of Health Care Data Sharing

Concerns about data accuracy and data quality have been previously addressed in the literature and are critical when extracting real data from the massive, variable, and complex health care datasets, as suboptimal data quality can lead to low utilization efficiency and, importantly, poorly informed decision making [27].

Interestingly, although the participants reported being concerned about patients' unwillingness to share data, studies assessing patients' perspectives do not seem to validate this perception. Public members express a widespread willingness to share health care data for secondary uses and, particularly, for research purposes, both in the United Kingdom (68.7%-91.4%) [5,6] and the United States (92%-100%) [7]. Nevertheless, patients' confidence about sharing data varies greatly depending on ethnicity, social class, and working status, with some remarkable negative perceptions among lower socioeconomic groups and ethnic minorities [28]. Furthermore, although members of the public are generally supportive of data to be shared and used for the purpose of helping others or improving health care, they tend to be less supportive for data to be used by commercial companies [28].

The main concern raised in this study was toward data security and, particularly, about the risk of data breaches, which can lead to patient exposure and possible exploitation. GPs and social care professionals want reassurance that partner organizations with whom they share data are protecting people's confidential data. Similar concerns were also raised by patients in other studies, including data leakage, loss or reidentification, unauthorized access, and sharing with third parties [29]. However, despite significant recent events (such as the accidental disclosure of a clinic list of email addresses at an HIV clinic [30], the WannaCry malware attack [31], and data sharing without consent in the Pharmacy 2U incident [32]) patients' confidence levels remain high. Overall, 77% of the public reported trusting the ability of the NHS to protect health care data. The health sector remained the most trusted by the public, ahead of sectors such as retail, banking, and other public services [33].

Strengths, Limitations, and Future Work

Our purposive sampling strategy ensured that HCPs from a variety of professional roles, health care settings, experience with data sharing, and geographic location were represented. Although we achieved sample diversity, we acknowledge that future surveys using randomized, more powerful samples could explore the determinants of HCPs' concerns toward patients' data sharing and assess the external validity of our key findings.

It must also be noted that the understanding and experiences of HCPs toward data sharing for nonindividual care were retrospectively assessed and that several relevant initiatives took place recently. This research temporally overlapped with a new campaign launched by the NHS in England, in partnership with the Information Commissioner's Office ("Your data matters–building trust and confidence"), aiming to increase the public's trust and confidence in how their data are both used and made available [34]. Importantly, this initiative provided resources to support health and care staff, including tailored materials for GP practices, nursing, midwifery, and care staff, and opened a register for the ones wanting to publicly pledge their support for their service users' data rights [34]. In addition, in Europe, the European Union General Data Protection Regulation, in May 2018, has also recently raised awareness, in a helpful way, in professionals and the public on data privacy [35]. We anticipate that longitudinal prospective studies could provide further light on temporal trends on this subject and, importantly, to explore the impact of these recent policies and initiatives on HCP perspective toward EHR data sharing.

Conclusions and Implications for Policy Makers and Public Communication

Although our results suggest a good understanding of the purposes of data sharing for secondary uses, some concerns still remain. HCPs seem particularly concerned with data accuracy and consent procedures; to mitigate these concerns, they advocate clear policies to appraise and monitor data quality as well as clear procedures to obtain consent. To address concerns related to data security and potential exposure and exploitation of patients' data, cybersecurity emerges as an important part of the health care culture. To that end, both infrastructure investment and culture change are crucial to minimize accidental or malicious data breaches [31] that can harm individuals and organizations and, importantly, erode both patients' and HCP's trust.

HCPs were also concerned about patients' unwillingness to share data, a perception that does not seem to be corroborated by studies assessing patients' perspectives [5-7] and that might negatively impact how they communicate the importance of data sharing to patients. Communication policies must provide HCPs with this evidence to reinforce their confidence in patients' overall receptivity and highlight the importance of targeted communication strategies to improve negative perceptions among lower socioeconomic groups and ethnic minorities.

In the broader context of health care data use for secondary purposes, HCPs are a key partner in creating a patient-centric system across the NHS [34]. A better understanding of their knowledge and concerns can inform national and international communication policies and, importantly, engage them as active contributors to a nuanced decision-making process with and for patients, thus supporting a more widespread learning culture to fully embrace the potential of health care data use.

Acknowledgments

This work is supported by the National Institute for Health Research (NIHR) Imperial Patient Safety Translation Research Centre. Infrastructure support was provided by the NIHR Imperial Biomedical Research Centre.

Authors' Contributions

The manuscript was written by ALN with contributions from all authors. ALN, DP, and EKM conceptualized this research and designed the protocol. DP designed the online interviews. ALN and LF performed the qualitative analysis. SG, KF, AD, and EKM contributed to the conceptualization and commented on the multiple versions of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topic guide.
[PDF File (Adobe PDF File)55 KB - jmir_v21i9e14135_app1.pdf]

Multimedia Appendix 2
Consolidated Criteria for Reporting Qualitative studies (COREQ) checklist.
[PDF File (Adobe PDF File)116 KB - jmir_v21i9e14135_app2.pdf]

References

1. Friedman DJ, Parrish RG, Ross DA. Electronic health records and US public health: current realities and future promise. Am J Public Health 2013 Sep;103(9):1560-1567. [doi: 10.2105/AJPH.2013.301220] [Medline: 23865646]
2. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. Nat Rev Cardiol 2016 Jun;13(6):350-359. [doi: 10.1038/nrcardio.2016.42] [Medline: 27009423]
3. Vuik SI, Mayer E, Darzi A. Enhancing risk stratification for use in integrated care: a cluster analysis of high-risk patients in a retrospective cohort study. BMJ Open 2016 Dec 19;6(12):e012903 [FREE Full text] [doi: 10.1136/bmjopen-2016-012903] [Medline: 27993905]
4. Tomasallo CD, Hanrahan LP, Tandias A, Chang TS, Cowan KJ, Guilbert TW. Estimating Wisconsin asthma prevalence using clinical electronic health records and public health data. Am J Public Health 2014 Jan;104(1):e65-e73 [FREE Full text] [doi: 10.2105/AJPH.2013.301396] [Medline: 24228643]
5. Buckley BS, Murphy AW, MacFarlane AE. Public attitudes to the use in research of personal health information from general practitioners' records: a survey of the Irish general public. J Med Ethics 2011 Jan;37(1):50-55. [doi: 10.1136/jme.2010.037903] [Medline: 21071570]
6. Papoutsi C, Reed JE, Marston C, Lewis R, Majeed A, Bell D. Patient and public views about the security and privacy of electronic health records (EHRs) in the UK: results from a mixed methods study. BMC Med Inform Decis Mak 2015 Oct 14;15:86 [FREE Full text] [doi: 10.1186/s12911-015-0202-2] [Medline: 26466787]
7. Bell EA, Ohno-Machado L, Grando MA. Sharing my health data: a survey of data sharing preferences of healthy individuals. AMIA Annu Symp Proc 2014;2014:1699-1708 [FREE Full text] [Medline: 25954442]
8. Jones KH, Laurie G, Stevens L, Dobbs C, Ford DV, Lea N. The other side of the coin: harm due to the non-use of health-related data. Int J Med Inform 2017 Jan;97:43-51 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.09.010] [Medline: 27919394]
9. Blumenthal D. Launching HITECH. N Engl J Med 2010 Feb 4;362(5):382-385. [doi: 10.1056/NEJMp0912825] [Medline: 20042745]
10. Blumenthal D, Tavenner M. The 'meaningful use' regulation for electronic health records. N Engl J Med 2010 Aug 5;363(6):501-504. [doi: 10.1056/NEJMp1006114] [Medline: 20647183]
11. The European Institute for Innovation through Health Data (i~HD). 2018. About i~HD: Enriching Knowledge and Enhancing Care Through Health Data URL: https://www.i-hd.eu/index.cfm/about/about-i-hd/ [accessed 2019-01-21]
12. NHS England. 2013. Nhs England Sets Out the Next Steps of Public Awareness About care.data URL: https://www.england.nhs.uk/2013/10/care-data/ [accessed 2018-07-11]

13. Appleby J. BMJ Blogs. 2014. Care.data—your Bits in Their Hands URL: https://blogs.bmj.com/bmj/2014/02/13/john-appleby-care-data-your-bits-in-their-hands/ [accessed 2019-02-07]
14. Safran C, Bloomrosen M, Hammond WE, Labkoff S, Markel-Fox S, Tang PC, Expert Panel. Toward a national framework for the secondary use of health data: an American Medical Informatics Association white paper. J Am Med Inform Assoc 2007;14(1):1-9 [FREE Full text] [doi: 10.1197/jamia.M2273] [Medline: 17077452]
15. Haluza D, Jungwirth D. ICT and the future of health care: aspects of doctor-patient communication. Int J Technol Assess Health Care 2014 Jul;30(3):298-305. [doi: 10.1017/S0266462314000294] [Medline: 25308693]
16. Riordan F, Papoutsi C, Reed JE, Marston C, Bell D, Majeed A. Patient and public attitudes towards informed consent models and levels of awareness of electronic health records in the UK. Int J Med Inform 2015 Apr;84(4):237-247 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.01.008] [Medline: 25649841]
17. Peters K, Halcomb E. Interviews in qualitative research. Nurse Res 2015 Mar;22(4):6-7. [doi: 10.7748/nr.22.4.6.s2] [Medline: 25783145]
18. Fontana A, Frey J. Handbook of Qualitative Research. Thousand Oaks, CA: Sage Publications; 1994.
19. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Huberman AM, Miles MB, editors. The Qualitative Researcher's Companion. London: Routledge; 1994.
20. Richie J, Lewis J. Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage Publications; 2003.
21. Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. Health Aff (Millwood) 2014 Jul;33(7):1123-1131. [doi: 10.1377/hlthaff.2014.0041] [Medline: 25006137]
22. Bergman BP, Mackay DF, Smith DJ, Pell JP. Long-term mental health outcomes of military service: national linkage study of 57,000 veterans and 173,000 matched nonveterans. J Clin Psychiatry 2016 Jun;77(6):793-798. [doi: 10.4088/JCP.15m09837] [Medline: 27135139]
23. John A, Marchant AL, McGregor JI, Tan JO, Hutchings HA, Kovess V, et al. Recent trends in the incidence of anxiety and prescription of anxiolytics and hypnotics in children and young people: an e-cohort study. J Affect Disord 2015 Sep 1;183:134-141 [FREE Full text] [doi: 10.1016/j.jad.2015.05.002] [Medline: 26005207]
24. Murdoch TB, Detsky AS. The inevitable application of big data to health care. J Am Med Assoc 2013 Apr 3;309(13):1351-1352. [doi: 10.1001/jama.2013.393] [Medline: 23549579]
25. Parsons A, McCullough C, Wang J, Shih S. Validity of electronic health record-derived quality measurement for performance monitoring. J Am Med Inform Assoc 2012;19(4):604-609 [FREE Full text] [doi: 10.1136/amiajnl-2011-000557] [Medline: 22249967]
26. Birkhead GS, Klompas M, Shah NR. Uses of electronic health records for public health surveillance to advance public health. Annu Rev Public Health 2015 Mar 18;36:345-359. [doi: 10.1146/annurev-publhealth-031914-122747] [Medline: 25581157]
27. Cai L, Zhu Y. The challenges of data quality and data quality assessment in the big data era. Data Sci J 2015 May 22;14:2 [FREE Full text] [doi: 10.5334/dsj-2015-002]
28. Hunn A. Health Research Authority. 2017. Survey of the General Public: Attitudes Towards Health Research URL: https://www.hra.nhs.uk/documents/872/survey-general-public-attitudes-towards-health-research.pdf [accessed 2019-02-07]
29. Stockdale J, Cassell J, Ford E. 'Giving something back': a systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. Wellcome Open Res 2018;3:6 [FREE Full text] [doi: 10.12688/wellcomeopenres.13531.2] [Medline: 30854470]
30. BBC Online. 2015. London Clinic Leaks HIV Status of Patients URL: https://www.bbc.co.uk/news/uk-england-london-34127740 [accessed 2019-02-07]
31. Martin G, Ghafur S, Kinross J, Hankin C, Darzi A. WannaCry-a year on. Br Med J 2018 Jun 4;361:k2381. [doi: 10.1136/bmj.k2381] [Medline: 29866711]
32. Andalo D. The Pharmaceutical Journal. 2015. Pharmacy2u Fined £130,000 for Selling Patient Data URL: https://www.pharmaceutical-journal.com/news-and-analysis/pharmacy2u-fined-130000-for-selling-patient-data/20069579.article?firstPass=false [accessed 2019-02-07]
33. Healthwatch England. 2018. How Do People Feel About Their Data Being Shared by the NHS? URL: https://www.healthwatch.co.uk/news/2018-05-17/how-do-people-feel-about-their-data-being-shared-nhs [accessed 2019-02-07]
34. Information Commissioner's Office. 2018. Your Data Matters – Building Trust and Confidence URL: https://ico.org.uk/for-organisations/resources-and-support/your-data-matters-campaign/ [accessed 2019-02-07]
35. European Commission. 2018. EU Data Protection Rules URL: https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en [accessed 2018-10-22]

Abbreviations

EHR: electronic health record
GP: general practitioner

HCP: health care professional
NHS: National Health Service
NIHR: National Institute for Health Research

Edited by G Eysenbach; submitted 20.04.19; peer-reviewed by L Laranjo, X Wu; comments to author 28.06.19; revised version received 08.07.19; accepted 28.07.19; published 26.09.19.

Please cite as:
Neves AL, Poovendran D, Freise L, Ghafur S, Flott K, Darzi A, Mayer EK
Health Care Professionals' Perspectives on the Secondary Use of Health Records to Improve Quality and Safety of Care in England: Qualitative Study
J Med Internet Res 2019;21(9):e14135
URL: https://www.jmir.org/2019/9/e14135
doi:10.2196/14135
PMID:31573898

Original Paper

The Content and Nature of Narrative Comments on Swiss Physician Rating Websites: Analysis of 849 Comments

Stuart McLennan[1,2], MBHL, PhD

[1]Institute of History and Ethics in Medicine, Technical University of Munich, Munich, Germany
[2]Institute for Biomedical Ethics, University of Basel, Basel, Switzerland

Corresponding Author:
Stuart McLennan, MBHL, PhD
Institute of History and Ethics in Medicine
Technical University of Munich
Ismaninger Straße 22
Munich, 81675
Germany
Phone: 49 089 4140 4041
Email: stuart.mclennan@tum.de

Abstract

Background: The majority of physician rating websites (PRWs) provide users the option to leave narrative comments about their physicians. Narrative comments potentially provide richer insights into patients' experiences and feelings that cannot be fully captured in predefined quantitative rating scales and are increasingly being examined. However, the content and nature of narrative comments on Swiss PRWs has not been examined to date.

Objective: This study aimed to examine (1) the types of issues raised in narrative comments on Swiss PRWs and (2) the evaluation tendencies of the narrative comments.

Methods: A random stratified sample of 966 physicians was generated from the regions of Zürich and Geneva. Every selected physician was searched for on 3 PRWs (OkDoc, DocApp, and Medicosearch) and Google, and narrative comments were collected. Narrative comments were analyzed and classified according to a theoretical categorization framework of physician-, staff-, and practice-related issues.

Results: The selected physicians had a total of 849 comments. In total, 43 subcategories addressing the physician (n=21), staff (n=8), and practice (n=14) were identified. None of the PRWs' comments covered all 43 subcategories of the categorization framework; comments on Google covered 86% (37/43) of the subcategories, Medicosearch covered 72% (31/43), DocApp covered 60% (26/43), and OkDoc covered 56% (24/43). In total, 2441 distinct issues were identified within the 43 subcategories of the categorization framework; 83.65% (2042/2441) of the issues related to the physician, 6.63% (162/2441) related to the staff, and 9.70% (237/2441) related to the practice. Overall, 95% (41/43) of the subcategories of the categorization framework and 81.60% (1992/2441) of the distinct issues identified were concerning aspects of performance (interpersonal skills of the physician and staff, infrastructure, and organization and management of the practice) that are considered assessable by patients. Overall, 83.0% (705/849) of comments were classified as positive, 2.5% (21/849) as neutral, and 14.5% (123/849) as negative. However, there were significant differences between PRWs, regions, and specialty regarding negative comments: 90.2% (111/123) of negative comments were on Google, 74.7% (92/123) were regarding physicians in Zurich, and 73.2% (90/123) were from specialists.

Conclusions: From the narrative comments analyzed, it can be reported that interpersonal issues make up nearly half of all negative issues identified, and it is recommended that physicians should focus on improving these issues. The current suppression of negative comments by Swiss PRWs is concerning, and there is a need for a consensus-based criterion to be developed to determine which comments should be published publicly. Finally, it would be helpful if Swiss patients are made aware of the current large differences between Swiss PRWs regarding the frequency and nature of ratings to help them determine which PRW will provide them with the most useful information.

(J Med Internet Res 2019;21(9):e14336) doi:10.2196/14336

KEYWORDS

physician rating websites; patient satisfaction

Introduction

Physician rating websites (PRWs) are a sign of the growing digitalization of the patient-health professional relationship, allowing patients to anonymously rate their physicians on the Web as a source of information for others [1-8]. Typically grounded in the assumptions of a theoretical consumer choice model [9], PRWs aim to improve patient welfare through (1) influencing patient decision making by increasing the chance that those patients who obtain information from PRWs will choose better quality physicians and benefit from this [10] and (2) driving quality improvement by identifying aspects of care needing improvement so that changes can be made in the practice [10].

In a recent systematic search of PRWs internationally, 143 different websites from 12 countries were analyzed [11]. The majority of websites were commercially operated by for-profit companies and were registered in the United States and Germany. It was found that 15.3% of these websites only provided the option to give feedback on a predefined quantitative rating scale and 4.2% of websites only provided the option for narrative comments, whereas 76.9% of websites provided the option to give both types of feedback [11].

Narrative comments potentially provide richer insights into patients' experiences and feelings that cannot be fully captured in predefined quantitative rating scales and are increasingly being examined with content analysis [4,8,12-15], and more recently with machine learning [16-17]. Narrative comments, however, have proved contentious because of concerns that they will be used for *doctorbashing* or defamation [4,18-20]. Although previous research suggests that this concern is largely unfounded [4,8,12-15], physicians' fear of receiving negative comments on PRWs can have a direct impact on the patient-health professional relationship. For instance, physicians may try to prevent patients from posting negative reviews on PRWs (eg, by asking patients to sign contracts stating they will not make negative comments) and legally challenge negative comments that are posted [3]. Due to the lack of expert knowledge of most patients regarding medicine, there are also concerns about the relevance and significance of their evaluation of physicians' performance [21]. Although recent research suggests that patients acknowledge their inability to rate physicians' technical competency [21], an analysis of 3000 narrative comments from the German PRW, jameda, from 2012 found that physicians' competence was the most frequent issue that patients mentioned in their comments (1874/3000, 62.46%) [4]. It is unclear whether this apparent contradiction, between patients' agreement about their inability to evaluate physicians' technical skills and their actual ratings, exists on other PRWs and countries, but the need for more research on this issue has been highlighted [21].

Although the first PRWs in Switzerland, OkDoc and Medicosearch, were established in 2008 at the same time as many international PRWs, there has been a lack of research conducted on PRWs in Switzerland to date. However, a study recently examined, for the first time, the frequency of quantitative ratings and narrative comments on Swiss PRWs [22]. It found that many of the selected physicians could not be identified (42.4% to 87.3%), and few of the identifiable physicians had been rated quantitatively (4.5% to 49.8%) or received narrative comments (4.5% to 31.2%) at least once. Rated physicians also had on average a low number of quantitative ratings (1.47 to 3.74 rating) and narrative comments (1.23 to 3.03 comments) [22]. However, there were significant differences between PRWs, with Google having substantially more quantitative ratings and narrative comments than the 3 dedicated Swiss PRWs in the past two and a half years [22].

Although the content and nature of narrative comments on Swiss PRWs has not been examined to date, the controversial nature of negative comments on Swiss PRWs has received media attention in Switzerland [23-26]. Furthermore, in response to a decision of a federal data commissioner that certain negative comments had to be removed on the PRW OkDoc in a particular case, OkDoc decided to only allow recommendations. OkDoc now explicitly states on its website that any negative comments will be deleted ("Only positive comments recommending your doctor will be accepted. Any negative post will be deleted. Thank you for respecting okdoc's principles!" [author translation]). Although the PRW Medicosearch allows negative comments, it informs the concerned physician before publishing it online so the physician can decide if the negative feedback is activated. However, if the physician refuses, the feedback function is deactivated, also removing the positive comments [23]. This situation potentially raises important implications not only for the frequency of ratings on Swiss PRWs but also for the types of comments that may be available for PRW users. Therefore, this study aimed to examine (1) the types of issues raised in narrative comments published on Swiss PRWs and (2) the evaluation tendencies of the narrative comments. Gaining better understanding regarding this issue may help identify issues that Swiss physicians should focus on to improve patient satisfaction and will also help inform future research and health policy in Switzerland in relation to PRWs.

Methods

Sample

A random stratified sample of 966 physicians was generated from the regions of Zürich and Geneva. Zürich is the largest city in Switzerland and is located in north-central Switzerland. Zürich has a total population of 402,762 (12/2016). Geneva is the second largest city in Switzerland and is located in south-western Switzerland. Geneva has a total population of 198,979 (12/2016). The regions of Zürich and Geneva were chosen because of language (German vs French) and a comparable number of total physician (Zürich 3254 physicians and Geneva 2780 physicians) considerations.

In November 2017, all physicians in these regions working in general practice, obstetrics and gynecology, pediatrics, and dermatology and venereology were searched for on the Swiss Medical Association's medical registry (Ärzteverzeichnis). From each region, a random sample was generated for each specialty based on a 95% confidence level and 5% confidence interval. From Zürich, the random sample consisted of 254 of 747 general practice physicians, 85 of 109 obstetrics and gynecology physicians, 74 of 92 pediatric physicians, and 53 of 61 dermatology and venereology physicians. Therefore, the Zürich sample of 466 physicians represents 46.18% of a total of 1009 physicians. From Geneva, the random sample consisted of 272 of 930 general practice physicians, 86 of 111 obstetrics and gynecology physicians, 96 of 128 pediatric physicians, and 46 of 52 dermatology and venereology physicians. Therefore, the Geneva sample of 500 physicians represents 40.95% of a total of 1221 physicians (see Table 1).

Table 1. Physician samples per region.

Specialty	Zurich		Geneva		Total	
	Total physicians found, N	Physicians selected for sample, n (%)	Total physicians found, N	Physicians selected for sample, n (%)	Total physicians found, N	Physicians selected for sample, n (%)
General practitioners	747	254 (34.0)	930	272 (29.2)	1677	526 (31.36)
Obstetrics and gynecology	109	85 (77.9)	111	86 (77.5)	220	171 (77.7)
Pediatrics	92	74 (80.4)	128	96 (75.0)	220	170 (77.3)
Dermatology and venereology	61	53 (86.8)	52	46 (88.5)	113	99 (87.6)
Total	1009	466 (46.18)	1221	500 (40.95)	2230	966 (43.32)

Data Collection

To identify PRWs on which patients can rate and review physicians in Switzerland, a systematic online search was conducted in June 2016 from a patient's perspective [22]. A website was included if it allowed users to view quantitative ratings and/or narrative comments about Swiss physicians in a structured manner without having to open an account or log onto the website. Websites that were not dedicated to Swiss physicians were excluded. A total of 3 PRWs were included: OkDoc, DocApp, and Medicosearch. In addition, Google itself allows users to rate and comment on physicians via Google reviews. Furthermore, although the health care information portal doktor does not provide the option for ratings, it links to Google reviews. Google was therefore also included in the study, and as far as this author is aware, this is the first time Google has been included in a study examining physician ratings internationally. The selected physicians were therefore searched for on a total of 4 websites: OkDoc, DocApp, Medicosearch, and Google. On each website, every selected physician was searched for between November 2017 and July 2018 and any narrative comments were recorded.

Data Analysis

The content of each narrative comment was analyzed and classified by the author according to a theoretical categorization framework of physician-, staff-, and practice-related issues. The categorization framework from Emmert et al was initially used [4], with modifications being made where necessary. This included removing categories that were not identified in the comments, adding categories that were identified but were not adequately covered by the previous framework, and separating categories (eg, friendliness and caring attitude) that were discussed in comments as distinct issues. Narrative comments were also classified as positive, neutral, and negative, overall. If a comment included both positive and negative aspects and no clear tendency could be determined, the comment was categorized as neutral. Narrative comments were analyzed in their original language. Descriptive statistics included means and standard deviations for continuous variables and percentages for categorical variables. To analyze whether differences exist between different groups, chi-squared tests were used for categorical data and *t* tests, for continuously distributed data. In relation to chi-squared tests with the 4 PRWs, posthoc tests using Bonferroni correction were conducted for the significant results to identify which PRW differed from the others. All analyses were performed with a significance level alpha set to .05 and 2-tailed tests, using Statistical Package for the Social Sciences (SPSS version 24 for Windows, IBM Corporation).

Results

Nature of Comments

The selected physicians in the sample had a total of 849 comments. Table 2 shows the breakdown of the number of comments by region, specialty, and gender. Overall, comments were significantly more likely to be regarding physicians in Zurich (668/849, 78.7%), specialists (545/849, 64%), and male physicians (477/849, 56.2%). However, there were important differences between PRWs. Although specialists (373/520, 71.7%) had significantly more comments on Google ($\chi 2_1$=98.2; P<.001), there were no significant differences between general practitioners and specialists on OkDoc, DocApp, or Medicosearch. Furthermore, although male physicians had more comments on okdoc (24/38, 63%) and significantly more ($\chi 2_1$=33.5; P<.001) on Google (326/520, 62.7%), female physicians had more comments on DocApp (30/57, 52%) and significantly more ($\chi 2_1$=4.9; P<.03) on Medicosearch (134/234, 57.3%).

The 849 comments had a mean length of 253.5 characters (SD 298), ranging from 15 to 3258 characters. There was a significant difference in the mean character length of the following groups:

- Positive comments (mean 222, SD 224) and negative comments (mean 436, SD 533); t_{130}=−4.4; P<.001.
- Physicians from Zurich (mean 231, SD 242) and physicians from Geneva (mean 335, SD 439); t_{210}=−3.1; P=.003.
- General practitioners (mean 193, SD 167) and specialists (mean 288, SD 347); t_{830}=−5.4; P<.001.

- Okdoc (mean 154, SD 126), DocApp (mean 296, SD 202), Medicosearch (mean 174, SD 146), and Google (mean 292, SD 354); F_3=10.4; P<.001.

However, there was no significant difference in the mean character length of male physicians (mean 256, SD 291) and female physicians (mean 250, SD 307); t_{847}=0.3; P=.77.

Categorization of Issues

The analysis of the 849 comments identified 43 subcategories addressing the physician (n=21), the staff (n=8), and the practice (n=14; see Textbox 1).

None of the PRWs' comments covered all 43 subcategories of the categorization framework (see Table 3); comments on Google covered 86% (37/43) of the subcategories, Medicosearch covered 72% (31/43), DocApp covered 60% (26/43), and OkDoc covered 55% (24/43).

Table 2. Physicians with comments.

Physician characteristics	OkDoc	DocApp	Medicosearch	Google	Total
Physician region					
Zurich, n/N (%)	20/38 (52.6)	56/57 (98.2)	206/234 (88)	386/520 (74.2)	668/849 (78.7)
Geneva, n/N (%)	18/38 (47.4)	1/57 (1.8)	28/234 (12)	134/520 (25.8)	181/849 (21.3)
Chi-squared test (*df*)	0.11 (1)	53.1 (*1*)	135.4 (*1*)	122.1 (*1*)	46.7 (*3*)
P value	.75	<.001	<.001	<.001	<.001
Physician specialty					
General practitioners, n/N (%)	23/38 (60.5)	28/57 (49.1)	108/234 (46.2)	147/520 (28.3)	306/849 (26)
Specialists, n/N (%)	15/38 (39.5)	29/57 (50.9)	126/234 (53.8)	373/520 (71.7)	543/849 (64)
Chi-squared test (*df*)	1.6 (*1*)	0.02 (*1*)	1.4 (*1*)	98.2 (*1*)	38.1(*3*)
P value	.19	.90	.24	<.001	<.001
Physician gender					
Male, n/N (%)	24/38 (63.2)	27/57 (47.4)	100/234 (42.7)	326/520 (62.7)	477/849 (56.2)
Female, n/N (%)	14/38 (36.8)	30/57 (52.6)	134/234 (57.3)	194/520 (37.3)	372/849 (43.8)
Chi-squared test (*df*)	2.6 (*1*)	0.20 (*1*)	4.9 (*1*)	33.5 (*1*)	28.7 (*3*)
P value	.11	.70	.03	<.001	<.001

Textbox 1. Categorization framework.

Physician (n=21)

- Overall assessment; Competence; Communication; Recommendation; Friendliness; Caring attitude; Satisfaction with treatment; Professionalism; Time spent with patient; Trust; Treatment cost/billing; Being taken seriously; Cooperation with medical specialists; Alternative medicine; Patient involvement; Telephone availability; Individualized service; House visits; Available outside normal hours; Privacy; Health insurance differentiation

Staff (n=8)

- Friendliness; Service/assistance; Overall assessment; Professionalism; Communication; Availability by telephone; Recommendation; Time spent with patient

Practice (n=14)

- Atmosphere; Waiting time within practice; Ability to get appointment; Overall assessment; Location; Organization; Equipment; Online appointment; Recommendation; Parking space; Consultation hours; Waiting room entertainment; Availability by telephone; Barrier-free access

Table 3. Subcategories covered by physician rating websites' comments.

Subcategories	OkDoc	DocApp	Medicosearch	Google
Physician (N=21), n (%)	16 (76)	14 (66)	17 (80)	18 (85)
Staff (N=8), n (%)	3 (37)	4 (50)	5 (62)	6 (75)
Practice (N=14), n (%)	5 (35)	8 (57)	9 (64)	13 (92)
Total (N=43), n (%)	24 (55)	26 (60)	31 (72)	37 (86)

In total, 2441 distinct issues were identified within the 43 subcategories of the categorization framework; 83.65% (2042/2441) of the issues were related to the physician, 6.63% (162/2441) related to the staff, and 9.70% (237/2441) related to the practice (see Table 4). Overall, the 2 most frequently issues mentioned were the overall assessment of the physician (300/849, 35.3%) and the physician's competence (300/849, 35.3%); the vast majority of these comments were positive (92.7% and 94.7%, respectively). Other frequently mentioned issues regarding the physician included 27.3% (232/849) of comments referred to the physician's communication (84.9% positive); 26.5% (225/849) recommended the physician (86.2% positive); 25.3% (215/849) referred to the physician's friendliness (88.8% positive); 22.6% (192/849) referred to the physician's caring attitude (87.5% positive); 17.6% (149/849) referred to satisfaction with treatment (79.2% positive); 15.2% (129/849) referred to the physician's professionalism (76.7% positive); 12.6% (107/849) referred to time spent with the patient (87.9% positive); and 9.7% (82/849) referred to the physician's trustworthiness (89% positive). In relation to staff issues, the most frequently mentioned issue was regarding the staff's friendliness (92/849, 10.8%), 84.8% of which were positive. Concerning practice issues, frequently mentioned issues included 6.9% (59/849) of comments that mentioned the atmosphere of the practice (91.5% positive), 6.8% (58/849) the waiting time within the practice (72.4% positive), and 4.6% (39/849) the ability to get an appointment (79.5% positive). Negative comments most frequently referred to treatment cost or billing (32/43, 74%), communication with the staff (7/13, 53%), the staff's professionalism (4/15, 26%), waiting time within practice (12/58, 20%), ability to get an appointment (8/39, 20%), the physician's professionalism (26/129, 20.2%), and satisfaction with treatment (27/149, 18.1%).

However, there were some significant differences between PRWs, regions, specialties, and gender (see Multimedia Appendices 1-4 for full results). Regarding the PRWs, there were significant differences between comments on Google and the 3 dedicated PRWs in a number of subcategories. For instance, there were significant differences between PRWs in relation to comments mentioning the physician's competency ($\chi2_3$=11.4; *P*=.01; V=0.12). Posthoc tests using Bonferroni correction revealed that comments on Google (161/520, 31.0%) mentioned the physician's competence significantly less than the overall sample. There were also significant differences between PRWs regarding satisfaction with treatment ($\chi2_3$=16.9; *P*<.001; V=0.14). Posthoc tests using Bonferroni correction revealed that comments on Google (111/520; 21.3%) referred to satisfaction with treatment significantly more than the overall sample. Furthermore, 97% (42/43) of the references to treatment cost or billing issues were made in comments from Google.

There were significant differences between comments regarding physicians from Zurich and Geneva in a number of subcategories. For instance, comments regarding physicians from Zurich mentioned the physician's competence (263/668, 39.3%) significantly more often ($\chi2_1$=22.3; *P*<.001) than comments from physicians from Geneva (37/181, 20.4%). However, physicians from Geneva had a higher percentage of comments referring to the physician's communication (60/181, 33.1% vs 172/668, 25.7%), the physician's caring attitude (50/181, 27.6% vs 142/668, 21.2%), the physician's professionalism (39/181, 21.5% vs 90/668, 13.4%), and trust in the physician (24/181, 13.2% vs 58/668, 8.6%). Comments regarding specialists significantly more often recommended the physician ($\chi2_1$=8.6; *P*=.004), the physician's caring attitude ($\chi2_1$=8.6; *P*=.004), satisfaction with treatment ($\chi2_1$=9.9; *P*=.002), treatment cost and billing ($\chi2_1$=9.6; *P*=.002), staff friendliness ($\chi2_1$=12.2; *P*<.001), and practice atmosphere ($\chi2_1$=6.8; *P*=.01). Comments regarding male physicians (102/477, 21.3%) were significantly more ($\chi2_1$=11.1; *P*=.001) likely to refer to satisfaction with treatment than comments about female physicians (47/372, 12.6%). However, comments regarding female physicians (22/372, 5.9%) were significantly more ($\chi2_1$=11.0; *P*=.001) likely to mention that the patient felt like they had been taken seriously than comments about male physicians (8/477, 1.6%).

Evaluation Results

Overall, 83.0% (705/849) of comments were classified as positive, 2.5% (21/849) as neutral, and 14.5% (123/849) as negative (see Table 5). However, there were significant differences between PRWs, regions, and specialty regarding negative comments: 90.2% (111/123) of negative comments were on Google ($\chi2_2$=180.1; *P*<.001), 74.7% (92/123) were regarding physicians in Zurich ($\chi2_1$=30.3; *P*<.001), and 73.2% (90/123) were regarding specialists ($\chi2_1$=26.4; *P*<.001). There was no significant difference ($\chi2_1$=2.4; *P*=.13) between males (70/123, 56.9%) and females (53/123, 43.1%) regarding negative comments.

Table 4. Categorization of issues.

Issue	Total (N=849), n (%)	Evaluation		
		Positive, n (%)	Neutral, n (%)	Negative, n (%)
Physician				
Overall assessment	300 (35.3)	278 (92.7)	7 (2.3)	15 (5.0)
Competence	300 (35.3)	284 (94.7)	5 (1.7)	11 (3.7)
Communication	232 (27.3)	197 (84.9)	2 (0.9)	33 (14.2)
Recommendation	225 (26.5)	194 (86.2)	0 (0.0)	31 (13.8)
Friendliness	215 (25.3)	191 (88.8)	5 (2.3)	19 (8.8)
Caring attitude	192 (22.6)	168 (87.5)	3 (1.6)	21 (10.9)
Satisfaction with treatment	149 (17.6)	118 (79.2)	4 (2.7)	27 (18.1)
Professionalism	129 (15.2)	99 (76.7)	4 (3.1)	26 (20.2)
Time spent with patient	107 (12.6)	94 (87.9)	2 (1.9)	11 (10.3)
Trust	82 (9.7)	73 (89)	0 (0)	9 (11)
Treatment cost/billing	43 (5.1)	10 (23)	1 (2)	32 (74)
Being taken seriously	30 (3.5)	25 (83)	0 (0)	5 (16)
Cooperation with medical specialists	11 (1.3)	11 (100)	0 (0)	0 (0)
Alternative medicine	5 (0.6)	5 (100)	0 (0)	0 (0)
Patient involvement	5 (0.6)	5 (100)	0 (0)	0 (0)
Telephone availability	5 (0.6)	4 (80)	0 (0)	1 (20)
Individualized service	4 (0.5)	4 (100)	0 (0)	0 (0)
House visits	3 (0.4)	3 (100)	0 (0)	0 (0)
Available outside normal hours	2 (0.2)	2 (100)	0 (0)	0 (0)
Privacy	2 (0.2)	2 (100)	0 (0)	0 (0)
Health insurance differentiation	1 (0.1)	0 (0)	0 (0)	1 (100)
Staff				
Friendliness	92 (10.8)	78 (84)	6 (6)	8 (8)
Service/assistance	19 (2.2)	17 (89)	0 (0)	2 (10)
Overall assessment	18 (2.1)	16 (88)	1 (5)	1 (5)
Professionalism	15 (1.8)	10 (66)	1 (6)	4 (26)
Communication	13 (1.5)	5 (38)	1 (7)	7 (53)
Availability by telephone	3 (0.4)	3 (100)	0 (0)	0 (0)
Recommendation	1 (0.1)	1 (100)	0 (0)	0 (0)
Time spent with patient	1 (0.1)	1 (100)	0 (0)	0 (0)
Practice				
Atmosphere	59 (6.9)	54 (91)	3 (5)	2 (3)
Waiting time within practice	58 (6.8)	42 (72)	4 (6)	12 (20)
Ability to get appointment	39 (4.6)	31 (79)	0 (0)	8 (20)
Overall assessment	22 (2.6)	20 (90)	1 (4)	1 (4)
Location	15 (1.8)	13 (86)	0 (0)	2 (13)
Organization	13 (1.5)	10 (76)	1 (7)	2 (15)
Equipment	9 (1.1)	8 (88)	0 (0)	1 (11)
Online appointment	5 (0.6)	5 (100)	0 (0)	0 (0)
Recommendation	5 (0.6)	5 (100)	0 (0)	0 (0)

Issue	Total (N=849), n (%)	Evaluation		
		Positive, n (%)	Neutral, n (%)	Negative, n (%)
Parking space	5 (0.6)	5 (100)	0 (0)	0 (0)
Consultation hours	2 (0.2)	2 (100)	0 (0)	0 (0)
Waiting room entertainment	2 (0.2)	2 (100)	0 (0)	0 (0)
Availability by telephone	2 (0.2)	1 (50)	0 (0)	1 (50)
Barrier free access	1 (0.1)	0 (0)	1 (100)	0 (0)

Table 5. Evaluation results.

Region	OkDoc, n/N (%)	DocApp, n/N (%)	Medicosearch, n/N (%)	Google, n/N (%)	Total, n/N (%)
Zurich					
Positive	19/20 (95)	54/56 (96)	192/206 (93.2)	293/386 (74.9)	558/668 (83.5)
Neutral	1/20 (5)	0/56 (0)	5/206 (2.4)	12/386 (3.1)	18/668 (2.7)
Negative	0/20 (0)	2/56 (3)	9/206 (4.4)	81/386 (21)	92/668 (13.8)
Geneva					
Positive	18/18 (100)	1/1 (100)	27/28 (96)	101/134 (75.4)	147/181 (81.2)
Neutral	0/18 (0)	0/1 (0)	0/28 (0)	3/134 (2.2)	3/181 (1.7)
Negative	0/18 (0)	0/1 (0)	1/28 (3)	30/134 (22.4)	31/181 (17.1)
Total					
Positive	37/38 (97)	55/57 (96)	219/234 (93.6)	394/520 (75.8)	705/849 (83)
Neutral	1/38 (2)	0/57 (0)	5/234 (2.1)	15/520 (2.9)	21/849 (2.5)
Negative	0/38 (0)	2/57 (3)	10/234 (4.3)	111/520 (21.3)	123/849 (14.5)

Discussion

As far as this author is aware, this is the first study to examine the content and nature of narrative comments on Swiss PRWs and has resulted in a number of key findings: (1) the vast majority of issues mentioned were concerning aspects of performance (interpersonal skills of physician and staff, infrastructure, and organization and management of practice) that are considered assessable by patients; (2) overall, the vast majority of comments were positive; and (3) there were significant differences between comments on Google and comments on the 3 dedicated PRWs.

Content of Comments

The 5 most frequently mentioned issues identified from the narratives comments were (1) the overall assessment of the physician (300/849, 35.3%) and the physician's competence (300/849, 35.3%); (2) the physician's communication (232/849, 27.3%); (3) recommending the physician (225/849, 26.5%); (4) the physician's friendliness (215/849, 25.3%); and 5) the physician's caring attitude (192/849, 22.6%). In contrast, the top 5 mentioned issues identified by Emmert et al's analysis of 3000 narrative comments from the German PRW, jameda, from 2012 were as follows: (1) the physician's competence (1874/3000, 62.46%); (2) the physician's friendliness and caring attitude (1148/3000, 38.26%); (3) the time the physician spent with the patient (987/3000, 32.90%); (4) the friendliness of the staff (667/3000, 22.23%); and (5) the information and advice from the physician (630/3000, 21.00%) [4].

Although both studies found that narrative comments most frequently mentioned the physician's competence, it should be noted that while this study kept the issues of *the physician's friendliness* and *the physician's caring attitude* separate, Emmert et al combined the 2 issues [4]. If this study also combined these 2 issues, the physician's friendliness and caring attitude would become the most frequently mentioned issue (407/849, 47.9%). Indeed, it is important to recognize that 95% (41/43) of the subcategories of the categorization framework and 81.60% (1992/2441) of the distinct issues identified were concerning aspects of performance (interpersonal skills of physician and staff, infrastructure, and organization and management of practice) that are considered to be assessable by patients [21]. Although a number of narrative comments also mentioned the physician's competency (300/849, 35.3%), the proportion of comments that mentioned this issue were substantially lower than that reported by Emmert et al (62.5%) [4].

Unsolicited critical comments on PRWs can be seen as a type of complaint, which can offer a *window of opportunity* to improve health services [27]. Indeed, one of the aims of PRWs is to drive quality improvement by identifying aspects of care needing improvement so that changes can be made in practice [10]. Overall, 123 comments were classified as negative. Within these negative comments, 293 distinct issues were identified. Nearly half of all negative issues (132/293, 45.1%) concerned interpersonal issues: the physician's communication (n=33), the physician's friendliness (n=19), the physician's caring attitude (n=21), the physician's professionalism (n=26), the physician's trustworthiness (n=9), being taken seriously by the physician (n=5), the friendliness of the staff (n=8), the professionalism of staff (n=4), and staff communication (n=7). Given these interpersonal issues make up nearly half of all negative issues, and that improving these issues will potentially also improve the patient's overall assessment and recommendation of physicians (46/293, 15.7% of negative issues), it is recommended that physicians should focus on improving interpersonal interactions with patients. However, the health care setting can be a very stressful and emotionally draining environment because of external (including workload, exposure to patient suffering, time pressures, documentation requirement, and financial issues) and internal (including personality characteristics and poor emotional regulation) factors [28]. This can lead to stress, dissatisfaction, increased cynicisms, burnout, and compassion fatigue among health care professionals and the staff [28,29]. In recent decades, the Swiss health care system has experienced a number of changes that have caused greater economic constraints, increased administrative workload, and decreased professional autonomy [30]. A study published in 2010 found that burnout levels among Swiss physicians had increased throughout the country over the last decade [30]. The increased burnout levels among Swiss physicians may be contributing to the suboptimal interpersonal issues reported in the narrative comments. Although there are strategies that individual physicians can use to improve their interpersonal skills [28], to really address this issue, whole-system approaches may be required to improve the well-being of physicians [29].

Nature of Comments

The analysis of the 849 narrative comments on Swiss PRWs reveals that 83% (705/849) of all comments were positive, 2.5% (21/849) were neutral, and 14.5% (123/849) were negative. This finding is very similar to the previous analysis of narrative comments on PRWs in other countries [4,8,12-15]. For example, Emmert et al's analysis of 3000 narrative comments from the German PRW, jameda, from 2012 found that 80% of all comments were positive, 4% were neutral, and 16% were negative [4]. Although this finding suggests that the users of Swiss PRWs are mostly satisfied with their physicians, the veracity of the level of satisfaction must be called into question given the explicit practice of the dedicated PRWs of not allowing negative comments or removing them if physicians do not want them published. On OkDoc, 0 of the 38 comments were negative; on DocApp, 2 of the 57 comments were negative; and on Medicosearch, 10 of the 234 comments were negative. Although Google had 90.2% (111/123) of negative comments, the author has become aware that some negative comments that were online during data collection have since been removed. It is, therefore, unclear how many negative comments are being supressed on Swiss PRWs. However, the current suppression of negative comments by Swiss PRWs is concerning and goes against their overall aim of achieving more transparency. There are, no doubt, challenges in finding the correct balance between protecting physicians from harm and promoting the health literacy benefits for patients. However, a blanket ban on negative comments or removing comments simply because the physician in question does not like a particular comment seems inappropriate and is leading to a biased and inaccurate picture of patients' experiences and satisfaction. There is a need for a consensus-based criterion that applies to all Swiss PRWs for determining which comments are to be and not to be published publically and which are clearly publicized so users of PRWs are aware of it. Indeed, a recent qualitative study conducted with a random sample of residents of 4 North German cities reported that a lack of rating guidance in terms of allowable content was a disincentive for rating a physician on a PRW [31]. It is also likely that the removal of a comment on the whims of a PRW operator is a disincentive for users to give further physician ratings in the future. Such criteria should also address how comments that PRWs suspect to be *fake reviews* should be handled, as there is some indication that physicians or practice staff sometimes pose as patients on PRWs to post either positive comments about themselves or negative comments about competitors [32].

Google

As far as this author is aware, this is the first time Google has been included in a project examining physician ratings internationally. It has already been reported that Google had the highest average number of quantitative ratings (3.74 ratings) and narrative comment (3.03 comments) ratings per identifiable physician [22]. This analysis of the content and nature of the narrative comments on Swiss PRWs reveals that the comments on Google are also far richer than the comments on the other Swiss PRWs; comments on Google covered the most subcategories of the categorization framework (37/43, 86%) and also had the majority of negative comments (111/123, 90.2%). It, therefore, appears that Google has not only become the most used website in Switzerland for physician ratings in recent years but is also potentially the most useful. It would be helpful if Swiss patients are made aware of the current large differences between Swiss PRWs regarding the frequency and nature of ratings to help them determine which PRW will provide them with the most useful information. However, future updates would be helpful to assess whether Google, given its general market dominance, will take an even bigger share of the PRW ratings away from the dedicated PRW competitors, or whether the dedicated PRWs will be able to increase the quantity and quality of ratings. Indeed, Medicosearch has already started to shift its business strategy toward online appointments, something that Google does not currently offer, which may allow them to gain a bigger market share and increase the number of ratings. However, it may be necessary for OkDoc to reflect on whether their continued existence in the Swiss PRW market is providing value or is in fact causing harm. It has already been reported that OkDoc had the lowest average number of quantitative ratings (1.47 ratings) and narrative comment (1.23 comments) ratings per identifiable physician, and it only had one comment posted for all 966 physicians in the sample during the last five and a half years (2012-2018) [22]. This analysis of the content and nature of the narrative comments has also found that OkDoc covered the least amount of subcategories of the categorization framework (24/43, 55%) and that it does not have any negative comments.

Limitations

This study has a number of limitations that should be taken into account when interpreting the results. First, although a systematic online search of Swiss PRWs was conducted, there may be other types of websites that allow Swiss physicians to be rated that were not included in this study. This is a fast-moving area, and it does appear that there are some websites that have started allowing ratings or making ratings publically available after this project had commenced (eg, deindoktor and doctena), which should be added to any future studies examining PRWs in Switzerland. Second, only German search terms were used for the systematic online search of Swiss PRWs. Although the author is confident that no important Swiss PRWs were missed at the time of developing and conducting the project, it would be preferable if French and Italian search terms are also included in future research in Switzerland to ensure that no PRWs are being missed. Third, the sample was only taken from 2 regions in Switzerland, which may limit the generalizability of the results. Although the study used a representative random sample from a German-speaking and French-speaking region of Switzerland with a comparable number of physicians, given the significant differences found between the 2 regions, it would be helpful for further research to include other regions to examine whether these differences are found between other German- and French-speaking regions and in the Italian-speaking region of Ticino. Fourth, a distinction was only made between general practitioners and specialists, and there may be further differences between the different specialties. Finally, the sociodemographic information of the rating patients is unknown and may not be representative of Swiss patients in general.

Acknowledgments

This study was funded by the Swiss Academy of Medical Sciences’ Käthe-Zingg-Schwichtenberg-Fonds, which had no role in the project design, in the collection, analysis, or interpretation of data, in the writing of the paper, or in the decision to submit the paper for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Categorization of issues by physician rating websites.
[PDF File (Adobe PDF File), 248 KB - jmir_v21i9e14336_app1.pdf]

Multimedia Appendix 2
Categorization of issues by regions.
[PDF File (Adobe PDF File), 208 KB - jmir_v21i9e14336_app2.pdf]

Multimedia Appendix 3
Categorization of issues by specialties.
[PDF File (Adobe PDF File), 205 KB - jmir_v21i9e14336_app3.pdf]

Multimedia Appendix 4
Categorization of issues by gender.
[PDF File (Adobe PDF File), 206 KB - jmir_v21i9e14336_app4.pdf]

References

1. Emmert M, Sander U, Esslinger AS, Maryschok M, Schöffski O. Public reporting in Germany: the content of physician rating websites. Methods Inf Med 2012;51(2):112-120. [doi: 10.3414/ME11-01-0045] [Medline: 22101427]
2. Emmert M, Meier F. An analysis of online evaluations on a physician rating website: evidence from a German public reporting instrument. J Med Internet Res 2013 Aug 6;15(8):e157 [FREE Full text] [doi: 10.2196/jmir.2655] [Medline: 23919987]
3. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. J Med Internet Res 2013 Feb 1;15(2):e24 [FREE Full text] [doi: 10.2196/jmir.2360] [Medline: 23372115]
4. Emmert M, Meier F, Heider A, Dürr C, Sander U. What do patients say about their physicians? An analysis of 3000 narrative comments posted on a German physician rating website. Health Policy 2014 Oct;118(1):66-73. [doi: 10.1016/j.healthpol.2014.04.015] [Medline: 24836021]
5. Hennig-Thurau T, Gwinner KP, Walsh G, Gremler DD. Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? J Interact Mark 2004 Jan;18(1):38-52. [doi: 10.1002/dir.10073]

6. Boulos MN, Wheeler S. The emerging web 2.0 social software: an enabling suite of sociable technologies in health and health care education. Health Info Libr J 2007 Mar;24(1):2-23 [FREE Full text] [doi: 10.1111/j.1471-1842.2007.00701.x] [Medline: 17331140]
7. Terlutter R, Bidmon S, Röttl J. Who uses physician-rating websites? Differences in sociodemographic variables, psychographic variables, and health status of users and nonusers of physician-rating websites. J Med Internet Res 2014 Mar 31;16(3):e97 [FREE Full text] [doi: 10.2196/jmir.3145] [Medline: 24686918]
8. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
9. Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. Med Care 2009 Jan;47(1):1-8. [doi: 10.1097/MLR.0b013e3181808bb5] [Medline: 19106724]
10. Strech D. Ethical principles for physician rating sites. J Med Internet Res 2011 Dec 6;13(4):e113 [FREE Full text] [doi: 10.2196/jmir.1899] [Medline: 22146737]
11. Rothenfluh F, Schulz PJ. Content, quality, and assessment tools of physician-rating websites in 12 countries: quantitative analysis. J Med Internet Res 2018 Jun 14;20(6):e212 [FREE Full text] [doi: 10.2196/jmir.9105] [Medline: 29903704]
12. Ellimoottil C, Hart A, Greco K, Quek ML, Farooq A. Online reviews of 500 urologists. J Urol 2013 Jun;189(6):2269-2273. [doi: 10.1016/j.juro.2012.12.013] [Medline: 23228385]
13. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. J Gen Intern Med 2012 Jun;27(6):685-692 [FREE Full text] [doi: 10.1007/s11606-011-1958-4] [Medline: 22215270]
14. Black EW, Thompson LA, Saliba H, Dawson K, Black NM. An analysis of healthcare providers' online ratings. Inform Prim Care 2009;17(4):249-253 [FREE Full text] [doi: 10.14236/jhi.v17i4.744] [Medline: 20359403]
15. Lagu T, Goff SL, Hannon NS, Shatz A, Lindenauer PK. A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States. Jt Comm J Qual Patient Saf 2013 Jan;39(1):7-15. [doi: 10.1016/S1553-7250(13)39003-5] [Medline: 23367647]
16. Li J, Liu M, Li X, Liu X, Liu J. Developing embedded taxonomy and mining patients' interests from web-based physician reviews: mixed-methods approach. J Med Internet Res 2018 Aug 16;20(8):e254 [FREE Full text] [doi: 10.2196/jmir.8868] [Medline: 30115610]
17. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. J Med Internet Res 2013 Nov 1;15(11):e239 [FREE Full text] [doi: 10.2196/jmir.2721] [Medline: 24184993]
18. Carvel J. The Guardian. 2008. Patients to Rate and Review Their GPs on NHS Website URL: https://www.theguardian.com/society/2008/dec/30/doctors-rating-website-nhs [accessed 2017-03-03] [WebCite Cache ID 6ogmjtvkj]
19. Ärztlicher Kreisverband Ebersberg. 2009. Ärztlicher Kreisverband Ebersberg Körperschaft Des Öffentlichen Rechts URL: https://www.aekv-ebersberg.de/aktuelles/105-baek-newsletter-2009-004-vom-25062009 [accessed 2017-03-03] [WebCite Cache ID 6ogmznsHJ]
20. Patel S, Cain R, Neailey K, Hooberman L. General practitioners' concerns about online patient feedback: findings from a descriptive exploratory qualitative study in England. J Med Internet Res 2015 Dec 8;17(12):e276 [FREE Full text] [doi: 10.2196/jmir.4989] [Medline: 26681299]
21. Rothenfluh F, Schulz PJ. Physician rating websites: what aspects are important to identify a good doctor, and are patients capable of assessing them? A mixed-methods approach including physicians' and health care consumers' perspectives. J Med Internet Res 2017 May 1;19(5):e127 [FREE Full text] [doi: 10.2196/jmir.6875] [Medline: 28461285]
22. McLennan S. Quantitative ratings and narrative comments on Swiss physician rating websites: frequency analysis. J Med Internet Res 2019 Jul 26;21(7):e13816 [FREE Full text] [doi: 10.2196/13816] [Medline: 31350838]
23. Roos N. Schweizer Radio und Fernsehen. 2014. Ärzte-Bewertungen im Internet Noch Wenig Aussagekräftig URL: https://www.srf.ch/sendungen/kassensturz-espresso/themen/gesundheit/aerzte-bewertungen-im-internet-noch-wenig-aussagekraeftig [accessed 2019-02-14] [WebCite Cache ID 76BgL5ONT]
24. Medinside - Das Portal für die Gesundheitsbranche. 2016. Immer Mehr Ärzte Sind Online-bewertungen Gegenüber Aufgeschlossen URL: https://www.medinside.ch/de/post/immer-mehr-aerzte-sind-online-bewertungen-gegenueber-aufgeschlossen [accessed 2019-02-14] [WebCite Cache ID 76BgVl2iK]
25. Atasoy S. Schweizer Radio und Fernsehen. 2014. Patienten Benoten Schweizer Ärzte URL: https://www.srf.ch/news/schweiz/patienten-benoten-schweizer-aerzte [accessed 2019-02-14] [WebCite Cache ID 76Bh6rFqc]
26. Troxler R. Luzerner Zeitung. 2018. Online-bewertungen: Bewertung Im Internet: Fünf Sterne Für Den Doktor URL: https://www.luzernerzeitung.ch/zentralschweiz/luzern/online-bewertungen-bewertung-im-internet-fuenf-sterne-fuer-den-doktor-ld.94116 [accessed 2019-02-14] [WebCite Cache ID 76BhINS1C]
27. Paterson R. The patients' complaints system in New Zealand. Health Aff (Millwood) 2002;21(3):70-79. [doi: 10.1377/hlthaff.21.3.70] [Medline: 12026005]

28. Amutio-Kareaga A, García-Campayo J, Delgado LC, Hermosilla D, Martínez-Taboada C. Improving communication between physicians and their patients through mindfulness and compassion-based strategies: a narrative review. J Clin Med 2017 Mar 17;6(3):E33 [FREE Full text] [doi: 10.3390/jcm6030033] [Medline: 28304333]
29. Brand SL, Coon JT, Fleming LE, Carroll L, Bethel A, Wyatt K. Whole-system approaches to improving the health and wellbeing of healthcare workers: a systematic review. PLoS One 2017;12(12):e0188418 [FREE Full text] [doi: 10.1371/journal.pone.0188418] [Medline: 29200422]
30. Arigoni F, Bovier PA, Sappino AP. Trend of burnout among Swiss doctors. Swiss Med Wkly 2010;140:w13070. [doi: 10.4414/smw.2010.13070] [Medline: 20809437]
31. McLennan S, Strech D, Kahrass H. Why are so few patients rating their physicians on German physician rating websites? A qualitative study. BMC Health Serv Res 2018 Aug 29;18(1):670 [FREE Full text] [doi: 10.1186/s12913-018-3492-0] [Medline: 30157842]
32. Bodkin H. The Daily Telegraph. 2018. GPs Are Posing as Patients and Posting 'Fake Reviews' Online, Health Chiefs Reveal URL: https://www.telegraph.co.uk/news/2018/07/16/gps-posting-fake-reviews-online-health-chiefs-reveal/ [accessed 2019-06-28]

Abbreviations

PRW: physician rating website

Edited by C Eichenberg; submitted 10.04.19; peer-reviewed by S Bidmon; comments to author 24.05.19; revised version received 28.06.19; accepted 15.07.19; published 30.09.19.

Please cite as:
McLennan S
The Content and Nature of Narrative Comments on Swiss Physician Rating Websites: Analysis of 849 Comments
J Med Internet Res 2019;21(9):e14336
URL: http://www.jmir.org/2019/9/e14336/
doi:10.2196/14336
PMID:31573918

Original Paper

Cost-Effectiveness of Therapist-Guided Internet-Based Cognitive Behavioral Therapy for Stress-Related Disorders: Secondary Analysis of a Randomized Controlled Trial

Elin Lindsäter[1], MSc; Erland Axelsson[1], PhD; Sigrid Salomonsson[1,2], PhD; Fredrik Santoft[1], PhD; Brjánn Ljótsson[1,2], PhD; Torbjörn Åkerstedt[1,3], PhD, Prof Dr; Mats Lekander[1,3,4], PhD, Prof Dr; Erik Hedman-Lagerlöf[1,4], PhD

[1]Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[2]Center for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[3]Stress Research Institute, Stockholm University, Stockholm, Sweden
[4]Osher Center for Integrative Medicine, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Elin Lindsäter, MSc
Division of Psychology
Department of Clinical Neuroscience
Karolinska Institutet
Nobels väg 9
17165 Solna
Stockholm,
Sweden
Phone: 46 735635233
Email: elin.lindsater@ki.se

Abstract

Background: Stress-related disorders are associated with significant suffering, functional impairment, and high societal costs. Internet-based cognitive behavioral therapy (ICBT) is a promising treatment for stress-related disorders but has so far not been subjected to health economic evaluation.

Objective: The objective of this study was to evaluate the cost-effectiveness and cost-utility of ICBT for patients with stress-related disorders in the form of adjustment disorder (AD) or exhaustion disorder (ED). We hypothesized that ICBT, compared with a waitlist control (WLC) group, would generate improvements at low net costs, thereby making it cost-effective.

Methods: Health economic data were obtained in tandem with a randomized controlled trial of a 12-week ICBT in which patients (N=100) were randomized to an ICBT (n=50) or a WLC (n=50) group. Health outcomes and costs were surveyed pre- and posttreatment. We calculated incremental cost-effectiveness ratios (ICERs) based on remission rates and incremental cost-utility ratios (ICURs) based on health-related quality of life. Bootstrap sampling was used to assess the uncertainty of our results.

Results: The ICER indicated that the most likely scenario was that ICBT led to higher remission rates compared with the WLC and was associated with slightly larger reductions in costs from pre- to posttreatment. ICBT had a 60% probability of being cost-effective at a willingness to pay (WTP) of US $0 and a 96% probability of being cost-effective at a WTP of US $1000. The ICUR indicated that ICBT also led to improvements in quality of life at no net societal cost. Sensitivity analyses supported the robustness of our results.

Conclusions: The results suggest that ICBT is a cost-effective treatment for patients suffering from AD or ED. Compared with no treatment, ICBT for these patients yields large effects at no or minimal societal net costs.

Trial Registration: ClinicalTrials.gov NCT02540317; https://clinicaltrials.gov/ct2/show/NCT02540317

(J Med Internet Res 2019;21(9):e14675) doi:10.2196/14675

KEYWORDS

psychological stress; adjustment disorder; exhaustion disorder; cognitive behavioral therapy; cost effectiveness; internet

Introduction

Background

Stress is considered to be one of the largest health challenges in the Western world [1]. Prolonged or repeated exposure to stress is associated with negative physical and mental health outcomes, decreased quality of life, and functional impairment [2,3]. The costs for employers and society are large because of high incidence of sickness absence, reduced productivity at work, and significant loss of potential labor supply [4]. For the afflicted individual, chronic stress can lead to high costs because of increased medical and insurance expenses and decreased income [1]. Taken together, the full scope of chronic stress can negatively impact a country's gross domestic product, with conservative estimates of annual costs ranging from US $23 billion in the European Union [1] to US $42 billion in the United States [5].

When clinically significant symptoms and functional impairment result as a consequence of chronic or intermittent life stressors, diagnoses such as adjustment disorder (AD) or exhaustion disorder (ED) may be warranted [3,6]. AD is one of the most commonly used diagnoses by clinical psychologists and psychiatrists worldwide [7,8], cited by the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) as being the principal diagnosis for 5% to 20% of mental health outpatients [9]. ED is a specification of the diagnostic category F43.8 (*other reactions to severe stress*) in the International Statistical Classification of Diseases and Related Health Problems, 10th edition [10] that has been accepted as a formal diagnosis by the Swedish National Board of Health and Welfare. Similar to the concepts of clinical burnout and allostatic overload [11,12], ED is characterized by severe mental and physical exhaustion, cognitive dysfunction, disturbed sleep, and somatoform complaints [3]. Compared with AD, ED is a more chronic and debilitating condition [13], but both are based on the longitudinal course of symptoms and behavioral changes in the context of stressful life events. Given the strain on health care systems to handle these stress-related disorders, health economic evaluations are important to provide decision makers with information about which treatments give maximum effect in relation to their cost [14]. Without knowledge of cost-effectiveness, there is a risk that health care resources are used inefficiently and that fewer can be offered treatment [15].

Cognitive behavioral therapy (CBT) is the most rigorously evaluated and effective psychological treatment for a range of mental health problems [16], including stress and stress-related symptoms [17,18]. To meet the high demand for CBT, delivering treatment via the internet has emerged as a viable option to increase treatment accessibility and reduce delivery costs. Therapist-guided internet-based CBT (ICBT) has been shown to be highly efficacious for many clinical conditions [19], often producing effect sizes in parity with face-to-face treatment while at the same time often requiring less than 10 min of therapist time weekly per patient [20]. A recent meta-analyses of Web- and computer-based interventions to reduce stress in diverse samples indicated that ICBT could be effective in reducing stress with small to moderate effect sizes [21]. Although these findings cannot be directly generalized to clinical populations suffering from stress-related disorders, we recently presented evidence that ICBT can also be effective in reducing symptoms in a clinical sample diagnosed with AD or ED [13].

Meta-analyses indicate that ICBT can be a cost-effective alternative for many clinical conditions [19,22]. Regarding interventions to reduce stress, however, health economic evaluations are scarce and generally suffer from low methodological quality [23]. A total of 2 recent studies have conducted health economic evaluations of internet-based stress-management interventions aimed at reducing stress in employees with elevated work-related stress [24,25]. Both the studies indicated that the interventions were cost-effective compared with waitlist control (WLC) conditions, from an employer's [24] and a full societal perspective [25]. Although these results are promising, to the best of our knowledge, no study to date has investigated the cost-effectiveness of ICBT for patients actually diagnosed with stress-related disorders.

Objectives

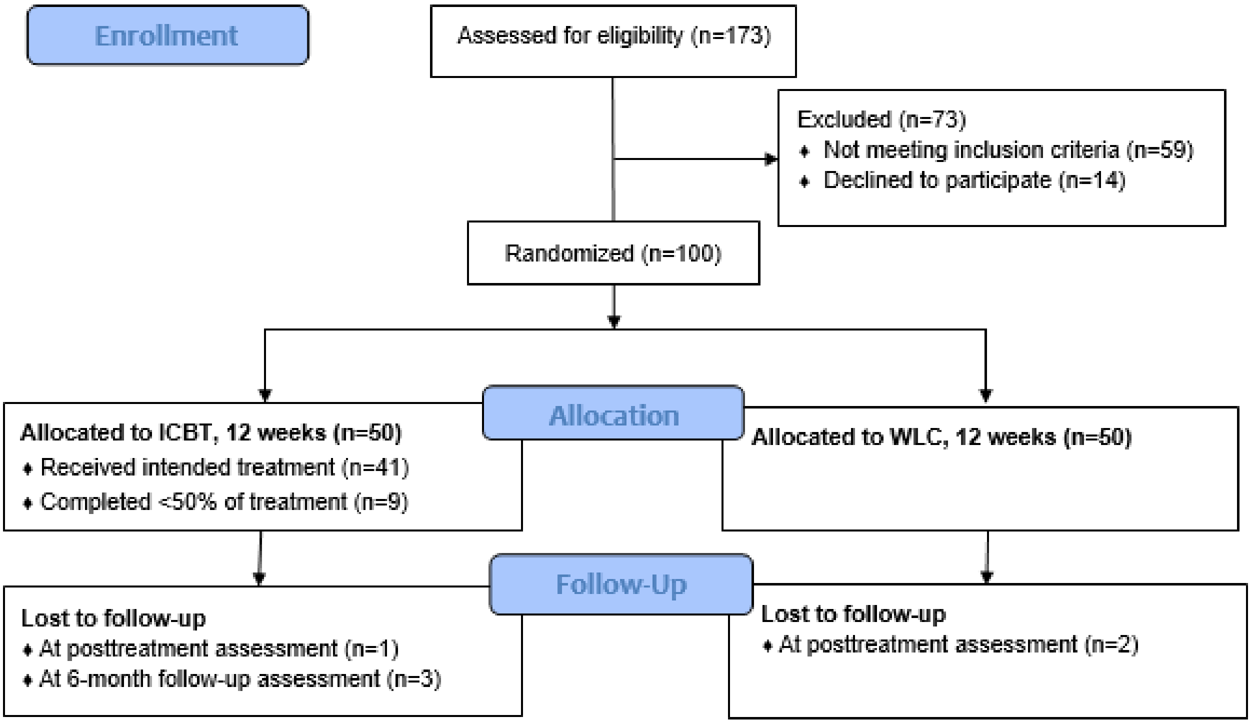
The aim of this study was to evaluate the cost-effectiveness and cost-utility of ICBT for stress-related disorders within the context of a randomized controlled trial [13], using both societal and health care perspectives. We hypothesized that ICBT, compared with a WLC group, would generate improvements at low net costs, thereby making the treatment cost-effective.

Methods

Design

Health economic data were collected within a randomized controlled trial [13] in which adults suffering from stress-related disorders (N=100) were randomized to either an ICBT (n=50) or a WLC (n=50) group, each of 12 weeks duration. Stress-related disorders were defined as a clinical diagnosis of AD according to DSM-5 [9] or ED as defined by the Swedish National Board of Health and Welfare [3,10]. Randomization was stratified by diagnosis (AD vs ED) and took place after inclusion to prevent allocation bias. Health economic data were collected at pre- and posttreatment assessment (after 12 weeks) as well as at a 6-month follow-up (6MFU). As participants in the WLC were crossed over to treatment after 12 weeks, cost-effectiveness and cost-utility were assessed from the pretreatment to the posttreatment assessment only. However, data from the 6MFU are presented to allow for crude and uncontrolled estimates of stability in clinical outcomes and costs over time. The trial was conducted at Karolinska Institutet, Stockholm, Sweden, between September 2015 and August 2016. It was approved by the Regional Ethics Review Board in Stockholm, Sweden (2015/415-31/5) and preregistered at ClinicalTrials.gov (ID: NCT02540317). All participants provided verbal and written informed consent for study participation. Figure 1 illustrates the trial design and study flow.

Figure 1. Flowchart illustrating trial design, adherence, and attrition. Note: The WLC was crossed over to treatment after the 12-week main phase of the trial and did not take part in the 6-month follow-up assessment. All health economic analyses are, therefore, based on the main phase of the trial. ICBT: internet-based cognitive behavioral therapy; WLC: waitlist control.



Study Sample

Detailed information about the recruitment procedure can be found in the study by Lindsäter et al [13]. The study was advertised through an ad in a national newspaper and through social media. Applicants were self-referred via a study Web portal, and primary inclusion criteria were (1) aged 18 to 65 years; (2) a primary diagnosis of AD or ED; (3) no substance abuse or dependence in the past 6 months; (4) no current or past psychosis or bipolar disorder; (5) no suicidal ideation; (6) if on medication with a monoamine agonist, this had been stable in the past month; and (7) no ongoing psychological treatment. Inclusion criteria were assessed by a licensed psychologist using information collected in a telephone-conducted clinical interview comprising the Mini-International Neuropsychiatric Interview [26] and data obtained from a Web-based screening. AD and ED were assessed with a clinical interview developed specifically for the trial that followed diagnostic criteria for respective diagnosis. The included sample consisted of 85% (85/100) women with a mean age of 47 (SD 8.8) years. A majority of patients (73%, 73/100) had a university education of 3 or more years, and 71% (71/100) of patients had a full-time employment at inclusion. Participants had, on average, suffered from their stress-related symptoms for 1.6 (SD 1.3) years. Patients with AD and ED were equally represented in the sample (53% [53/100] and 47% [47/100], respectively).

Interventions

The CBT protocol used in the trial, on which this health economic study is based [13], has previously been tested in 2 randomized trials in the form of face-to-face therapy [27] and guided self-help [28] with promising results. A central component of the treatment was recovery training, which has previously been found to be important both in prevention and rehabilitation of chronic stress [29]. The treatment targeted increased recovery through scheduled recuperating activities and psychoeducation about sleep hygiene and stimulus control for improved sleep. Other central aspects of the treatment were behavioral activation [30] and exposure [31] to help patients break fear-avoidance patterns related to, for example, perfectionism, excessive worry, and assertiveness. Patients were encouraged to plan their days to create a balance between effortful expenditure and recovery.

The 12-week ICBT was delivered via a secure (encrypted traffic and 2-factor authentication) Web platform where psychoeducation, worksheets, and exercises were presented primarily through text. Patients received weekly written feedback on their work and progress by a licensed clinical psychologist who gave gradual access to the 12 treatment modules, similar to chapters of a book. Patients could contact their psychologist at any time and expect a reply within 48 business hours. Weekly automated text message reminders were sent to all patients in the ICBT group to log on to the treatment platform and work with the treatment. The average number of completed modules in ICBT was 9.2 (SD 3.2) out of 12, which is typically considered adequate and on par with many other ICBT protocols for common mental health conditions [32]. The mean therapist time per patient was 87 (SD 36) min, that is, on average, a little over 7 min weekly per patient.

Patients in the WLC group did not receive any treatment during the 12-week experimental phase and were encouraged not to make any changes in psychotropic medication or seek other psychological treatment for their stress-related disorder during that time. After 12 weeks, they were immediately crossed over to treatment.

Clinical Assessment

For the cost-effectiveness analyses, effectiveness was defined as remission rate, estimated based on the 14-item Perceived Stress Scale (PSS) [33]. The PSS (which was administered online) measures how often one has perceived life as unpredictable, uncontrollable, and overloading in the past month. Patients respond on a 5-point response scale from 0 (never) to 4 (very often), with a total scale range of 0 to 56 (a higher score indicating higher level of perceived stress). The PSS has good construct validity and sensitivity to change in samples with stress-related disorders [34].

For the cost-utility analyses, the EuroQol 3L Questionnaire (EQ-5D) was used to assess health-related quality of life [35]. The EQ-5D is a nondisease-specific self-assessment questionnaire that measures 5 health domains: mobility, self-care, usual activities, pain and/or discomfort, and anxiety and/or depression [36]. An advantage of the EQ-5D is that an overall utility score of quality of life can be obtained, which can facilitate comparisons with trials using other health states and other disease conditions [37].

Economic Evaluation

The economic evaluation framework of this study was a within-trial (12-weeks) cost-effectiveness and cost-utility analyses. Cost-effectiveness concerns the association between costs and efficacy (ie, remission rates on PSS), and cost-utility concerns the association between costs and health-related quality of life. Analyses were conducted from a societal perspective, including all direct and indirect medical costs (eg, health care consumption, costs for medicine, supplements, and alternative care/support) as well as nonmedical costs (eg, work cutback, domestic costs, and sick leave). We also conducted separate cost-effectiveness and cost-utility analyses from a health care perspective, based on direct medical costs only.

Measurement of Resource Utilization

To enable the estimation of costs, resource utilization data were obtained using the self-rated Trimbos and Institute Medical Technology Assessment of Costs Questionnaire for Psychiatry (TIC-P) [38]. This instrument measures resource use over the past month (except for medication use, which concerns the past 2 weeks). The TIC-P is used to estimate (1) direct medical costs (ie, health care utilization and medication costs), (2) direct nonmedical costs (eg, use of alternative medicine and medicaments), and (3) indirect nonmedical costs (costs associated with production loss generated by, for example, sick leave or work cutback due to presenteeism). The TIC-P has previously been shown to be a feasible and reliable instrument for collecting data on medical consumption and productivity losses [39]. All costs were initially assessed in the local currency (Swedish Krona or SEK) and converted into US Dollars ($) using 2016 as the reference year (yielding 1 SEK equivalent to US $0.1148). Scoring of the TIC-P and transformation of TIC-P scores to costs was done by the first author who, during this process, was blind to patient condition (ICBT vs WLC).

Estimation of Costs

We used public health care tariffs to estimate medical cost (see Multimedia Appendix 1). For each patient, tariffs were multiplied by the corresponding number of health service units (eg, consultations and sessions) from the TIC-P. Costs for medications (prescription and otherwise) were estimated using market prices in Sweden. As the TIC-P only asks for medication use in the past 2 weeks, these costs were multiplied by 2 to represent a time frame of 1 month. The human capital approach was used to estimate indirect costs of sick leave, unemployment, and work cutback, meaning that lost gross earnings were taken as a measure of production loss [37]. Salaries were estimated based on the average monthly earnings in Sweden, stratified by the level of education and gender, as listed by Statistics Sweden for the year 2015. To estimate costs for work cutback, participants reported the number of days worked when ill in the past month and an inefficiency score of how much productivity was reduced as compared with productivity when in good health. Reduced monthly productivity (based on days worked when ill) was multiplied with the percentage of production loss (inefficiency score) and with the estimated daily earning [40]. Costs of domestic work cutback and informal care from family and friends was estimated to US $10 per hour based on data from the study by Smit et al [41], an estimate that has been used in several recently published health economic trials [42,43]. Owing to the short period under study (12 weeks), we did not discount costs.

Direct medical costs associated with the intervention mainly comprised the clinician’s time for the 12 weeks of ICBT. Clinician time spent was logged for every contact with a patient and included reviewing the patient’s work and giving written feedback. As clinicians in this study were exclusively licensed psychologists working at a primary health care clinic, we used the cost tariff for a 45-min session with a psychologist within primary health care (US $103) multiplied by the clinician time for each individual patient.

Statistical Analysis

All analyses adhered to the intention-to-treat principle, meaning that all randomized participants were eligible for analysis. Owing to the low degree of data loss (1 out of 50 in ICBT and 2 out of 50 in the WLC post treatment), no imputation of missing data was deemed necessary (we employed listwise deletion). Analyses were conducted in SPSS 20.0 (IBM) and Stata/IC 14.2 (Stata Corporation).

Analysis of Remission Rates

Remission was operationalized as clinically significant improvement on the primary outcome PSS. This meant that patients were required to (1) make a reliable change in accordance with the Jacobson and Truax [44] criteria (ie, a reduction of 7 units on the PSS) and (2) have a posttreatment rating on PSS that was closer to a normal population than to a clinical population. On the basis of the data from the study by Lavoie and Douglas [45], the postrating cutoff was defined as a PSS score <31. We analyzed differences in remission rates pre- to posttreatment using chi-square tests.

Analysis of Health-Related Quality of Life

The answers given in EQ-5D were combined to generate a utility score of health states ranging from 0 to 1, with 0 representing death and 1 representing full health [36], based on the Swedish experience-based time trade-off value set [46]. Owing to the relatively short time perspective (12 weeks), we used the difference in utility scores between the pre- and posttreatment assessment to measure the effect in terms of health-related quality of life in the different cost-utility analyses. To generate a frame of reference, we also conducted a separate cost-utility analysis from a societal perspective in which the utility scores were converted to quality-adjusted life years (QALYs) gained over the 12-week treatment phase. As utility scores demonstrated skewness and kurtosis of residuals, differences between the ICBT and the WLC were analyzed using bootstrapped linear mixed models (5000 replications) where time (pre-to posttreatment), group (ICBT vs WLC), and the time × group interaction were independent variables.

Analysis of Cost Changes

As in the case of utility scores, cost residuals were skewed and showed evidence of high kurtosis. We, therefore, analyzed between-group changes in costs (gross and net total costs as well as changes in specific cost-domains) from pre- to posttreatment using bootstrapped linear mixed-models (5000 replications).

Analysis of Cost-Effectiveness and Cost-Utility

To estimate cost-effectiveness and cost-utility, we used the incremental cost-effectiveness ratio (ICER) and the incremental cost-utility ratio (ICUR), respectively. These were calculated by taking the net cost difference between conditions (ICBT vs WLC) at posttreatment compared with baseline, divided by the difference in remission rate (ICER) or utility improvement (ICUR) between the groups over the 12-week treatment period [14]. In the analysis based on QALYs, the denominator of the ICUR was the ICBT versus WLC difference in per capita QALYs over the 12-week treatment period.

Assessment of Uncertainty

As ICER and ICUR point estimates are difficult to interpret [14], we modeled the uncertainty of these based on bootstrapping (5000 samples) of the treatment group's incremental costs and effects (remission rate and utility) as compared with the WLC. The bootstrapped values then formed the basis for cost-effectiveness and cost-utility planes, with effects on the x-axis and costs on the y-axis. Cost-effectiveness and cost-utility planes allow a probabilistic decision-making approach [47], where a majority of ICERs or ICURs in the southeast quadrant of these planes indicate a larger effect of ICBT at a lower cost compared with the WLC, and a majority of ICERs or ICURs in the northeast quadrant suggest a larger effect of ICBT at a higher cost compared with the WLC. In the latter case, the amount of money a decision maker is willing to pay for, for example, 1 additional patient in remission, is crucial in determining whether a new treatment is to be adopted or not. Hence, we investigated the probability of the treatment being cost-effective compared with the WLC at a range of different willingness to pay (WTP) scenarios. WTP for an additional responder can be illustrated by the means of cost-effectiveness and cost-utility acceptability curves [48].

Additional sensitivity analyses were conducted to further confirm the robustness of results. First, we replaced the intervention cost for each patient in the ICBT group with US $663 (ie, more than thrice the mean intervention cost), based on an estimate of the cost of running ICBT in a psychiatric clinic [49]. The new intervention cost included an average therapist time for psychiatric patients in ICBT and costs for assessment procedures, referral and follow-up visits, other health care staff, hospital space, and platform administration [49]. Second, as previous investigations have shown that patients with AD and ED differ in symptom severity and functional disability [13], we calculated ICERs separately for the diagnostic groups and also constructed cost-effectiveness planes and acceptability curves from a societal perspective. Third, as costs for domestic work and informal care from family and friends are difficult to estimate [49], we calculated ICERs and modeled cost-effectiveness planes using 2 different scenarios: (1) a scenario in which we used an updated cost-estimate of US $19 (ie, almost doubling the cost-estimate used in the main analyses), representing the average gross hourly wage earned by a domestic worker as suggested by Bock et al [49], and (2) another scenario in which we removed the costs for domestic work and informal care altogether. Finally, although there are indications that production loss because of presenteeism is a major cost driver for a range of disease conditions, comprising on average 52% of total costs [50], there is to date no consensus regarding methods to accurately measure and value it. We, therefore, conducted a sensitivity analysis with the purpose of illuminating to what extent cost related to work cutback may impact the total cost-effectiveness of ICBT in our trial. We calculated ICERs and modeled cost-effectiveness planes from a societal perspective, both when excluding work cutback as a cost-domain and when doubling these costs.

Results

Data Completion and Receiving Treatment Outside of the Study

There was a 100% (100/100) data completion at pretreatment. At posttreatment assessment, data completion was 98% (49/50) in the ICBT group and 96% (48/50) in the WLC. At the 6MFU, 94% (47/50) of patients in the ICBT group completed assessments. A total of 3 patients (6%) in ICBT and 4 (8%) in the WLC reported having received other treatment for stress-related problems during the intervention period. Fisher exact test revealed no significant difference between the groups in this regard (*P*=.47). In the ICBT group, 2 patients received psychological treatment and 1 initiated psychotropic medication. In the WLC, 2 patients received psychological treatment and 2 initiated psychotropic medication.

Clinical Efficacy: Remission Rates and Utility

Table 1 shows means and SDs at pretreatment, posttreatment, and 6MFU for PSS and utility scores. In this study, remission rate on PSS was the main clinical outcome. As previously reported in the main outcome study [13], 31 of 50 patients (62%)

in the ICBT group were in remission (ie, met criteria for clinically significant improvement) post treatment, as compared with 5 of 50 patients (10%) in the WLC group, which was a statistically significant difference (X^2_1=29.3; *P*<.001). At the 6MFU, 34 of 50 patients (68%) in the ICBT group were in remission. Regarding utility, patients in the ICBT group significantly increased their health-related quality of life from pre- to posttreatment (beta=.05; 95% CI 0.03 to 0.08; *Z*=3.91; *P*<.001), whereas patients in the WLC group did not (beta=.02; 95% CI −0.01 to 0.05; *Z*=1.41; *P*=.16). However, there was no significant group × time interaction effect on utility from pre- to posttreatment (beta=.03; 95% CI −0.01 to 0.07; *Z*=1.68; *P*=.09). There was no significant change from posttreatment to 6MFU in utility in the ICBT group, indicating stability of improvements (beta=.01; 95% CI −0.02 to 0.04; *Z*=0.51; *P*=.61).

Table 1. Statistical values for primary outcome measure and health-related quality of life.

Measure and group	Pretreatment, mean (SD)	Posttreatment, mean (SD)	6-month follow-up, mean (SD)
Perceived Stress Scale			
ICBT[a]	37.2 (7.1)	24.2 (8.6)	21.9 (7.7)
WLC[b]	36.4 (7.3)	33.2 (7.9)	—[c]
Utility[d]			
ICBT	0.82 (0.12)	0.87 (0.09)	0.89 (0.09)
WLC	0.82 (0.11)	0.84 (0.11)	—

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

[b] WLC: waitlist control.

[c] Not applicable. Patients in the WLC group were crossed over to treatment after the posttreatment assessment.

[d] Utility scores are based on EuroQol 3L Questionnaire health states.

Cost Changes

Table 2 presents the per capita costs at the pre- and posttreatment assessment (for spatial reasons, costs at the 6MFU are presented in the Multimedia Appendix 1). There was no significant difference between groups from pre- to posttreatment regarding gross total costs (ie, societal costs excluding the intervention costs; beta=−260.22; 95% CI −877.91 to 357.47; *Z*=−0.83; *P*=.41) or net total costs (ie, societal costs including the intervention costs; beta=−57.61; 95% CI −685.67 to 570.45; *Z*=−0.18; *P*=.86). Analyses of specific cost-domains revealed a significant interaction effect between time (from pre- to posttreatment) and group (ICBT vs WLC) regarding costs related to work cutback, where patients in the ICBT group significantly reduced their costs compared with those in the WLC group (beta=−292.24; 95% CI −525.98 to −58.50; *Z*=−2.45; *P*=.014). No other between-group interaction effects were found in specific cost-domains based on the TIC-P. Gross total costs in the ICBT group remained stable between posttreatment and the 6MFU (beta=173.76; 95% CI −383.98 to −731.46; *Z*=0.62; *P*=.54).

Table 2. Per capita costs at pre- and posttreatment assessment (all costs are in US dollar, converted from the Swedish Krona).

Cost-domains	Pretreatment[a]				Posttreatment			
	ICBT[b], (n=50)		WLC[c], (n=50)		ICBT (n=49)		WLC (n=48)	
	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median
Direct medical	187 (320)	6	316 (547)	10	172 (255)	41	350 (845)	213
Health care visits	169 (288)	0	312 (546)	0	166 (254)	0	345 (845)	211
Medication	18 (83)	2	4 (5)	1	6 (9)	2	5 (7)	3
Direct nonmedical	80 (120)	20	160 (445)	35	78 (130)	0	109 (173)	5
Indirect nonmedical	1115 (1436)	579	1043 (1629)	132	884 (1348)	313	1040 (1368)	515
Unemployment	150 (746)	0	80 (568)	0	182 (910)	0	0 (0)	0
Sick leave	389 (980)	0	639 (1406)	0	438 (929)	0	624 (1276)	0
Work cutback[d]	408 (636)	149	221 (507)	0	212 (421)	0	317 (463)	178
Domestic	168 (397)	56	103 (189)	26	53 (78)	13	99 (155)	31
Gross total costs	1382 (1533)	843	1518 (2080)	466	1134 (1528)	492	1499 (1841)	762
Intervention costs	—[e]	—	—	—	203 (78)	205	—	—
Net total costs	1382 (1533)	843	1518 (2080)	466	1314 (1553)	721	1499 (1841)	762

[a] There were no significant differences between groups (ICBT vs WLC) in any cost-domain at the pretreatment assessment (*P*=.06-.57).

[b] ICBT: internet-based cognitive behavioral therapy.

[c] WLC: waitlist control.

[d] Significant interaction effect between groups (ICBT vs WLC) from pre- to posttreatment (*P*=.014).

[e] Not applicable.

Societal Perspective: Cost-Effectiveness and Cost-Utility

Taking all costs into account (ie, conducting the analysis from a societal perspective), the ICER was −77.24/0.49=−157.63, favoring ICBT over the WLC. This indicates that the most likely scenario was that ICBT led to higher remission rates compared with the WLC and was associated with slightly larger reductions in costs from pre- to posttreatment. Figure 2 (top left) presents the scatter plot of simulated ICERs across the 4 quadrants of the cost-effectiveness plane, illustrating the degree of uncertainty associated with the estimated parameter. A majority of the simulated ICERs (60.86%, 3043/5000) are located within the southeast quadrant (larger effect of ICBT at a lower cost), whereas 39.14% (1957/5000) are located in the northeast quadrant (larger effect of ICBT at a higher cost). This suggests that ICBT is a cost-effective treatment from a societal perspective.

To estimate the cost-utility of ICBT relative to the WLC, we repeated the analysis using health-related quality of life (change in utility) as outcome. The ICUR was −77.24/0.03=−2574.67. As illustrated in Figure 2 (bottom left), a majority (57.20%, 2860/5000) of the simulated ICURs are located within the southeast quadrant, indicating that the most likely outcome is that ICBT, in comparison with no treatment, leads to lower net costs while increasing health-related quality of life. When QALYs were instead used as outcome, the ICUR was −77.24/0.0043=−17962.79. Of the simulated ICURs, 48.14% (2407/5000) fell in the southeast quadrant of the cost-utility plane, 34,52% (1726/5000) fell in the northeast quadrant, 11.90% (595/5000) fell in the northwest quadrant, and 5.44% (272/5000) in the southwest quadrant (see Multimedia Appendix 1 for cost-utility plane). This suggests that the most likely scenario is that the treatment generates more QALYs at lower net costs compared with no treatment.

Figure 3 illustrates the acceptability curves based on the same data as used above in the cost-effectiveness and cost-utility analyses (using change in utility) from a societal perspective. As can be seen, ICBT has a 60% probability of being cost-effective from pre- to posttreatment if society is willing to pay US $0 for 1 additional case of remission or for increased utility. This was true also when using QALYs as outcome (see Multimedia Appendix 1). If society instead were willing to pay US $1000 for 1 additional case of remission, the probability of the treatment being cost-effective would increase to 96% (Figure 3, top). Assuming a WTP of US $8000, the probability of ICBT being cost-effective with regard to increasing quality of life would increase to 80% (Figure 3, bottom). Using QALYs as outcome, a WTP of US $25,000 and US $50,000 would increase the probability of 1 QALY gain in ICBT, compared with the WLC, to 71% and 79%, respectively.

Figure 2. Cost-effectiveness (top) and cost-utility (bottom) planes comprising 5000 bootstrapped incremental cost-effectiveness ratios and incremental cost-utility ratios, respectively, comparing internet-based cognitive behavioral therapy with waitlist control over the 12-week treatment period. Remission was operationalized as the proportion of patients who made a clinically significant improvement on the 14-item Perceived Stress Scale (PSS). Utility was based on EuroQol Questionnaire, 3L version (EQ-5D 3L) health states. In the 2 left-hand planes, all costs from a societal perspective are included. The 2 right-hand planes include costs from a health care perspective (ie, using only direct medical costs). Please note that the y-scale of the health care perspective graphs differs from that of the societal perspective graphs.

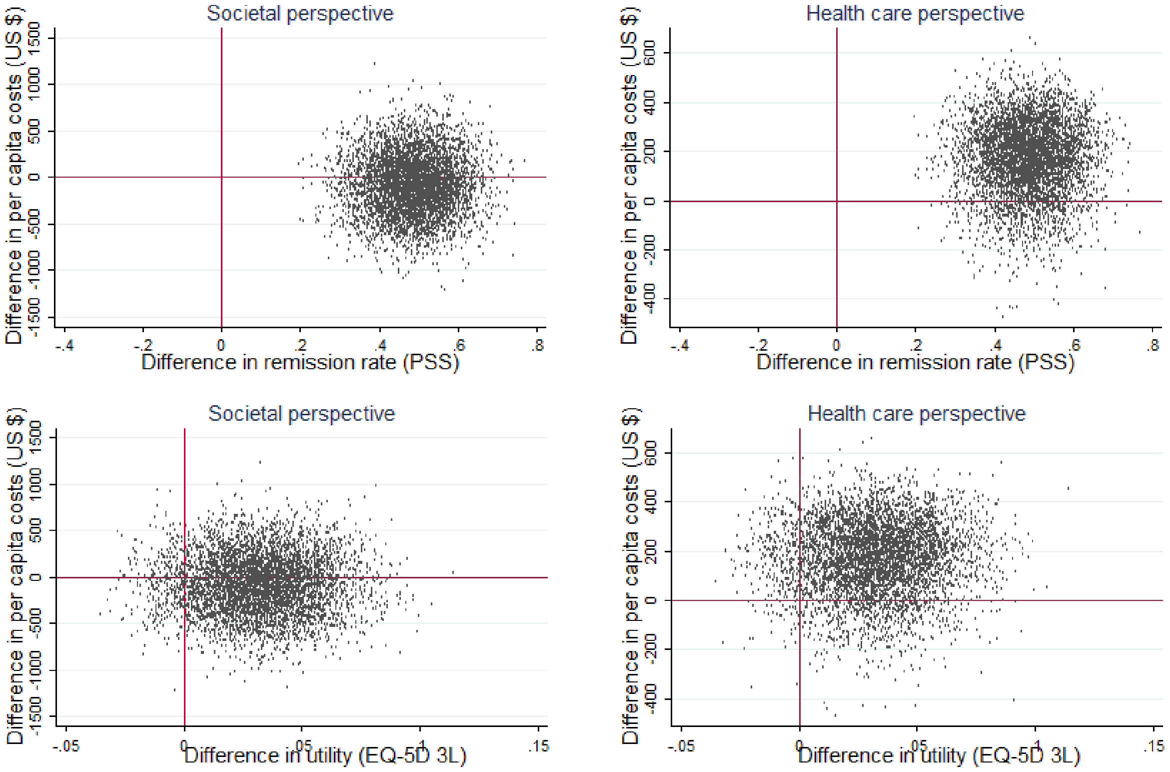
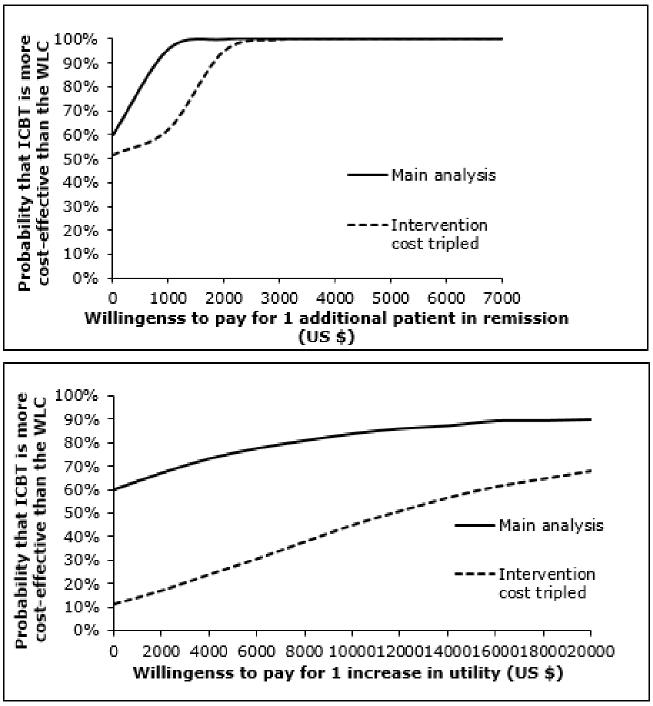


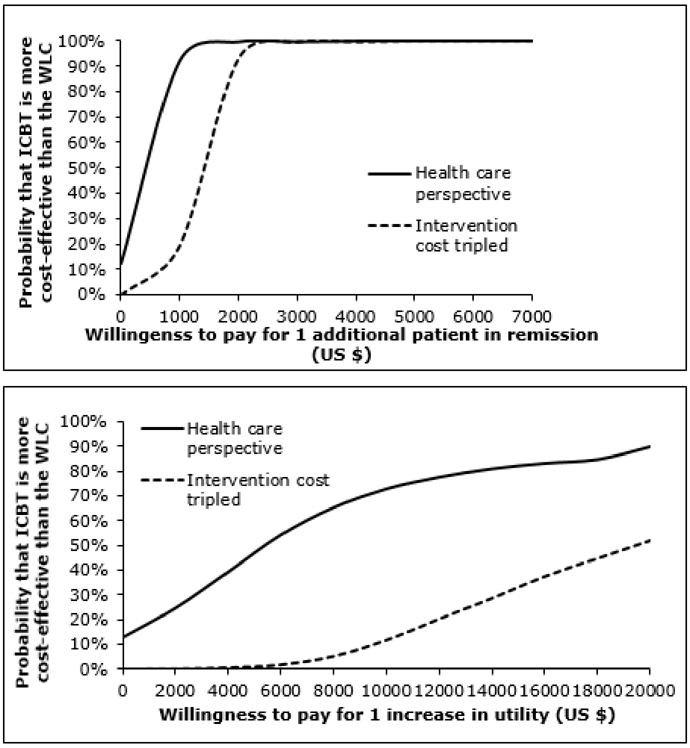
Figure 3. Cost-effectiveness (above) and cost-utility (below) acceptability curves from a societal perspective comparing internet-based cognitive behavioral therapy (ICBT) with waitlist control (WLC) over the 12-week treatment period. Note: Solid lines represent the probability of ICBT being more cost-effective in the standard scenario. Dotted lines represent the probability of ICBT being more cost-effective based on a sensitivity analysis in which a 3-fold higher intervention cost was used (US $663).



Health Care Perspective: Cost-Effectiveness and Cost-Utility

Taking only the direct medical costs into account, that is, a health care perspective, analysis using remission rate as efficacy measure generated an estimated ICER of 171.12/0.49=349.00. This means that each additional case of remission in ICBT relative to the WLC was associated with a slight increase in health care costs in the ICBT compared with the WLC. When conducting the same analysis using change in utility as outcome, the corresponding ICUR was 171.12/0.03=5704. Figure 2 shows the scatter plots of the simulated ICERs (top right) and ICURs (bottom right) from a health care perspective. Although treatment benefits of ICBT are associated with a cost for the health care provider, the cost-effectiveness acceptability curve (Figure 4, top) shows that, at a relatively low WTP of US $1000, the probability of ICBT being cost-effective from a health care perspective is 97%. Regarding cost-utility from a health care perspective, the ICBT has a higher probability of cost-utility compared with the WLC at a WTP of approximately US $6000 (Figure 4, bottom).

Figure 4. Cost-effectiveness (above) and cost-utility (below) acceptability curves from a health care perspective comparing internet-based cognitive behavioral therapy (ICBT) with waitlist control (WLC) over the 12-week treatment period. Note: Solid lines represent the probability of ICBT being more cost-effective in the standard scenario. Dotted lines represent the probability of ICBT being more cost-effective based on a sensitivity analysis in which a 3-fold higher intervention cost was used (US $663).

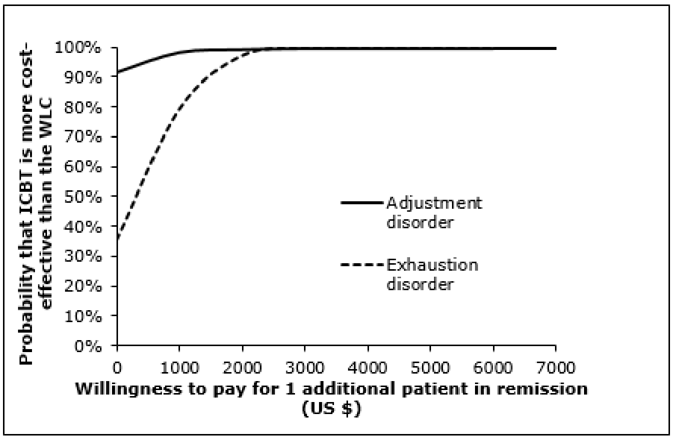


Sensitivity Analyses

In the sensitivity analysis in which an intervention cost of US $663 (including, eg, assessment procedure, platform administration, and other health care staff and overhead costs) replaced the intervention cost of ICBT in this study, ICBT remained more likely to be cost-effective from a societal perspective compared with the WLC even at a WTP of US $0 (see Figure 3, top, dotted line). The cost-utility acceptability curve (Figure 3, bottom, dotted line) indicates that a WTP of approximately US $17,000 would be required for the ICBT to be preferable to the WLC in terms of cost-utility. From a health care perspective (see Figure 4, dotted lines), the new intervention cost estimate would require a WTP of approximately US $1500 for ICBT to be more likely than the WLC to be cost-effective from pre- to posttreatment. To render cost-utility of ICBT, compared with the WLC, from a health care perspective, a WTP of US $20,000 would be needed.

Figure 5 illustrates the cost-effectiveness acceptability curves for the respective diagnostic groups (AD and ED), using remission rate on PSS as efficacy measure. For patients diagnosed with AD, there was a 92% probability that ICBT was more cost-effective than the WLC at a WTP of US $0. For patients diagnosed with ED, there was an 80% probability that ICBT was more cost-effective compared with the WLC at a WTP of US $1000. More information can be found in Multimedia Appendix 1 about outcomes, costs, and cost-effectiveness planes for AD and ED respectively (Multimedia Appendix 1).

Figure 5. Cost-effectiveness acceptability curve from a societal perspective, comparing patients with adjustment disorder and exhaustion disorder in internet-based cognitive behavioral therapy (ICBT) to their respective wait-list control (WLC) over the 12-week treatment period. Remission is operationalized as clinically significant improvement on the Perceived Stress Scale (PSS).



The sensitivity analyses investigating different cost scenarios for informal care and domestic work indicated that the overall cost-effectiveness of ICBT remained unchanged when a higher cost-estimate was used and when costs for informal care and domestic work were removed from the total net costs. In the first scenario (where US $19 was used as a cost estimate of the average gross hourly wage earned by a domestic worker), the ICER was −126.69/0.49=−258.55 in favor of the ICBT, with 64.96% (3248/5000) of simulated ICERs falling in the southeast quadrant of the cost-effectiveness plane and 35.04% (1752/5000) falling in the northeast quadrant. In the second scenario (where costs for informal care and domestic work were removed), the ICER was −37.91/0.49=−1.99, supporting the superiority of ICBT with 55.10% (2755/5000) of simulated ICERs in the southeast quadrant and 44.90% (2245/5000) in the northeast quadrant (see Multimedia Appendix 1 for cost-effectiveness planes). Regarding presenteeism costs, the ICER in which costs for work cutback were removed indicated that each patient in remission was associated with a higher cost in the ICBT compared with the WLC (ICER: 214.43/0.49=437.61), with only 27.02% (1351/5000) of simulated ICERs in the southeast quadrant and 72.98% (3649/5000) in the northeast quadrant. At a WTP of US $2000, however, the probability of ICBT being cost-effective increased to 98% (see Multimedia Appendix 1). When costs for presenteeism were doubled, the resulting ICER was −368.91/0.49=−752.8 in favor of the ICBT, with 86.92% (4346/5000) of simulated ICERs in the southeast quadrant and 13.01% (654/5000) in the northeast quadrant, indicating that this scenario would render the treatment highly cost-effective as compared with the WLC. This sensitivity analysis supports that the net difference in these costs is largely in favor of the ICBT and that productivity loss because of presenteeism to a large extent contributes to the high cost-effectiveness of the treatment.

Discussion

Principal Findings

To our knowledge, this is the first health economic evaluation conducted of an ICBT specifically targeting stress-related disorders in the form of AD and ED. The results showed that ICBT generated higher remission rates compared with the WLC at no additional societal costs. The cost-utility analysis also showed that the ICBT was more likely to be preferable to the WLC from a societal perspective, both when change in utilities and QALYs were used as outcome. Taking a health care perspective (ie, including only direct medical costs in the ICER), there was a 97% probability that ICBT would generate an additional patient in remission compared with the WLC at a WTP of US $1000. A WTP of approximately US $6000 would, however, be required for the ICBT to be preferable to the WLC in terms of cost-utility. The overall indication of treatment cost-effectiveness was supported by sensitivity analyses. In sum, using data from a randomized controlled trial, this study showed that ICBT for stress-related disorders is likely to be a cost-effective treatment compared with no treatment.

The results of our study are in line with other health economic evaluations of ICBT for a range of clinical conditions, where ICBT in general has more than 50% probability of being cost-effective compared with no treatment or conventional CBT at a WTP of US $0 for an additional improvement [19]. As no previous health economic evaluations of ICBT for patients diagnosed with stress-related disorders have been conducted, to our knowledge, there are no available cost-effectiveness estimates of the target population with which we can directly compare our results. Nevertheless, 2 recently published studies evaluated cost-effectiveness of ICBT for employees with elevated levels of stress, 1 from an employer's perspective (including only costs relevant for the employer, ie, presenteeism, absenteeism, and treatment costs) [24] and 1 adopting a full societal perspective [25]. Results of these studies were similar to those found in this study, with probabilities of treatment cost-effectiveness of approximately 70% compared with WLC at a WTP of US $0 [24,25]. In these studies, as in ours, costs because of presenteeism were large and constituted central cost drivers in the cost-effectiveness analyses. These clinical findings support previous reports stating that presenteeism constitutes a large cost in stress-related mental illness and many other clinical conditions [1,50], in fact reducing output by at least as much as absenteeism [50,51]. The indication that ICBT may significantly reduce costs within this domain is encouraging and needs to be further investigated. Although the potential cost savings related to reduced work cutback after ICBT may not directly benefit the health care provider (the payer of the intervention in this study), the low intervention cost and high scalability of ICBT makes the treatment cost-effective also from a health care perspective.

Although this study showed that ICBT for stress-related disorders is likely to be cost-effective, the degree of cost-effectiveness and cost-utility versus the WLC was, to some extent, moderated by 3 factors. First, the ICER was slightly more favorable when using a societal perspective compared with a narrower health care perspective. This is reflective of the absence of effect of ICBT on direct medical costs (health care visits and medication), which may be explained by the fact that patients suffering from chronic stress often have comorbid health conditions [52] that ICBT would be unlikely to affect in the short term. Second, the ICER (based on remission rates) was somewhat more favorable than the ICUR (based on utility change or QALYs). This was because the controlled effect on the primary outcome PSS was substantially larger than that on EQ-5D. This is expected given the generic nature of EQ-5D, which includes items such as "I am unable to wash and dress myself. " Third, even if ICBT was cost-effective for both diagnostic groups, it was more so for patients with AD compared with those with ED. These results can be understood by the fact that, relative to their respective controls, both diagnostic groups made large symptomatic improvements, but cost changes were smaller for patients with ED compared with patients with AD. A possible conclusion that can be drawn is that it is likely beneficial to offer treatment early on in the development of chronic stress (ie, for patients suffering from AD as opposed to ED), as a means of optimizing treatment cost-effectiveness and preventing increased societal costs. Nonetheless, it is encouraging that ICBT is cost-effective even in the very short term also for patients suffering from severe symptoms of chronic stress.

Limitations and Strengths

There are some limitations to this study. First, the use of a WLC that (for ethical reasons) received treatment immediately after the 12-week experimental phase did not allow for between-group comparisons of costs and effects at the 6MFU. Although costs seem to remain stable between posttreatment and 6MFU in the ICBT condition, we do not know whether this is because of treatment or other factors. The reason for using a WLC in the randomized controlled trial was that we judged the intervention research on stress-related disorders to still be in its early stages. In this context, the use of WLC has been suggested to be a viable option because it gives protection against basic threats to internal validity (eg, regression toward the mean and spontaneous remission) and reduces the risk of type-II error, which are central factors in the early phase of treatment development [53]. Furthermore, the WLC represents a realistic scenario, given that no established treatment guidelines exist for AD or ED and many suffering individuals never get access to psychological treatment [54].

A second limitation was that we relied on self-report data to obtain information about costs, which are prone to certain biases (eg, memory bias). Using self-report data to estimate costs in economic evaluations conducted together with randomized trials is, however, the most common procedure [55], and the TIC-P has demonstrated reasonable validity and reliability [39]. Moreover, there are indications of high convergence between self-report of health care visits and registry data [56].

A third issue is that of generalizability. In this study, the fact that the treatment was delivered via the internet likely improves generalizability, as it makes the results less dependent on factors such as regional differences in clinical practice and resource use (ie, patients were not tied to any specific setting). However, the high percentage of educated women in our sample limits generalizability to the full spectrum of individuals suffering from stress-related disorders. Even though the effect of gender on outcome in CBT has been found to be limited [57], and women with a moderate to high level of education tend to be overrepresented among people who seek health care for stress-related disorders (refer to, eg, the studies by Salomonsson et al [27], Heber et al [58], and Netterstrom et al [59]), more research on the cost-effectiveness of ICBT is needed using different types of patient samples and treatment settings.

Central strengths of this study were the randomized design, which enabled control over confounders; low attrition rates; and the use of both societal and health care perspectives in the health economic evaluation. Another central strength was that we used a clinical sample for which no past health economic evaluations have been conducted. We characterized heterogeneity through sensitivity analysis of the diagnostic subgroups AD and ED, which is often overlooked in health economic trials [60].

Conclusions

In conclusion, ICBT for the stress-related disorders AD and ED is likely to be a cost-effective treatment compared with WLC. Most probably, ICBT leads to substantial improvements at no net societal costs, meaning that the treatment costs associated with administrating ICBT are likely to be quickly returned to society. In this trial, nearly two-thirds of patients who received ICBT were in remission after 12 weeks. Considering the scalability of this treatment and the low net costs, we believe that if ICBT is implemented in routine health care, it could play an important role in making effective treatment accessible to patients suffering from chronic stress in the form of AD and ED.

Acknowledgments

This study was funded by the Doctoral School in Health Care Sciences, Karolinska Institutet, Stockholm Stress Center, and Stockholm County Council. These are public institutions, neither of which had any role in the design, execution, or publication of the study.

Authors' Contributions

All authors made significant contributions to the planning, recruitment, execution, analysis, and publication of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary material presenting estimation of costs, cost-utility analysis based on Quality-Adjusted Life Years, and additional results from sensitivity analyses.

[DOCX File , 402 KB - jmir_v21i9e14675_app1.docx]

References

1. Hassard J, Teoh K, Cox T, Dewe P, Cosmar M, Gründler R, et al. European Agency for Safety and Health at Work. 2014. Calculating the Cost of Work-Related Stress And Psychosocial Risks: European Risk Observatory Literature Review URL: https://osha.europa.eu/en/publications/literature_reviews/calculating-the-cost-of-work-related-stress-and-psychosocial-risks/view [accessed 2018-02-01]
2. McEwen BS. Physiology and neurobiology of stress and adaptation: central role of the brain. Physiol Rev 2007 Jul;87(3):873-904 [FREE Full text] [doi: 10.1152/physrev.00041.2006] [Medline: 17615391]
3. Grossi G, Perski A, Osika W, Savic I. Stress-related exhaustion disorder--clinical manifestation of burnout? A review of assessment methods, sleep impairments, cognitive disturbances, and neuro-biological and physiological changes in clinical burnout. Scand J Psychol 2015 Dec;56(6):626-636. [doi: 10.1111/sjop.12251] [Medline: 26496458]
4. Layard R. Mental health: the new frontier for labour economics. IZA J Labor Policy 2013;2:2 [FREE Full text] [doi: 10.1186/2193-9004-2-2]
5. Kalia M. Assessing the economic impact of stress--the modern day hidden epidemic. Metabolism 2002 Jun;51(6 Suppl 1):49-53. [doi: 10.1053/meta.2002.33193] [Medline: 12040542]
6. Arends I, Bruinvels DJ, Rebergen DS, Nieuwenhuijsen K, Madan I, Neumeyer-Gromen A, et al. Interventions to facilitate return to work in adults with adjustment disorders. Cochrane Database Syst Rev 2012 Dec 12;12(12):CD006389. [doi: 10.1002/14651858.CD006389.pub2] [Medline: 23235630]
7. Reed GM, Correia JM, Esparza P, Saxena S, Maj M. The WPA-WHO global survey of psychiatrists' attitudes towards mental disorders classification. World Psychiatry 2011 Jun;10(2):118-131 [FREE Full text] [doi: 10.1002/j.2051-5545.2011.tb00034.x] [Medline: 21633689]
8. Evans SC, Reed GM, Roberts MC, Esparza P, Watts AD, Correia JM, et al. Psychologists' perspectives on the diagnostic classification of mental disorders: results from the WHO-IUPsyS global survey. Int J Psychol 2013;48(3):177-193 [FREE Full text] [doi: 10.1080/00207594.2013.804189] [Medline: 23750927]
9. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. Fifth Edition. Arlington, VA: American Psychiatric Publishing; 2013.
10. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Description and Diagnostic Guidelines. Geneva, Switzerland: World Health Organization; 1992.
11. Schaufeli W, Enzmann D. The Burnout Companion To Study And Practice: A Critical Analysis. Boca Raton, Florida: CRC Press; 1998.
12. Fava GA, Cosci F, Sonino N. Current psychosomatic practice. Psychother Psychosom 2017;86(1):13-30 [FREE Full text] [doi: 10.1159/000448856] [Medline: 27884006]
13. Lindsäter E, Axelsson E, Salomonsson S, Santoft F, Ejeby K, Ljótsson B, et al. Internet-based cognitive behavioral therapy for chronic stress: a randomized controlled trial. Psychother Psychosom 2018;87(5):296-305. [doi: 10.1159/000490742] [Medline: 30041167]
14. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press; 2015.
15. Saha S, Hoerger TJ, Pignone MP, Teutsch SM, Helfand M, Mandelblatt JS, Cost Work Group, Third US Preventive Services Task Force. The art and science of incorporating cost effectiveness into evidence-based recommendations for clinical preventive services. Am J Prev Med 2001 Apr;20(3 Suppl):36-43. [doi: 10.1016/S0749-3797(01)00260-4] [Medline: 11306230]
16. Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
17. Bhui KS, Dinos S, Stansfeld SA, White PD. A synthesis of the evidence for managing stress at work: a review of the reviews reporting on anxiety, depression, and absenteeism. J Environ Public Health 2012;2012:515874 [FREE Full text] [doi: 10.1155/2012/515874] [Medline: 22496705]

18. Richardson KM, Rothstein HR. Effects of occupational stress management intervention programs: a meta-analysis. J Occup Health Psychol 2008 Jan;13(1):69-93. [doi: 10.1037/1076-8998.13.1.69] [Medline: 18211170]
19. Hedman E, Ljótsson B, Lindefors N. Cognitive behavior therapy via the internet: a systematic review of applications, clinical efficacy and cost-effectiveness. Expert Rev Pharmacoecon Outcomes Res 2012 Dec;12(6):745-764. [doi: 10.1586/erp.12.67] [Medline: 23252357]
20. Andersson G. Using the internet to provide cognitive behaviour therapy. Behav Res Ther 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010] [Medline: 19230862]
21. Heber E, Ebert DD, Lehr D, Cuijpers P, Berking M, Nobis S, et al. The benefit of web- and computer-based interventions for stress: a systematic review and meta-analysis. J Med Internet Res 2017 Feb 17;19(2):e32 [FREE Full text] [doi: 10.2196/jmir.5774] [Medline: 28213341]
22. Donker T, Blankers M, Hedman E, Ljótsson B, Petrie K, Christensen H. Economic evaluations of internet interventions for mental health: a systematic review. Psychol Med 2015 Dec;45(16):3357-3376. [doi: 10.1017/S0033291715001427] [Medline: 26235445]
23. Hamberg-van Reenen HH, Proper KI, van den Berg M. Worksite mental health interventions: a systematic review of economic evaluations. Occup Environ Med 2012 Nov;69(11):837-845. [doi: 10.1136/oemed-2012-100668] [Medline: 22864248]
24. Ebert DD, Kählke F, Buntrock C, Berking M, Smit F, Heber E, et al. A health economic outcome evaluation of an internet-based mobile-supported stress management intervention for employees. Scand J Work Environ Health 2018 Mar 1;44(2):171-182. [doi: 10.5271/sjweh.3691] [Medline: 29144535]
25. Kählke F, Buntrock C, Smit F, Berking M, Lehr D, Heber E, et al. Economic evaluation of an internet-based stress management intervention alongside a randomized controlled trial. JMIR Ment Health 2019 May 15;6(5):e10866 [FREE Full text] [doi: 10.2196/10866] [Medline: 31094355]
26. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The mini-international neuropsychiatric interview (MINI): 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]
27. Salomonsson S, Santoft F, Lindsäter E, Ejeby K, Ljótsson B, Öst LG, et al. Cognitive-behavioural therapy and return-to-work intervention for patients on sick leave due to common mental disorders: a randomised controlled trial. Occup Environ Med 2017 Dec;74(12):905-912. [doi: 10.1136/oemed-2017-104342] [Medline: 28756415]
28. Salomonsson S, Santoft F, Lindsäter E, Ejeby K, Ljótsson B, Öst LG, et al. Stepped care in primary care - guided self-help and face-to-face cognitive behavioural therapy for common mental disorders: a randomized controlled trial. Psychol Med 2018 Jul;48(10):1644-1654. [doi: 10.1017/S0033291717003129] [Medline: 29095133]
29. Sonnentag S, Venz L, Casper A. Advances in recovery research: what have we learned? What should be done next? J Occup Health Psychol 2017 Jul;22(3):365-380. [doi: 10.1037/ocp0000079] [Medline: 28358572]
30. Lejuez CW, Hopko DR, Acierno R, Daughters SB, Pagoto SL. Ten year revision of the brief behavioral activation treatment for depression: revised treatment manual. Behav Modif 2011 Mar;35(2):111-161. [doi: 10.1177/0145445510390929] [Medline: 21324944]
31. Abramowitz JS, Deacon BJ, Whiteside SP. Exposure Therapy for Anxiety: Principles and Practice. New York: Guilford Press; 2012.
32. van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]
33. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
34. Eklund M, Bäckström M, Tuvesson H. Psychometric properties and factor structure of the Swedish version of the perceived stress scale. Nord J Psychiatry 2014 Oct;68(7):494-499. [doi: 10.3109/08039488.2013.877072] [Medline: 24460116]
35. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
36. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
37. Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: recommendations for the design, analysis, and reporting of studies. Int J Technol Assess Health Care 2005;21(2):165-171. [doi: 10.1017/S0266462305050221] [Medline: 15921055]
38. Haakkart-van Roijen L. Trimbos/iMTA Questionnaire for Costs Associated with Psychiatric Illness (TiC-P). Rotterdam, Netherlands: Institute for Medical Technology Assessment; 2010.
39. Bouwmans C, de Jong K, Timman R, Zijlstra-Vlasveld M, van der Feltz-Cornelis C, Swan ST, et al. Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). BMC Health Serv Res 2013 Jun 15;13:217 [FREE Full text] [doi: 10.1186/1472-6963-13-217] [Medline: 23768141]
40. Osterhaus JT, Gutterman DL, Plachetka JR. Healthcare resource and lost labour costs of migraine headache in the US. Pharmacoeconomics 1992 Jul;2(1):67-76. [doi: 10.2165/00019053-199202010-00008] [Medline: 10146980]

41. Smit F, Willemse G, Koopmanschap M, Onrust S, Cuijpers P, Beekman A. Cost-effectiveness of preventing depression in primary care patients: randomised trial. Br J Psychiatry 2006 Apr;188:330-336. [doi: 10.1192/bjp.188.4.330] [Medline: 16582059]
42. Axelsson E, Andersson E, Ljótsson B, Hedman-Lagerlöf E. Cost-effectiveness and long-term follow-up of three forms of minimal-contact cognitive behaviour therapy for severe health anxiety: results from a randomised controlled trial. Behav Res Ther 2018 Aug;107:95-105. [doi: 10.1016/j.brat.2018.06.002] [Medline: 29936239]
43. Hedman-Lagerlöf M, Hedman-Lagerlöf E, Ljótsson B, Wicksell RK, Flink I, Andersson E. Cost-effectiveness and cost-utility of internet-delivered exposure therapy for fibromyalgia: results from a randomized, controlled trial. J Pain 2019 Jan;20(1):47-59. [doi: 10.1016/j.jpain.2018.07.012] [Medline: 30107241]
44. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [doi: 10.1037//0022-006x.59.1.12] [Medline: 2002127]
45. Lavoie JA, Douglas KS. The perceived stress scale: evaluating configural, metric and scalar invariance across mental health status and gender. J Psychopathol Behav Assess 2011 Nov 27;34(1):48-57. [doi: 10.1007/s10862-011-9266-1]
46. Burström K, Sun S, Gerdtham UG, Henriksson M, Johannesson M, Levin LA, et al. Swedish experience-based value sets for EQ-5D health states. Qual Life Res 2014 Mar;23(2):431-442 [FREE Full text] [doi: 10.1007/s11136-013-0496-4] [Medline: 23975375]
47. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ 1994;3(5):309-319. [doi: 10.1002/hec.4730030505] [Medline: 7827647]
48. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. Br J Psychiatry 2005 Aug;187:106-108. [doi: 10.1192/bjp.187.2.106] [Medline: 16055820]
49. Bock JO, Brettschneider C, Seidl H, Bowles D, Holle R, Greiner W, et al. [Calculation of standardised unit costs from a societal perspective for health economic evaluation]. Gesundheitswesen 2015 Jan;77(1):53-61. [doi: 10.1055/s-0034-1374621] [Medline: 25025287]
50. Kigozi J, Jowett S, Lewis M, Barton P, Coast J. The estimation and inclusion of presenteeism costs in applied economic evaluation: a systematic review. Value Health 2017 Mar;20(3):496-506 [FREE Full text] [doi: 10.1016/j.jval.2016.12.006] [Medline: 28292496]
51. Layard R. IZA World of Labor. 2017. The Economics of Mental Health URL: https://wol.iza.org/articles/economics-of-mental-health/long [accessed 2019-02-22]
52. McEwen BS. Epigenetic interactions and the brain-body communication. Psychother Psychosom 2017;86(1):1-4 [FREE Full text] [doi: 10.1159/000449150] [Medline: 27884000]
53. Mohr DC, Spring B, Freedland KE, Beckner V, Arean P, Hollon SD, et al. The selection and design of control conditions for randomized controlled trials of psychological interventions. Psychother Psychosom 2009;78(5):275-284. [doi: 10.1159/000228248] [Medline: 19602916]
54. Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. Int Rev Psychiatry 2011 Aug;23(4):318-327 [FREE Full text] [doi: 10.3109/09540261.2011.606803] [Medline: 22026487]
55. Lensberg BR, Drummond MF, Danchenko N, Despiégel N, François C. Challenges in measuring and valuing productivity costs, and their relevance in mood disorders. Clinicoecon Outcomes Res 2013;5:565-573 [FREE Full text] [doi: 10.2147/CEOR.S44866] [Medline: 24273412]
56. Patel A, Rendu A, Moran P, Leese M, Mann A, Knapp M. A comparison of two methods of collecting economic data in primary care. Fam Pract 2005 Jun;22(3):323-327. [doi: 10.1093/fampra/cmi027] [Medline: 15824056]
57. Cuijpers P, Weitz E, Twisk J, Kuehner C, Cristea I, David D, et al. Gender as predictor and moderator of outcome in cognitive behavior therapy and pharmacotherapy for adult depression: an 'individual patient data' meta-analysis. Depress Anxiety 2014 Nov;31(11):941-951. [doi: 10.1002/da.22328] [Medline: 25407584]
58. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: a randomized controlled trial. J Med Internet Res 2016 Jan 27;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline: 26818683]
59. Netterstrøm B, Friebel L, Ladegaard Y. Effects of a multidisciplinary stress treatment programme on patient return to work rate and symptom reduction: results from a randomised, wait-list controlled trial. Psychother Psychosom 2013;82(3):177-186. [doi: 10.1159/000346369] [Medline: 23548852]
60. Mihalopoulos C, Chatterton ML. Economic evaluations of interventions designed to prevent mental disorders: a systematic review. Early Interv Psychiatry 2015 Apr;9(2):85-92. [doi: 10.1111/eip.12156] [Medline: 24962626]

Abbreviations

6MFU: 6-month follow-up
AD: adjustment disorder
CBT: cognitive behavioral therapy
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th edition

ED: exhaustion disorder
EQ-5D: EuroQol 3L Questionnaire
ICBT: internet-based cognitive behavioral therapy
ICER: incremental cost-effectiveness ratio
ICUR: incremental cost-utility ratio
PSS: Perceived Stress Scale
QALY: quality-adjusted life year
SEK: Swedish Krona
TIC-P: Trimbos and Institute Medical Technology Assessment of Costs Questionnaire for Psychiatry
WLC: waitlist control
WTP: willingness to pay

Edited by G Eysenbach; submitted 13.05.19; peer-reviewed by P Romero-Sanchiz, C Buntrock; comments to author 03.06.19; revised version received 14.07.19; accepted 27.07.19; published 13.09.19.

Please cite as:
Lindsäter E, Axelsson E, Salomonsson S, Santoft F, Ljótsson B, Åkerstedt T, Lekander M, Hedman-Lagerlöf E
Cost-Effectiveness of Therapist-Guided Internet-Based Cognitive Behavioral Therapy for Stress-Related Disorders: Secondary Analysis of a Randomized Controlled Trial
J Med Internet Res 2019;21(9):e14675
URL: https://www.jmir.org/2019/9/e14675
doi:10.2196/14675
PMID:31586370

Original Paper

Telemonitoring of Crohn’s Disease and Ulcerative Colitis (TECCU): Cost-Effectiveness Analysis

Javier Del Hoyo[1*], MD; Pilar Nos[1,2,3], MD, PhD; Guillermo Bastida[1,2,3], MD, PhD; Raquel Faubel[4,5], PhD; Diana Muñoz[1], RN; Alejandro Garrido-Marín[1], MD; Elena Valero-Pérez[1], MD; Sergio Bejar-Serrano[1], MD; Mariam Aguas[1,2,3*], MD, PhD

[1]Gastroenterology Department, La Fe University and Polytechnic Hospital, Valencia, Spain
[2]Networked Biomedical Research Center for Hepatic and Digestive Diseases, Valencia, Spain
[3]Health Research Institute La Fe, Valencia, Spain
[4]Physiotherapy Department, University of Valencia, Valencia, Spain
[5]Joint Research Unit in Biomedical Engineering-eRPSS, Health Research Institute La Fe–Polytechnic University of Valencia, Valencia, Spain
[*]these authors contributed equally

Corresponding Author:
Mariam Aguas, MD, PhD
Gastroenterology Department
La Fe University and Polytechnic Hospital
Av Fernando Abril Martorell 106
Valencia, 46026
Spain
Phone: 34 961245859
Fax: 34 961246257
Email: m.aguas.peris@gmail.com

Abstract

Background: Although electronic health interventions are considered safe and efficient, evidence regarding the cost-effectiveness of telemonitoring in inflammatory bowel disease is lacking.

Objective: We aimed to evaluate the cost-effectiveness and cost-utility of the *Telemonitorización de la Enfermedad de Crohn y Colitis Ulcerosa* (Telemonitoring of Crohn’s Disease and Ulcerative Colitis [TECCU]) Web platform (G_TECCU intervention group) for telemonitoring complex inflammatory bowel disease, compared with standard care (G_control) and nurse-assisted telephone care (G_NT intervention group).

Methods: We analyzed cost-effectiveness from a societal perspective by comparing the 3 follow-up methods used in a previous 24-week randomized controlled trial, conducted at a tertiary university hospital in Spain. Patients with inflammatory bowel disease who initiated immunosuppressants or biologic agents, or both, to control inflammatory activity were recruited consecutively. Data on the effects on disease activity (using clinical indexes) and quality-adjusted life-years (using the EuroQol 5 dimensions questionnaire) were collected. We calculated the costs of health care, equipment, and patients’ productivity and social activity impairment. We compared the mean costs per patient, utilities, and bootstrapped differences.

Results: We included 63 patients (21 patients per group). TECCU saved €1005 (US $1100) per additional patient in remission compared with G_control (95% CI €–13,518 to 3137; US $–14,798 to 3434), with a 79.96% probability of being more effective at lower costs. Compared with G_NT, TECCU saved €2250 (US $2463) per additional patient in remission (95% CI €–15,363 to 11,086; US $–16,817 to 12,135), and G_NT saved €538 (US $589) compared with G_control (95% CI €–6475 to 5303; US $–7088 to 5805). G_TECCU and G_NT showed an 84% and 67% probability, respectively, of producing a cost saving per additional quality-adjusted life-year (QALY) compared with G_control, considering those simulations that involved negative incremental QALYs as well.

Conclusions: There is a high probability that the TECCU Web platform is more cost-effective than standard and telephone care in the short term. Further research considering larger cohorts and longer time horizons is required.

Trial Registration: ClinicalTrials.gov NCT02943538; https://clinicaltrials.gov/ct2/show/NCT02943538 (http://www.webcitation.org/746CRRtDN)

(J Med Internet Res 2019;21(9):e15505) doi:10.2196/15505

KEYWORDS

telemedicine; eHealth; cost-effectiveness; inflammatory bowel diseases; Crohn disease; colitis, ulcerative

Introduction

Background

Health systems are facing problems of financial sustainability, and the burden and health care costs associated with the management of inflammatory bowel disease (IBD) continue to rise [1,2]. IBD is one of the most expensive gastrointestinal conditions [3]. In this context, interest in electronic health (eHealth) interventions as a potential means to improve health care services at a lower cost has grown in recent years, especially for the management of chronic diseases such as IBD [4,5].

Unlike other chronic pathologies, IBD mainly affects young individuals in their optimal period of personal and professional development. As such, IBD is related to high levels of school absenteeism and work disability [6], interference in social activities, and impaired health-related quality of life [7]. Therefore, IBD has a significant medical, social, and financial impact, which accounts for direct and indirect costs to both health care systems and society. In fact, although the estimated indirect costs differ between countries depending on their health policies, they are an important percentage of the IBD-related economic burden. A recent report estimated them to account for 46.5% of the total IBD-related costs in Spain [8].

Web-based telemonitoring systems applied to IBD are safe and feasible not only for adults but also for adolescents [9-11], and empowerment through these systems reduces outpatient visits and hospital admissions, with potential cost savings [12-14]. However, even if eHealth is considered a promising option to improve the quality of care while reducing costs, its efficacy in terms of disease outcomes has not been consistent across studies [10-13,15,16], with high attrition rates despite the continued adaptation of Web platforms and the evolution of mobile health over the last decade.

These systems aim to shift the emphasis from hospital and personal visits toward remote encounters, although it is still necessary to determine whether such telemonitoring systems actually decrease direct and indirect costs [12,13,16]. While cost savings have been attributed to such systems in the IBD setting, these are almost exclusively related to direct costs [12,14]. Indeed, these studies did not consider either the costs associated with the implementation and maintenance of the remote monitoring systems or indirect costs. Thus, evidence regarding the cost-effectiveness and cost-utility of telemonitoring in IBD is lacking [17], without economic data available to evaluate its use in patients with complex IBD, hindering the long-term implementation of these systems by health care services [18].

Objective

Our research group developed a Web-based telemanagement program known as *Telemonitorización de la Enfermedad de Crohn y Colitis Ulcerosa* (Telemonitoring of Crohn’s Disease and Ulcerative Colitis [TECCU]) [19]. The results of a pilot clinical trial suggested that this remote monitoring system is a safe strategy to improve health outcomes in patients with complex IBD, while reducing the use of health care resources when compared with nurse-assisted telephone care and standard face-to-face visits [20]. Due to the limited knowledge about the efficiency of telemonitoring in IBD and given the absence of health economics data in a non–remission setting, we aimed to evaluate the cost-effectiveness and cost-utility of the Web-based TECCU system for remote management of patients with complex IBD from a societal perspective, comparing telemonitoring versus standard care and telephone care.

Methods

Study Design

We performed a cost-effectiveness analysis alongside a previously published 3-arm parallel-group randomized controlled trial, which had been carried out at a referral university hospital in Spain [20]. We evaluated the impact of the TECCU Web-based telemanagement system (G_TECCU intervention group), nurse-assisted telephone care (G_NT intervention group), and standard face-to-face visits (G_control) on the health outcomes and the direct and indirect costs associated with patients with complex IBD, after a 24-week follow-up. This follow-up period allowed us to evaluate the impact of these 3 interventions and the events that occurred during the initiation of therapy with immunosuppressants and biologic agents.

With the aim of analyzing cost-effectiveness and cost-utility from a societal perspective, we used a standard economic evaluation methodology to measure the costs and effects on disease activity and quality of life associated with each of the 3 interventions [21]. First, we determined the categories to be included in both the costs and effects analyses. We included health care, equipment, and productivity and social activities in the costs study, while the effects measured were disease activity and quality of life. Subsequently, to analyze the costs in each category, we measured the number of units of health care and non–health care resources consumed, thereafter calculating the cost in euros (and including the exchange in US $) by multiplying the number of units used by their price weight. Finally, we correlated the cost and effect data to obtain a cost-effectiveness and cost-utility comparison between the 3 follow-up strategies.

Patient Selection and Setting

Patients were recruited consecutively at the Outpatient Clinic of the IBD Unit at La Fe University and Polytechnic Hospital, Valencia, Spain, or at the Gastroenterology Department if they were admitted for a flare-up. This is a tertiary hospital that serves more than 1500 patients with IBD, and it has 2 specialist IBD nurses, also providing its patients with an email and telephone consultation service. All participants had IBD diagnosed according to internationally accepted criteria at least 6 months prior to their recruitment [22,23]. The inclusion criteria were age 18 years or older; and having initiated therapy with corticosteroids, immunosuppressants, or biologic agents, or a combination of these, due to disease activity. The exclusion criteria were inability to speak and read Spanish; inability to manage a mobile phone or tablet, or the internet, or not having a telephone line; participation in other clinical trials during the inclusion period; having other uncontrolled medical or psychiatric disease; the presence of ileorectal or ileal pouch-anal anastomosis; having received a definitive ileostomy; having associated perianal disease; and being pregnant. All participants provided their written informed consent to participate in the study. Enrollment began in October 2014 and ended in June 2016. The follow-up ended in December 2016. Eligible patients were randomly assigned to 1 of the 3 groups to G_TECCU, G_NT, or G_control in a 1:1:1 ratio. A block randomization method was used via a Web-based tool [19] to generate a random allocation sequence and ensure allocation concealment.

Interventions

Regardless of the assigned arm, all patients completed 3 face-to-face visits, at baseline, 12 weeks, and 24 weeks, in addition to their routine visits to the IBD clinic, telephone consultations, or Web telemonitoring in accordance with their group assignment. Patients treated with immunosuppressants alone or in combination with biologic agents were monitored every 1 to 2 weeks during the first month, every 2 to 4 weeks between month 1 and month 3, and every 4 weeks from month 3 until the end of the follow-up. Patients treated with biologic agents alone were monitored every 2 to 4 weeks during the entire follow-up period. Patients from all 3 arms who initiated the same drug complied with these follow-up schedules, which differed only in the monitoring method used for the study group they were assigned to. In any of the 3 arms, additional clinical visits were made when necessary if the patient's evolution so required. The 3 interventions evaluated in this clinical trial have been described previously [19] but we present a summary below.

TECCU Web Platform

In the G_TECCU follow-up and monitoring was performed telematically using the TECCU app. The patients used a computer, or an app on a mobile phone or tablet, to connect via the internet and self-complete questionnaires related to their IBD symptoms. Through these questionnaires, they also provided information regarding any possible adverse effects. The health care providers used the information obtained from questionnaires and biological markers to adapt medication and follow-up schedules through an intelligent prioritization system. These changes were communicated through the platform's messaging system, in combination with telephone calls or in-person visits when necessary.

Usual Care Provided at the Inflammatory Bowel Disease Unit

Patients from the G_control received usual care provided by the IBD unit (Outpatient Clinic) based on national and European clinical guidelines [22-24]. Treatment was adjusted at face-to-face visits or via information provided through telephone calls based on the evolution of disease activity. Disease activity was measured using specific indexes and through biological markers.

Nursing Care by Telephone

In the G_NT, patients were asked about their health during telephone calls with the nursing staff at the IBD Unit. Periodic telephone assessment was carried out using structured interviews to evaluate the patient's health status based on the same clinical indexes and biological markers used for the other 2 groups. The interventions depended on the results of the interview, and any changes to medication or the follow-up schedule were established by the nurses with the aid of a physician through alerts and action plans incorporated into the intervention protocol [19].

Cost Measures

We established 3 major cost categories to perform an economic evaluation from a societal perspective: health care costs, equipment costs, and costs related to patients' work productivity and impaired social activity. We expressed all costs in euros (including the exchange in US $) and corrected for price inflation in 2016 according to the Spanish consumer price index [25].

Health Care Costs

We measured health care resource use on the basis of the number of outpatient visits, telephone calls, emergency visits, hospitalizations, and IBD-related surgeries detected in the hospital registry over the study period. We calculated their associated costs by multiplying the number of services registered by the official regional rates [26]. We also considered telemonitoring contacts through the Web platform in G_TECCU. As there is as yet no rate for telemonitoring contacts in Spain, we calculated this rate by multiplying the mean time spent by health care providers in each telemonitoring contact (8 minutes, based on consultation with nurses and physicians) by the mean cost of their salary per minute (€0.21 [US $0.23]/minute for a nurse and €0.38 [US $0.42]/minute for a physician), using data published by the Spanish National Statistical Institute [27].

Equipment Costs

Equipment costs included those related to software development, Web security, and technical support. These equipment costs were incorporated into the health care cost of Web telemonitoring, as the software is a necessary expenditure for this health care initiative.

Productivity and Social Activity Costs

To assess the number of hours that patients lost from work and social activities, we used the Work Productivity and Activity Impairment questionnaire [28]. The Spanish version of this questionnaire has been validated in patients with Crohn disease. We measured the number of working hours lost due to absenteeism and presenteeism associated with disease activity at baseline and at 12 and 24 weeks, and we also registered the number of hours lost from the patients' leisure time in the same periods.

We calculated the cost of absenteeism by multiplying the hours lost (question 2 of the Work Productivity and Activity Impairment questionnaire) by the average hourly wage for age and sex according to the Spanish National Statistical Institute [29]. We estimated lost work hours due to presenteeism by multiplying the percentage impairment while working (question 5) by the hours worked over the past 7 days (question 4), and then calculated the cost similarly to the cost of absenteeism hours. To calculate the cost of leisure time lost, we used the percentage of impairment of social activities (question 6) and data regarding the costs of leisure time based on manuals for economic evaluations previously published in Europe [30]. Due to the limited time horizon, we did not discount costs and we considered the human capital approach to evaluate costs associated with the loss of productivity [31].

In addition, we calculated the cost of absenteeism associated with medical visits, assuming a loss of 3.3 hours per visit, as previously reported for patients with Crohn disease in Spain [32]. We calculated the cost of the time that patients spent on telephone contacts assuming that they were conducted during working hours, because they were made between 8 AM and 3 PM. We considered the cost of leisure time for contacts through the Web telemonitoring platform, assuming that patients accessed this tool out of office hours.

Effect Measures

The primary outcome used for the cost-effectiveness analysis was the effect of each intervention (G_TECCU, G_NT, and G_control) on the percentage of patients in remission throughout the study. The clinical indexes used for the follow-up in the 3 arms were the Harvey-Bradshaw index for patients with Crohn disease [33], and the Simple Clinical Colitis Activity Index and the partial Mayo score for patients with ulcerative colitis [34,35]. As described in the study protocol [19], we considered remission to be indicated by a Harvey-Bradshaw index score of 4 or less in patients with Crohn disease, and a Simple Clinical Colitis Activity Index and partial Mayo score of 2 or less in patients with UC. We compared the proportion of patients in remission at baseline versus that after 24 weeks.

We also measured the number of quality-adjusted life-years (QALYs) for the cost-utility analysis. To measure quality of life, patients answered the EuroQol 5 dimensions questionnaire (EQ-5D) [36] at baseline and at 24 weeks, from which we calculated QALYs using specific coefficients established for the Spanish population [37]. We also briefly described patients' perceptions and satisfaction with the care received at baseline (previous standard care) and at 24 weeks (assigned follow-up intervention) by using a previously adapted version of the Client Satisfaction Questionnaire [38] for the study purpose, which we described in more detail elsewhere [20].

Cost-Effectiveness and Cost-Utility Analyses

We divided the calculated costs by the effects of the 3 interventions over 24 weeks to assess cost-effectiveness (considering the improvement in disease activity) and cost-utility (considering the impact on QALYs). The differences in costs between the 3 interventions divided by the differences in their effects allowed us to obtain the incremental cost-effectiveness ratio (ICER). In decision making, ICERs are more useful when the new intervention is more costly but generates an improved health effect [39].

To evaluate the strength of the economic evaluation and the influence of the cost values on the median ICERs calculated, we carried out sensitivity analyses in which we increased and reduced the main cost drivers independently. We also evaluated the uncertainty in the cost-effectiveness and cost-utility analyses. We visually represented the distribution of possible values that the ICERs could acquire as dots plotted on a cost-effectiveness plane (see Figure 1). In this plane, the horizontal axis represents the differences in health outcomes and the vertical axis represents the differences in costs between the 3 interventions compared (G_TECCU vs G_control, G_TECCU vs G_NT, and G_NT vs G_control).

We drew cost-effectiveness acceptability curves to map (on the vertical axis) the evolution of the probability that one intervention is cost-effective compared with another, as a function of the willingness to pay (WTP) for 1 additional unit of effect in a range of €0 to €20,000 (US $0 to 21,893, represented in the horizontal axis; see Figure 2). As part of the sensitivity analysis, we also assessed the influence of alternative costing scenarios on the cost-effectiveness acceptability curve.

Figure 1. Generic cost-effectiveness plane (left) and an example illustrating the bootstrapped incremental cost-effectiveness ratios plotted (right).

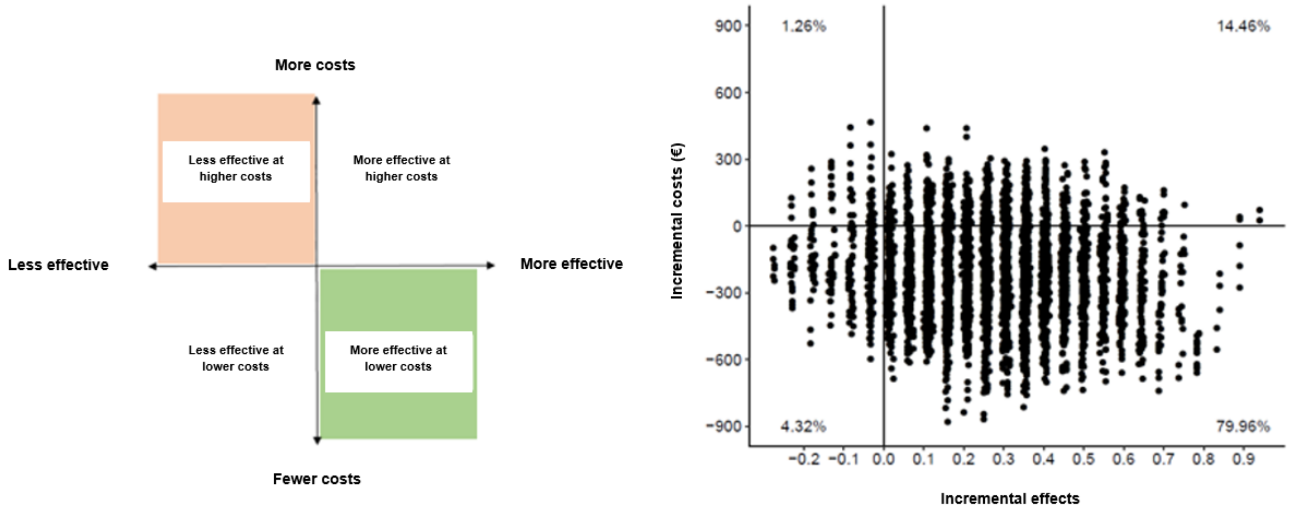
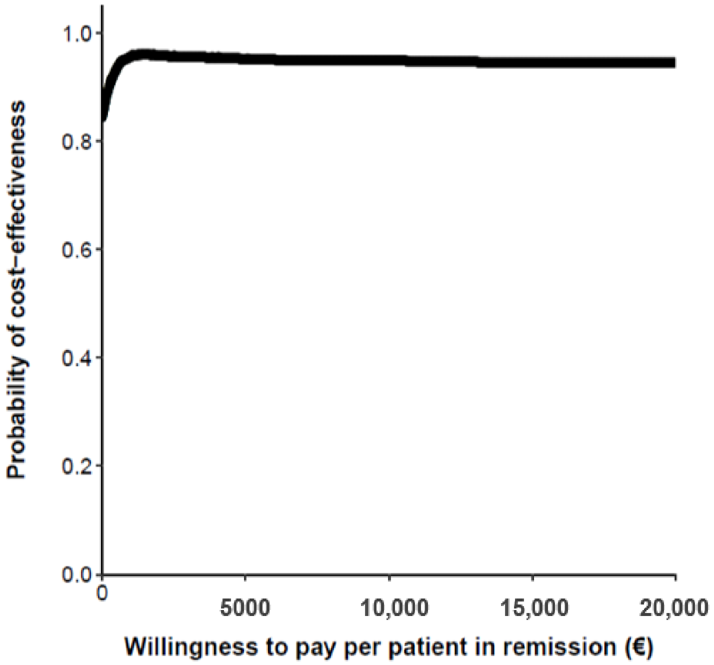


Figure 2. Cost-effectiveness acceptability curve representing the probability that one intervention is cost-effective relative to another, depending on the willingness-to-pay value.



Statistical Analysis

We handled missing observations in costs and effects data using multiple imputation with the Amelia II for R software package version 1 [40]. We imputed the original dataset 5 times and then analyzed each of these 5 datasets separately, subsequently combining the outcomes using Rubin rules. We calculated the differences and statistical uncertainty in the disease activity, QALYs, and costs using nonparametric bootstrap estimations, which consisted of extracting 1000 random samples (n=21 per trial arm) from each of the 5 imputations. For each of these samples, we calculated the incremental costs, incremental effects, and ICERs. We performed these analyses using R version 3.5.1 (R Foundation).

Ethical Considerations

The study protocol was reviewed and approved by the local independent ethics committee at La Fe University and Polytechnic Hospital, Valencia, Spain; by the regional independent ethics committee (*Comité Ético Autonómico de Estudios Clínicos de Medicamentos y Productos Sanitarios de la Comunitat Valenciana*); and by the Spanish Agency of Medicines and Medical Devices (*Agencia Española de Medicamentos y Productos Sanitarios*). According to the physicians involved in the study, the risks did not outweigh the potential benefits, and each participant provided their informed consent without coercion before inclusion in the study. The randomized controlled trial is registered at ClinicalTrials.gov (NCT02943538; Multimedia Appendix 1 [41]).

Results

Patient Characteristics

We invited a total of 68 patients with complex IBD to participate in this study, of whom 3 (4%) declined to participate, as they did not have internet access at their home, and 2 (3%) did not meet the inclusion criteria. The remaining 63 eligible patients were randomly assigned to the 3 groups (21 patients in each group). During the study period, 3 patients in G_TECCU did not adhere to the study protocol, while the remaining 18 patients (86%) showed good adherence, as compared with 19 patients (90%) in G_control and 20 patients (95%) in G_NT. The mean age of the patients was 39.50 (SD 12.06) years, and 52% (33/63) of the patients were women. At baseline, the patients in G_TECCU and G_NT had higher fecal calprotectin levels and satisfaction scores with previous in-person care, with a lower adherence to medication in G_NT according to the validated 4-item Morisky-Green questionnaire [42], which is easy to answer and has been previously used to evaluate adherence to telemonitoring in IBD patients [15]. The other baseline characteristics did not differ markedly between the groups, although quality-of-life scores were slightly higher in G_TECCU and G_NT (Table 1).

Table 1. Patients' baseline characteristics.

Characteristics	Control group (n=21)	Telephone care group (n=21)	TECCU[a] care group (n=21)
Age (years), median (range)	39.31 (22-61)	40.91 (24-60)	41.32 (19-66)
Sex, n (%)			
Male	12 (57)	12 (57)	9 (43)
Female	9 (43)	9 (43)	12 (57)
Education, n (%)			
Primary	4 (19)	4 (19)	5 (23)
Secondary	9 (43)	6 (29)	6 (29)
University	8 (38)	11 (52)	10 (48)
Disease profile, n (%)			
Crohn disease	14 (67)	13 (62)	13 (62)
Ulcerative colitis	7 (33)	8 (38)	8 (38)
Median time since diagnosis, months (range)	123.32 (6-427)	108.27 (7-452)	146.72 (7-424)
Treatment, n (%)			
Immunomodulators	10 (48)	10 (48)	9 (43)
Biologic monotherapy	4 (19)	4 (19)	4 (19)
Combination therapy	6 (29)	5 (24)	6 (29)
Corticosteroids	1 (5)	2 (10)	2 (10)
Calprotectin (µg/g), median (IQR[b])	330 (103-617)	526 (115-1724)	490 (23-2016)
Quality-of-life scores, median (IQR)			
IBDQ-9[c]	38.50 (33.25-46.75)	37.50 (28.75-46.25)	42.00 (33.75-47.50)
EQ-5D[d]	0.82 (0.75-0.91)	0.83 (0.71-0.92)	0.83 (0.58-0.91)
Visual analog scale (%)	60.50 (50-85)	62.50 (50-80)	60 (40-90)
Medication adherence, n (%)	14 (67)	7 (33)	12 (57)
Work Productivity and Activity Impairment questionnaire responses			
Not working, n (%)	8 (38)	7 (33)	5 (24)
Work hours missed, % median (IQR)	27.50 (0-52)	40 (15-62.50)	32.50 (7.50-57.50)
Work impairment score, median (IQR)	7 (2.75-10)	7 (3-10)	10 (2.25-10)
Social impairment score, median (IQR)	3.50 (1-5.75)	3.50 (2-7)	6 (2.75-8)
Satisfaction score, median (IQR)	49.50 (42.50-53.75)	53 (50-59)	52 (47.50-55)

[a] TECCU: Telemonitoring of Crohn's Disease and Ulcerative Colitis.

[b] IQR: interquartile range.

[c] IBDQ-9: Inflammatory Bowel Disease Questionnaire 9.

[d] EQ-5D: EuroQol 5 dimensions questionnaire.

Effects

Disease Activity

In terms of the effect on disease activity, the proportion of patients in clinical remission at baseline was 48% (10/21) in G_TECCU, 38% (8/21) in G_NT, and 57% (12/21) among the controls. The proportion of patients in remission improved after 24 weeks in all 3 groups, although this improvement was stronger among the G_TECCU participants, even after considering in the likelihood-based analysis the 3 patients who did not complete the follow-up schedule in G_TECCU: 81% (17/21) of G_TECCU patients were in remission as opposed to 67% (14/21) in the G_NT and 71% (15/21) in the G_control. The proportion of patients in remission increased by 0.33 in G_TECCU, 0.29 in G_NT, and 0.14 in G_control. Thus, the incremental efficacy of G_TECCU was 0.33–0.14=0.19 relative to G_control (median incremental efficacy calculated with the bootstrapping procedure was 0.21, 95% CI –0.07 to 0.66), and the incremental efficacy of TECCU was 0.33–0.29=0.04 relative to the G_NT (median incremental efficacy calculated with the bootstrapping procedure was 0.06, 95% CI –0.16 to 0.43).

Quality of Life

The median EQ-5D score improved after 24 weeks in all 3 arms, from 0.816 to 1.00 in G_control and from 0.825 to 1.00 in G_NT and G_TECCU (overall intervention effect on the EQ-5D score: odds ratio 1.99, 95% CI 1.09-3.63; *P*<.001). Consequently, the improvement in the median EQ-5D score in the control group was 0.184 (ie, 1.00–0.816), and that in the G_NT and G_TECCU was 0.175 (ie, 1.00–0.825). Considering that participant mortality was zero and given the 24-week (6-month) time frame of this study, the number of QALYs gained by each patient treated with the standard intervention (controls) was 0.184×(6/12)=0.092 QALYs, and for each patient treated in G_NT and G_TECCU it was 0.175×(6/12)=0.088 QALYs.

All patients considered the care received through the telemonitoring platform to be useful. The 89% (16/18) surveyed patients in G_TECCU indicated that the distance follow-up took little time. The 94% (17/18) surveyed patients rated the quality of the services received remotely as high (8 or more out of 10 points), which helped them to more effectively face their disease-related problems. In all 3 groups, satisfaction scores were high at baseline and at 24 weeks. According to the adapted version of the Client Satisfaction Questionnaire, with a maximum score of 60 points, patient satisfaction improved from a median score of 52 to 57 in G_TECCU and from 49.50 to 55 in G_control (overall intervention effect: odds ratio 8.93, 95% CI 2.97-26.84; *P*<.001) at 24 weeks. However, the satisfaction score remained unchanged at 53 points in G_NT.

Costs

We calculated the costs per unit of health care resources, equipment, and work productivity and social activities (Table 2). The mean costs and utilities per patient, as well as the bootstrapped differences, were compared between G_TECCU vs G_control (Tables 3 and 4), G_TECCU vs G_NT (Tables 5 and 6), and G_NT vs G_control (Tables 7 and 8).

Table 2. Costs per unit of health care resources, equipment, and productivity and social activities.

Costs	Unit	Cost per unit, € (US $)
Health care costs		
Emergency room visits	Visit	189 (206.89)
Outpatient visits	Visit	40.02 (43.81)
Hospitalization	Day	310.17 (339.53)
Hospitalization due to surgical intervention	Day	378 (413.78)
Telephone calls	Contact	15 (16.42) nurse 21.47 (23.50) physician
Cost of TECCU[a] controls	Contact	1.68 (1.84) nurse 3.04 (3.33) physician
Equipment costs		
TECCU rental costs	Patient/month	3.99 (4.37)
Productivity costs		
Work absenteeism (sick leave)	Hour	12.04-25.23 (13.18-27.62)
Work presenteeism (due to disease activity)	Hour	12.04-25.23 (13.18-27.62)
Absenteeism for medical visit	Hour	12.04-25.23 (13.18-27.62)
Absenteeism for telephone call	Hour	12.04-25.23 (13.18-27.62)
Leisure time used in TECCU contacts	Hour	9.18 (10.05)

[a] TECCU: Telemonitoring of Crohn’s Disease and Ulcerative Colitis.

Table 3. Costs per patient in the group receiving remote monitoring (G_TECCU) vs standard care (G_control).

Costs	TECCU[a] (n=21), mean (SD)		Controls (n=21), mean (SD)		Bootstrapped difference in costs (rounded values), € (US $)	
	Number of units	Cost per patient, €; US $	Number of units	Cost per patient, €; US $	Median	95% CI
Health care costs						
Emergency room visits and nonscheduled outpatient visits	0.27 (0.68)	51.03 (115.06); 55.86 (125.95)	0.19 (0.87)	36.18 (165.03); 39.60 (180.65)	17 (19)	–57 to 116 (62 to 127)
Outpatient visits	3.79 (1.38)	151.67 (58.72); 166.03 (64.28)	6.30 (0.57)	252.13 (22.91); 276.00 (25.08)	–96 (–105)	–126 to –65 (–138 to –71)
Hospitalization	0.11 (0.32)	32.63 (97.70); 35.72 (106.95)	0.05 (0.22)	14.79 (67.65); 16.19 (74.05)	16 (18)	–29 to 63 (–32 to 69)
Hospitalization due to surgical intervention	0.19 (0.89)	71.82 (312.89); 78.62 (342.51)	0.17 (0.87)	65.10 (297.74); 71.26 (325.93)	3 (3)	–195 to 205 (–213 to 224)
Telephone calls	0.51 (0.76)	9.18 (11.31); 10.05 (12.38)	2.04 (1.62)	32.05 (21.41); 35.08 (23.44)	–24 (–26)	–33 to –14 (–36 to –15)
Cost of TECCU controls	33.15 (5.78)	68.96 (12.07); 75.49 (13.21)	N/A[b]	N/A	67 (73)	59 to 75 (65 to 82)
Equipment costs						
TECCU rental costs	6 months	23.94 (0); 26.21 (0)	N/A	N/A	24 (26)	22 to 27 (24 to 30)
Productivity costs						
Work absenteeism (sick leave)	15.09 (21.19)	213.97 (300.51); 234.23 (328.96)	25.77 (27.25)	384.22 (418.00); 420.59 (457.57)	–42 (–46)	–234 to 146 (–256 to 160)
Work presenteeism (due to disease activity)	11.88 (17.26)	168.42 (224.04); 184.36 (245.25)	24.35 (34.53)	363.06 (549.48); 397.43 (601.50)	–89 (–97)	–299 to 91 (–327 to 100)
Absenteeism for medical visits	13.20 (4.93)	187.22 (62.30); 204.94 (68.20)	20.79 (4.79)	309.98 (113.41); 339.32 (124.15)	–125 (–137)	–189 to –58 (–207 to –63)
Absenteeism for telephone calls	0.12 (0.15)	1.71 (2.24); 1.87 (2.45)	0.36 (0.27)	5.39 (3.94); 5.90 (4.31)	–3.81 (–4.17)	–5.82 to –2 (–6.37 to –2.19)
Leisure time used in TECCU contacts	1.79 (0.31)	25.37 (4.42); 27.77 (4.84)	N/A	N/A	25 (27)	22 to 29 (24 to 32)
Total productivity costs per patient (rounded values)	N/A	407 (339); 445 (371)	N/A	678 (686); 742 (751)	–260 (–285)	–600 to 71 (–657 to 78)
Total costs per patient (rounded values)	N/A	807 (623); 883 (682)	N/A	1066 (678); 1167 (742)	–211 (–231)	–600 to 180 (–657 to 197)

[a] TECCU: Telemonitoring of Crohn’s Disease and Ulcerative Colitis.

[b] N/A: not applicable.

Table 4. Utilities per patient in the group receiving remote monitoring (G_TECCU) vs standard care (G_control).

Utilities	TECCU[a] (n=21)	Controls (n=21)	Bootstrapped difference in effects (rounded values)	
			Median	95% CI
Effects				
EQ-5D[b] score week 24, mean (SD)	0.90 (0.19)	0.93 (0.15)	–0.03	–0.14 to 0.08
EQ-5D score improvement week 0-24, mean (SD)	0.09 (0.28)	0.10 (0.19)	–0.02	–0.16 to 0.11
Weeks in remission, mean (SD)	17.89 (7.03)	14.27 (8.13)	3.87	–1.09 to 8.72
Calprotectin at week 24 (μg/g), median (interquartile range)	126 (47.24)	230 (48.67)	–104	–504 to 75
Remission at week 24, n (%)	17 (80.95)	15 (71.43)	0.19	–0.01 to 0.42
Improvement in remission week 0-24, n (%)	7 (33.33)	3 (14.28)	0.21	–0.07 to 0.66
ICER[c] remission response (rounded values), € (US $)	N/A[d]	N/A	–1005 (–1100)	–13,518 to 3137 (–14,798 to 3434)
ICER quality-adjusted life-years (rounded values), € (US $)	N/A	N/A	9078 (9937)	–56,547 to 44,628 (–61,900 to 48,853)

[a] TECCU: Telemonitoring of Crohn’s Disease and Ulcerative Colitis.

[b] EQ-5D: EuroQol 5 dimensions questionnaire.

[c] ICER: incremental cost-effectiveness ratio.

[d] N/A: not applicable.

Table 5. Costs per patient in the group receiving remote monitoring (G_TECCU) vs nurse-assisted telephone care (G_NT).

Costs	TECCU[a] (n=21), mean (SD)		Telephone care (n=21), mean (SD)		Bootstrapped difference in costs (rounded values), € (US $)	
	Number of units	Cost per patient, €; US $	Number of units	Cost per patient, €; US $	Median	95% CI
Health care costs						
Emergency room visits and nonscheduled outpatient visits	0.27 (0.68)	51.03 (115.06); 55.86 (125.95)	0.57 (1.17)	108.15 (220.18); 118.39 (241.02)	–48 (–53)	–150 to 55 (–164 to 60)
Outpatient visits	3.79 (1.38)	151.67 (58.72); 166.03 (64.28)	4.25 (1.92)	170.03 (76.68); 186.13 (83.94)	–13 (–14)	–59 to 30 (–65 to 33)
Hospitalization	0.11 (0.32)	32.63 (97.70); 35.72 (106.95)	0.24 (0.77)	73.83 (238.20); 80.82 (260.75)	–43 (–47)	–161 to 48 (–176 to 53)
Hospitalization due to surgical intervention	0.19 (0.89)	71.82 (312.89); 78.62 (342.51)	0.26 (1.31)	97.41 (446.49); 106.63 (488.76)	–29 (–32)	–292 to 204 (–320 to 223)
Telephone calls	0.51 (0.76)	9.18 (11.31); 10.05 (12.38)	5.27 (1.35)	85.03 (20.31); 93.08 (22.23)	–76 (–83)	–85 to –66 (–93 to –72)
Cost of TECCU controls	33.15 (5.78)	68.96 (12.07); 75.49 (13.21)	N/A[b]	N/A	67 (73)	59 to 75 (65 to 82)
Equipment costs						
TECCU rental costs	6 months	23.94 (0); 26.21 (0)	N/A	N/A	24 (26)	22 to 27 (24 to 30)
Productivity costs						
Work absenteeism (sick leave)	15.09 (21.19)	213.97 (300.51); 234.23 (328.96)	13.52 (12.51)	206.74 (194.61); 226.31 (213.03)	30 (33)	–106 to 185 (–116 to 203)
Work presenteeism (due to disease activity)	11.88 (17.26)	168.42 (224.04); 184.36 (245.25)	11.58 (20.53)	177.09 (317.12); 193.85 (347.14)	–40 (–44)	–186 to 102 (–204 to 112)
Absenteeism for medical visits	13.20 (4.93)	187.22 (62.30); 204.94 (68.20)	14.03 (6.32)	214.52 (91.72); 234.83 (100.40)	–27 (–30)	–93 to –36 (–102 to –39)
Absenteeism for telephone calls	0.12 (0.15)	1.71 (2.24); 1.87 (2.45)	1.06 (0.31)	15.69 (4.59); 17.18 (5.02)	–14 (–15)	–16 to –12 (–18 to –13)
Leisure time used in TECCU contacts	1.79 (0.31)	25.37 (4.42); 27.77 (4.84)	N/A	N/A	25 (27)	22 to 27 (24 to 30)
Total productivity costs per patient (rounded values)	N/A	407 (339); 445 (371)	N/A	466 (397); 510 (435)	–51 (–56)	–292 to 199 (–320 to 218)
Total costs per patient (rounded values)	N/A	807 (623); 883 (682)	N/A	992 (804); 1086 (880)	–135 (–148)	–579 to 290 (–634 to 317)

[a] TECCU: Telemonitoring of Crohn’s Disease and Ulcerative Colitis.

[b] N/A: not applicable.

Table 6. Utilities per patient in the group receiving remote monitoring (G_TECCU) vs nurse-assisted telephone care (G_NT).

Utilities	TECCU[a] (n=21)	Telephone care (n=21)	Bootstrapped difference in effects (rounded values)	
			Median	95% CI
Effects				
EQ-5D[b]score week 24, mean (SD)	0.90 (0.19)	0.89 (0.16)	0.01	–0.10 to 0.12
EQ-5D score improvement week 0-24, mean (SD)	0.09 (0.28)	0.08 (0.18)	–0.01	–0.15 to 0.13
Weeks in remission, mean (SD)	17.89 (7.03)	17.24 (8.38)	1.62	–3.53 to 6.67
Calprotectin at week 24 (µg/g), median (interquartile range)	126 (47.24)	168 (49.38)	–12.31	–167 to 133
Remission at week 24, n (%)	17 (80.95)	14 (66.67)	0.17	–0.01 to 0.48
Improvement in remission week 0-24, n (%)	7 (33.33)	6 (28.57)	0.06	–0.16 to 0.43
ICER[c] remission response (rounded values), € (US $)	N/A[d]	N/A	–2250 (–2463)	–15,363 to 11,086 (–16,817 to 12,135)
ICER quality-adjusted life-years (rounded values), € (US $)	N/A	N/A	5761 (6306)	–36,109 to 47,231 (–39,527 to 51,702)

[a] TECCU: Telemonitoring of Crohn's Disease and Ulcerative Colitis.

[b] EQ-5D: EuroQol 5 dimensions questionnaire.

[c] ICER: incremental cost-effectiveness ratio.

[d] N/A: not applicable.

Table 7. Costs per patient in the group receiving nurse-assisted telephone care (G_NT) vs standard care (G_control).

Costs	Telephone care (n=21), mean (SD)		Controls (n=21), mean (SD)		Bootstrapped difference in costs (rounded values), € (US $)	
	Number of units	Cost per patient, €; US $	Number of units	Cost per patient, €; US $	Median	95% CI
Health care costs						
Emergency room visits and nonscheduled outpatient visits	0.57 (1.17)	108.15 (220.18); 118.39 (241.02)	0.19 (0.87)	36.18 (165.03); 39.60 (180.65)	72 (79)	–45 to 180 (–49 to 197)
Outpatient visits	4.25 (1.92)	170.03 (76.68); 186.13 (83.94)	6.30 (0.57)	252.13 (22.91); 276.00 (25.08)	–84 (–92)	–119 to –39 (–130 to –43)
Hospitalization	0.24 (0.77)	73.83 (238.20); 80.82 (260.75)	0.05 (0.22)	14.79 (67.65); 16.19 (74.05)	59 (65)	–29 to 177 (–32 to 194)
Hospitalization due to surgical intervention	0.26 (1.31)	97.41 (446.49); 106.63 (488.76)	0.17 (0.87)	65.10 (297.74); 71.26 (325.93)	32 (35)	–195 to 292 (–213 to 320)
Telephone calls	5.27 (1.35)	85.03 (20.31); 93.08 (22.23)	2.04 (1.62)	32.05 (21.41); 35.08 (23.44)	52 (57)	40 to 64 (44 to 70)
Cost of TECCU[a] controls	N/A[b]	N/A	N/A	N/A	N/A	N/A
Equipment costs						
TECCU rental costs	N/A	N/A	N/A	N/A	N/A	N/A
Productivity costs						
Work absenteeism (sick leave)	13.52 (12.51)	206.74 (194.61); 226.31 (213.03)	25.77 (27.25)	384.22 (418.00); 420.59 (457.57)	–72 (79)	–247 to 79 (–270 to 86)
Work presenteeism (due to disease activity)	11.58 (20.53)	177.09 (317.12); 193.85 (347.14)	24.35 (34.53)	363.06 (549.48); 397.43 (601.50)	–49 (–54)	–286 to 143 (–313 to 157)
Absenteeism for medical visits	14.03 (6.32)	214.52 (91.72); 234.83 (100.40)	20.79 (4.79)	309.98 (113.41); 339.32 (124.15)	–98 (–107)	–173 to –18 (–189 to –20)
Absenteeism for telephone calls	1.06 (0.31)	15.69 (4.59); 17.18 (5.02)	0.36 (0.27)	5.39 (3.94); 5.90 (4.31)	10 (11)	8 to 13 (9 to 14)
Leisure time used in TECCU contacts	N/A	N/A	N/A	N/A	N/A	N/A
Total productivity costs per patient (rounded values)	N/A	466 (397); 510 (435)	N/A	678 (686); 742 (751)	–209 (–229)	–570 to 125 (–624 to 137)
Total costs per patient (rounded values)	N/A	992 (804); 1086 (880)	N/A	1066 (678); 1167 (742)	–77 (–84)	–524 to 378 (–574 to 414)

[a] TECCU: Telemonitoring of Crohn's Disease and Ulcerative Colitis.

[b] N/A: not applicable.

Table 8. Utilities per patient in the group receiving nurse-assisted telephone care (G_NT) vs standard care (G_control).

Utilities	Telephone care (n=21)	Controls (n=21)	Bootstrapped difference in effects (rounded values)	
			Median	95% CI
Effects				
EQ-5D[a] score week 24, mean (SD)	0.89 (0.16)	0.93 (0.15)	–0.04	–0.13 to 0.05
EQ-5D score improvement week 0-24, mean (SD)	0.08 (0.18)	0.10 (0.19)	–0.02	–0.11 to 0.09
Weeks in remission, mean (SD)	17.24 (8.38)	14.27 (8.13)	2.28	–2.38 to 7.33
Calprotectin at week 24 (µg/g), median (interquartile range)	168 (49.38)	230 (48.67)	–91	–505 to 117
Remission at week 24, n (%)	14 (66.67)	15 (71.43)	–0.05	–0.33 to 0.24
Improvement in remission week 0-24, n (%)	6 (28.57)	3 (14.28)	0.14	–0.19 to 0.48
ICER[b] remission response (rounded values), € (US $)	N/A[c]	N/A	–538 (–589)	–6475 to 5303 (–7088 to 5805)
ICER quality-adjusted life-years (rounded values), € (US $)	N/A	N/A	3316 (3630)	–58,652 to 35,482 (–64,204 to 38,841)

[a] EQ-5D: EuroQol 5 dimensions questionnaire.

[b] ICER: incremental cost-effectiveness ratio.

[c] N/A: not applicable.

G_TECCU Versus G_control

After the 24-week follow-up, the total mean cost per G_TECCU patient was €807 (US $883) as opposed to €1066 (US $1167) for the control patients, representing a median cost reduction from a societal perspective of €211 (US $231) per patient (95% CI €–600 to 180 per patient; US $–657 to 197 per patient). The main drivers of health care costs were the reduction of €96 (US $105) per patient in outpatient visits (95% CI €–126 to –65; US $–138 to –71) and that of €24 (US $26) per patient in telephone consultations (95% CI €–33 to –14; US $–36 to –15). Productivity costs were reduced in the G_TECCU intervention group by €260 (US $285) per patient (95% CI €–600 to 71; US $–657 to 78). Reduced absenteeism due to outpatient visits (€–125 per patient, 95% CI €–189 to –58; US $–137 per patient, 95% CI US $–207 to –63) and telephone consultations (€–3.81 per patient, 95% CI €–5.82 to –2; US $–4.17 per patient, 95% CI US $–6.37 to –2.19) represented a significant cost saving (see Table 3).

G_TECCU Versus G_NT

At 24 weeks, the total mean cost per patient in the G_NT was €992 (US $1086) as opposed to €807 (US $883) for G_TECCU patients. Thus, there was a median cost reduction of €135 (US $148) per patient in the G_TECCU after 24 weeks (95% CI €–579 to 290 per patient; US $–634 to 317 per patient), which was associated with several health care and work productivity factors. The reduction of costs was significant in terms of telephone consultations (€–76 per patient, 95% CI €–85 to –66; US $–83 per patient, 95% CI US $–93 to –72) and the associated absenteeism to attend to those telephone calls (€–14 per patient, 95% CI €–16 to –12; US $–15 per patient, 95% CI US $–18 to –13; see Table 5).

G_NT Versus G_control

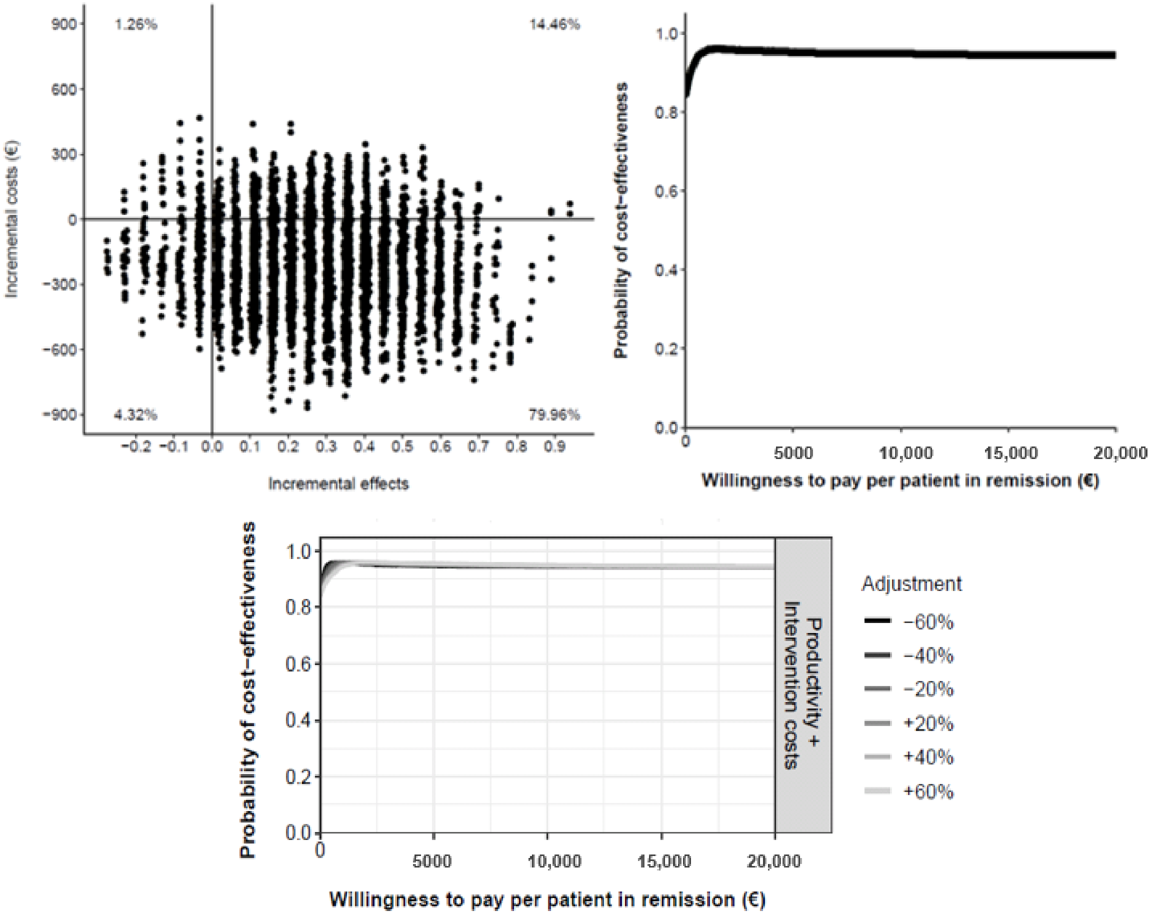
The mean costs per patient in the G_control and G_NT represented a median cost reduction of €77 (US $84) per patient in G_NT after 24 weeks (95% CI €–524 to 378 per patient; US $–574 to 414 per patient). This costs saving was mainly due to the reduction of €84 (US $92) per patient linked to fewer outpatient visits by patients in G_NT (95% CI €–119 to –39; US $–130 to –43) and a significant reduction of €98 (US $107) per patient for absenteeism due to medical visits (95% CI €–173 to –18; US $–189 to 20; see Table 7).

Cost-Effectiveness Analysis

G_TECCU Versus G_control

We obtained the mean ICER for TECCU care compared with standard care by dividing the incremental costs by the differences in the increment for the percentage of patients in remission. For TECCU, this was €–211/0.191=€–1105 (US $1210) for 1 additional patient in remission in G_TECCU, 24 weeks after inclusion. The bootstrapping procedure gave an estimated median ICER of €–1005 (95% CI €–13,518 to 3137; US $1100, 95% CI US $–14,798 to 3434).

Figure 3. Cost-effectiveness plane (top left) and cost-effectiveness acceptability curve (top right) comparing the Telemonitoring of Crohn’s Disease and Ulcerative Colitis (TECCU) intervention versus standard care. Bottom: impact of different cost values on the cost-effectiveness acceptability curves.



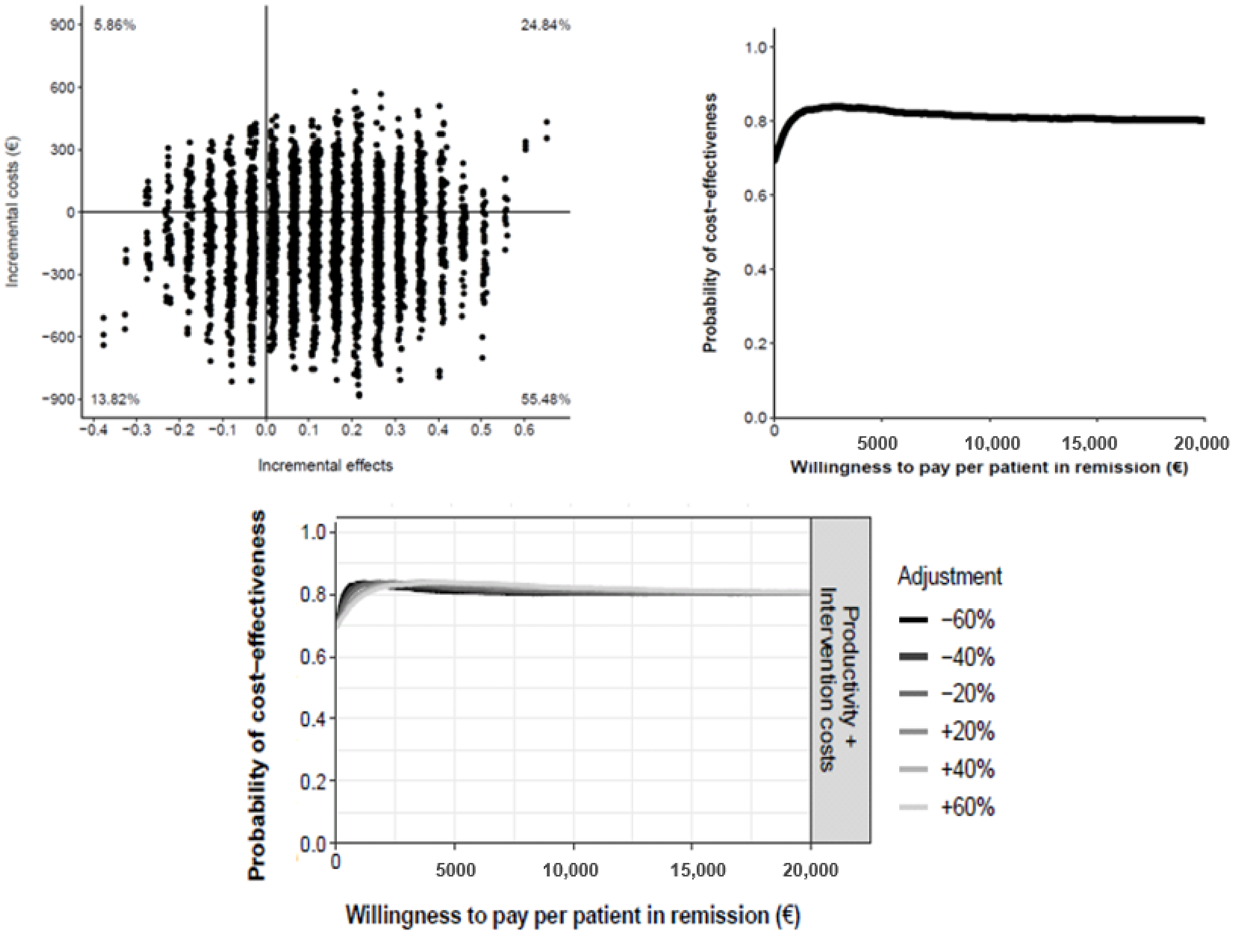
In the cost-effectiveness plane (Figure 3), we represented all the estimated ICERs with dots, and there was a 79.96% probability that TECCU improved the proportion of patients in remission at a lower societal cost than for the control patients (dominant quadrant). In an additional 14.46% of simulations, TECCU produced stronger effects but with higher costs than for the standard caare. The probability that TECCU was cost-effective in comparison with standard care at a WTP of €20,000 (US $21,893) per additional patient in remission was 95%, and the probability that TECCU was cost saving at a WTP of €0 was 84%. These percentages remained stable even after adjusting both the health care (including equipment costs) and indirect costs over a range of ±60%, and the different cost-effectiveness acceptability curves calculated were very similar (Figure 3).

G_TECCU Versus G_NT

Considering that TECCU saved €135 per patient relative to G_NT and the difference in the efficacy on disease activity between these 2 interventions was 0.048, the mean ICER of G_TECCU relative to G_NT was € – 2812 (€–135/0.048) after 24 weeks. This means that when TECCU achieved 1 additional patient in remission compared with G_NT, the cost savings was €2812 (US $3078). Using the bootstrapping procedure we estimated that the median ICER was €–2250 (95% CI €–15,363 to 11,086; US $–2463, 95% CI US $–16,817 to 12,135).

In the cost-effectiveness plane (Figure 4), we found that 55.48% of the dots fell in the dominant quadrant, indicating that this was the probability of TECCU being more effective than telephone care at a lower societal cost. In another 24.84% of the simulations, TECCU had a stronger effect but higher costs than telephone care. The probability that TECCU was cost-effective relative to G_NT at a WTP of €20,000 (US $21,893) per additional patient in remission was 81%, and the probability that TECCU was more cost saving at a WTP of €0 was 69%. In all sensitivity scenarios (±60%), the probability of TECCU being cost-effective at a WTP of €0 (cost saving) was 69%, and this probability increased to a stable 80% to 81% at a WTP of €20,000 or more (Figure 4).

Figure 4. Cost-effectiveness plane (top left) and cost-effectiveness acceptability curve (top right) comparing Telemonitoring of Crohn's Disease and Ulcerative Colitis (TECCU) versus telephone care. Bottom: impact of different cost values on the cost-effectiveness acceptability curves.

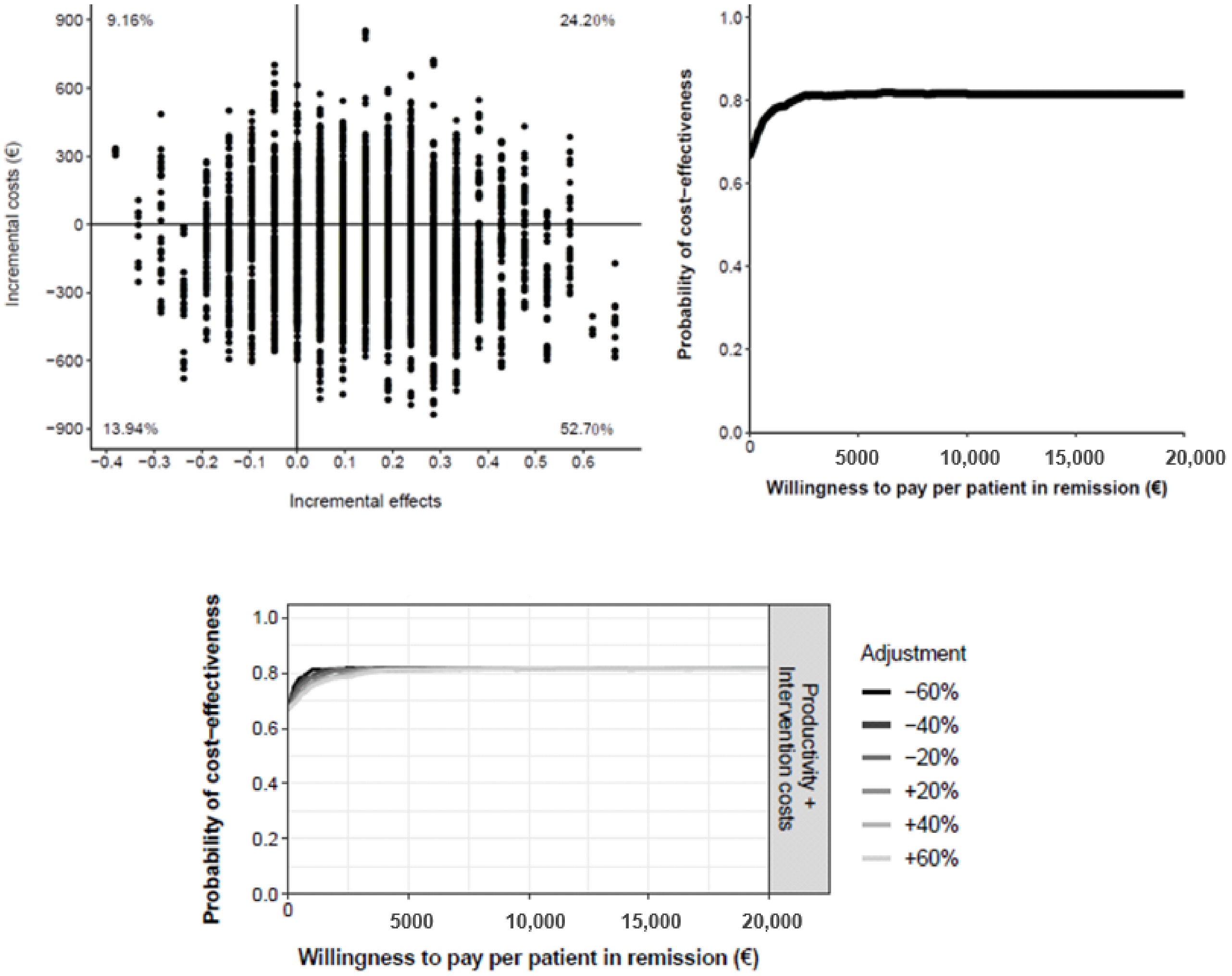


G_NT Versus G_control

Comparing the mean ICER of telephone care with that of standard care, by dividing the incremental costs by the differences in the increase of the percentage of patients in remission, it showed a cost of €–77/0.143= €–538 (US $–589) for 1 additional patient in remission 24 weeks after inclusion. Using the bootstrapping procedure, we estimated that the median ICER was €–538 (95% CI €–6475 to 5303; US $–589, 95% CI US $–7088 to 5805).

In the cost-effectiveness plane (Figure 5), 52.70% fell into the dominant quadrant, indicating that there was a 52.70% probability that nurse-assisted telephone care was more effective than standard care at a lower societal cost. In 24.20% of the simulations, telephone care was more effective than standard care at a higher societal cost. The probability that telephone care was cost saving relative to standard care was 67% at a WTP of €0 per additional patient in remission, whereas at a WTP of €20,000 (US $21,893) the probability of telephone care being cost-effective was 81%. After modifying the costs over a range of ±60%, the probability of telephone care being more cost saving remained unchanged at 67% at a WTP of €0. The probability that telephone care was more cost-effective at a WTP of €20,000 or more was stable at 81% (Figure 5).

Figure 5. Cost-effectiveness plane (top left) and cost-effectiveness acceptability curve (top right) comparing telephone care versus standard care. Bottom: impact of different cost values on the cost-effectiveness acceptability curves.



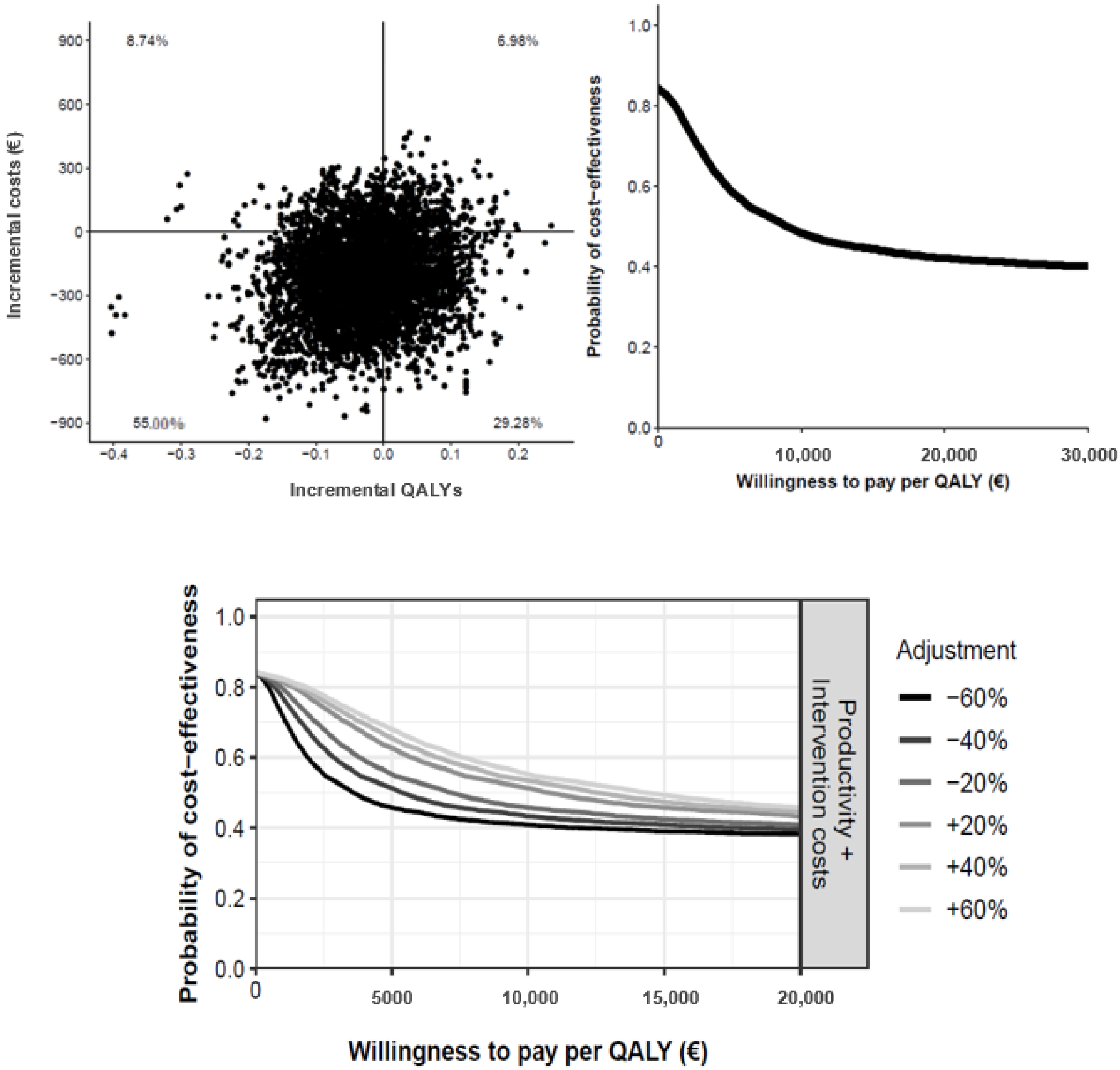
Cost-Utility Analysis

G_TECCU Versus G_control

The baseline EQ-5D scores were slightly higher in G_TECCU and G_NT, and thus the improvement over the 24-week follow-up was not greater than that in G_control. Using the bootstrapping procedure, the median ICER for 1 additional QALY was estimated to be €9078 (95% CI €–56,547 to 44,628; US $9937, 95% CI US $–61,900 to 48,853), which means that €9078 was saved in G-TECCU for each extra QALY gained in the G_control.

There was a 29.28% probability that the TECCU intervention was associated with a higher gain in QALYs at a lower cost (Figure 6). In another 55.00% of the simulations, TECCU was still less expensive than standard care, although with a weaker improvement in QALYs gained. Thus, considering the statistical uncertainty of the costs and QALYs calculated, about 84% of the bootstrapped ICERs were associated with cost savings, which is the probability of TECCU being cost-effective relative to standard care at a WTP of €0. However, the cost-effectiveness acceptability curve was a decreasing function of WTP because approximately 64% of simulations did not involve health gains [43]. Thus, the probability of TECCU being cost-effective at a WTP of €20,000 (US $21,893) fell to 42%, and this probability remained stable (40% to 45%) after modifying the costs in the sensitivity analysis (Figure 6).

Figure 6. Cost-effectiveness plane (top left) and cost-effectiveness acceptability curve (top right) comparing the effect on quality-adjusted life-years (QALYs) of Telemonitoring of Crohn’s Disease and Ulcerative Colitis (TECCU) versus standard care. Bottom: impact of different cost values on the cost-utility acceptability curves.

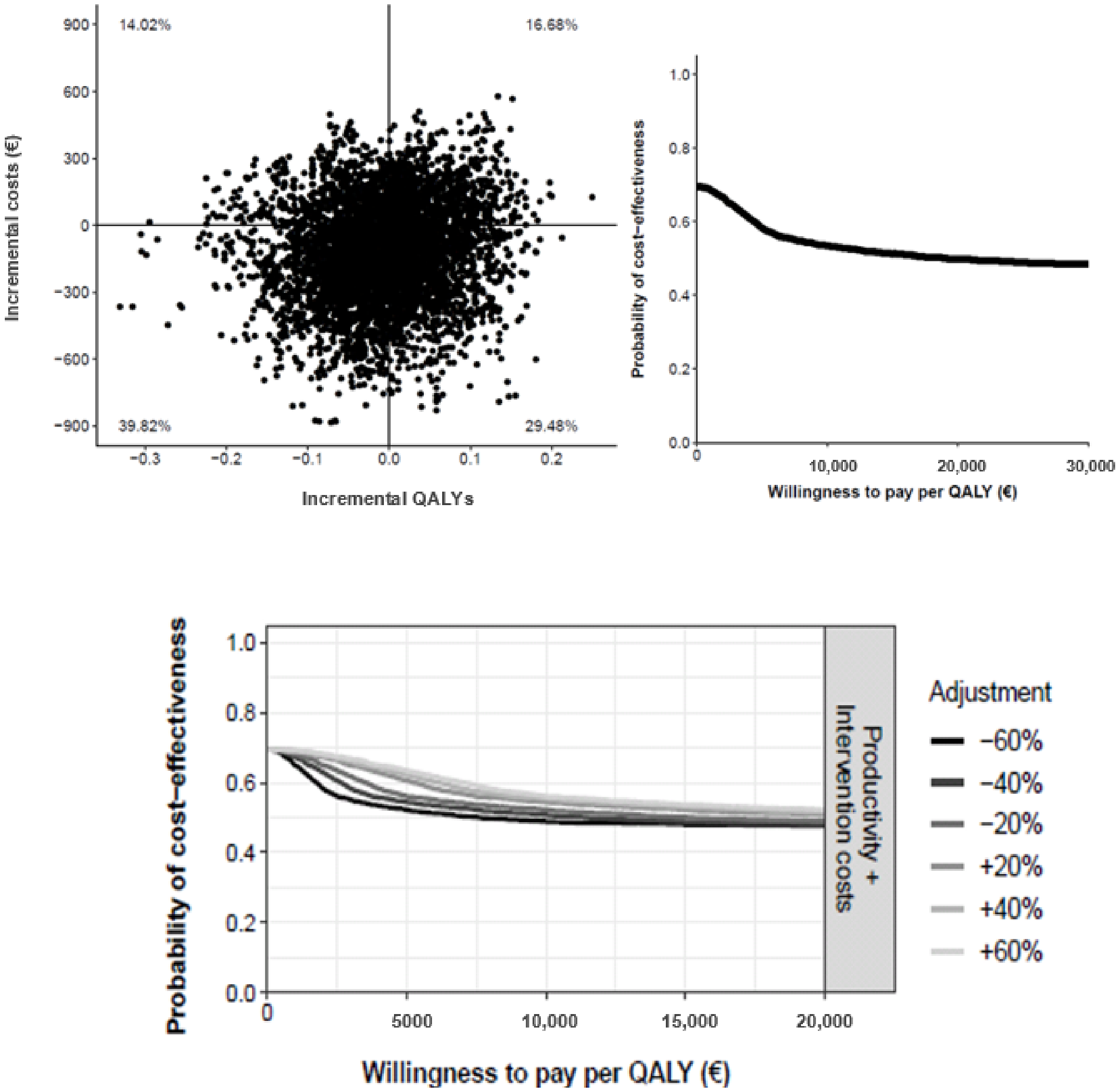


G_TECCU Versus G_NT

The median cost-utility ratio per QALY gained with TECCU relative to telephone care was €5761 (95% CI €–36,109 to 47,231; US $6306, 95% CI US $–39,527 to 51,702), meaning that €5761 was saved in G_TECCU for each extra QALY gained in the G_NT.

The probability that TECCU led to a higher QALY gain at lower costs was 29.48% (Figure 7), while in another 39.82% of the simulations TECCU was still associated with lower costs but with a weaker improvement in QALYs gained. Considering statistical uncertainty, the cost-effectiveness acceptability curve suggested a 69% probability that TECCU was more cost saving at a WTP of €0 than telephone care in terms of QALYs. However, as approximately 54% of the simulations did not involve health gains, the probability of TECCU being cost-effective at a WTP of €20,000 (US $21,893) fell to 50%. In the different cost scenarios calculated in the sensitivity analysis, the range of this probability was tight (48% to 51%; Figure 7).

Figure 7. Cost-effectiveness plane (top left) and cost-effectiveness acceptability curve (top right) comparing the effect on quality-adjusted life-years (QALYs) of Telemonitoring of Crohn’s Disease and Ulcerative Colitis (TECCU) versus telephone care. Bottom: impact of different cost values on the cost-utility acceptability curves.

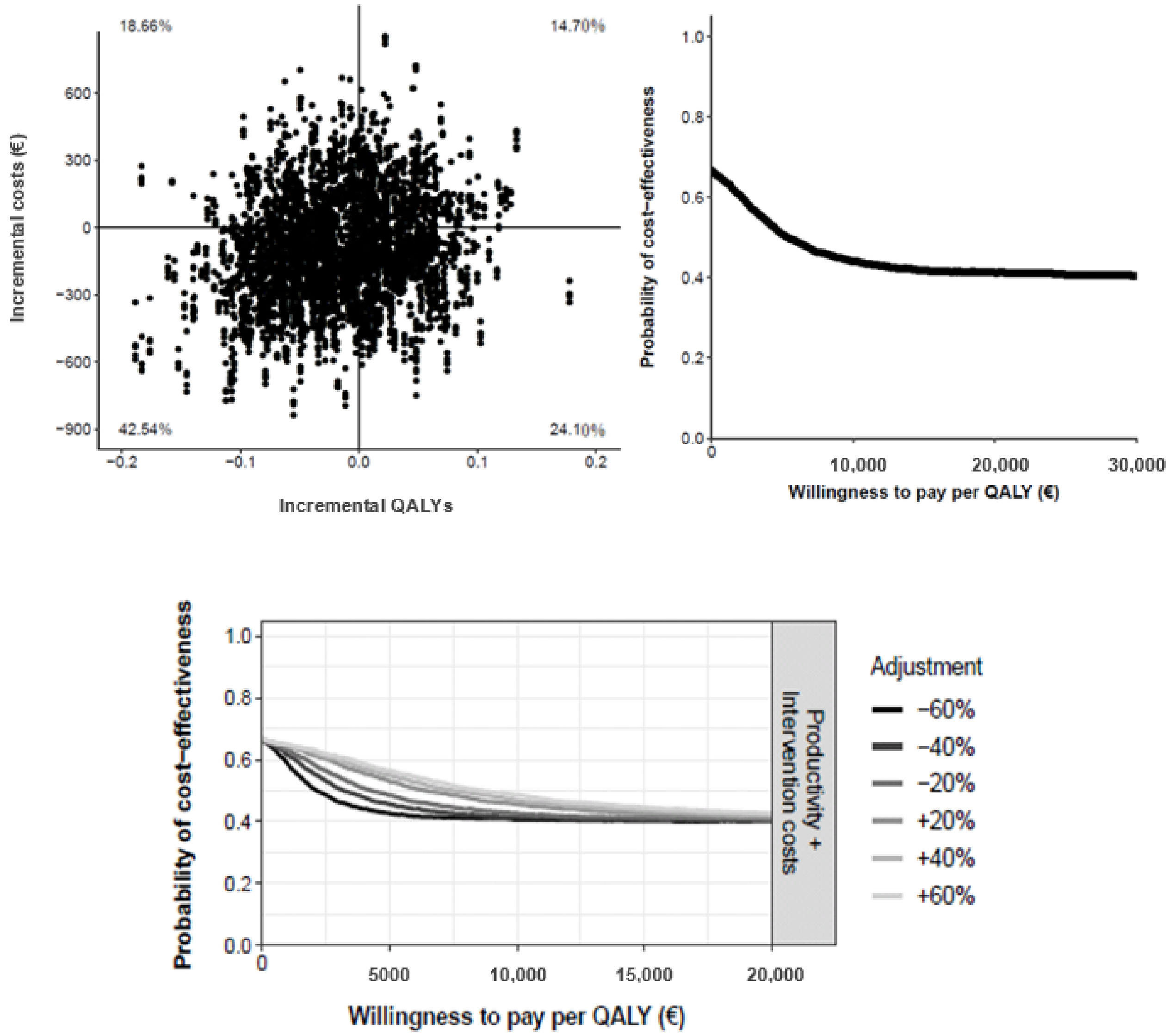


G_NT Versus G_control

The median incremental societal costs per QALY gained in the G_NT relative to the controls was €3316 (95% CI €–58,652 to 35,482; US $3630, 95% CI US $–64,204 to 38,841), which means that telephone care saved €3316 for each additional QALY gained in the G_control.

There was a 24.10% probability that nurse-assisted telephone care was associated with a higher QALY gain at lower costs (Figure 8). In a further 42.54% of simulations, telephone care was less expensive than standard care but with a lower improvement in QALYs gained. With the combination of these percentages, approximately 67% of simulations estimated that telephone care was less costly than standard care (cost-effective at a WTP €0). However, as about 61% of simulations did not involve health gains, the probability that telephone care was cost-effective fell to 41% at a WTP of €20,000 (US $21,893). Considering scenarios in which health care and indirect costs decrease or increase over 60%, at a WTP of €20,000 the probability of telephone care being cost-effective fell to 40% and 42%, respectively (Figure 8).

Figure 8. Cost-effectiveness plane (top left) and cost-effectiveness acceptability curve (top right) comparing the effect on quality-adjusted life-years (QALYs) of telephone care versus standard care. Bottom: impact of different cost values on the cost-utility acceptability curves.



Discussion

Principal Findings

To the best of our knowledge, this is the first study to evaluate the cost-effectiveness and cost-utility of a Web telemonitoring platform for IBD patients from a societal perspective. We used a standard methodology [21] that examined both health care and non–health care–related costs against the effect of the TECCU Web platform on health outcomes, and to compare this with telephone care and standard care, the main strategies currently used for the follow-up of IBD patients. This economic evaluation included topics and outcome measures recommended in the model for assessment of telemedicine [44]. Other approaches such as the monitoring and assessment framework for the European Innovation Partnership are available to estimate the health and economic outcomes of eHealth interventions [45]. In any case, as the 24-week period after initiation of immunosuppressants or biologic agents may be different from the maintenance therapy period, we avoided performing economic estimations with longer time horizons.

In this regard, TECCU had a stronger effect on disease activity, with associated savings of €1005 (US $1100) and €2250 (US $2463) for each additional patient in remission when compared with standard care and telephone care, respectively. Conversely, while quality-of-life scores improved in all 3 groups, neither TECCU nor telephone care produced a stronger improvement than standard care, possibly due to the higher baseline scores of G_TECCU and G_NT patients, and to the relatively small sample size, hindering the possibility of detecting statistical differences after 24 weeks. However, TECCU and telephone care were associated with an 84% and 67% probability, respectively, of reducing costs per additional QALY relative to the controls.

The WTP threshold that can be considered acceptable to consider an intervention cost-effective is not clear, and in Spain it was recently estimated to lie between €22,000 (US $24,082) and €25,000 (US $27,367) per QALY [46]. Far from these values, the probability of TECCU being more effective than standard care was 80%, even at a lower societal cost. Furthermore, the probability of being cost-effective was 95%

at a WTP of €20,000 (US $21,893) per additional patient in remission. The differences with respect to telephone care were lower, although comparing these interventions there was still an 81% likelihood that telemonitoring was cost-effective at a WTP of €20,000 per additional patient in remission. Moreover, all these probabilities were stable after considering alternative costing scenarios.

While we evaluated the cost-effectiveness and cost-utility of the TECCU Web platform, economics data regarding eHealth interventions in IBD have been scarce thus far. This represents a barrier to their implementation in real life because they are associated with a good cost-effectiveness profile in some chronic pathologies such as cardiovascular diseases [47], but not in others [48]. Considering IBD, 2 clinical trials previously assessed a Web-based approach to guide patient self-treatment, demonstrating a significant reduction in direct costs by replacing outpatient visits with distance care [12,49]. A remote management program developed in the United Kingdom for patients with stable IBD also estimated that virtual clinics could potentially save £119,000 (US $143,072) per year [14]. However, these savings analyses did not consider indirect costs, and a reduction in travel time was only described in an uncontrolled pilot trial of patients who used telehealth [50]. It is surprising that even if the use of telemonitoring for IBD is associated with a shift from in-person visits toward remote encounters [12,13,16], previous studies did not include costs associated with the purchase of the necessary equipment and with remote contacts.

The use of TECCU was associated with a median saving of €211 (US $231) per patient relative to standard care. Considering health care costs, TECCU saved €94 (US $103) per patient in outpatient visits and €24 (US $26) per patient in telephone calls, these savings in outpatient visits over 6 months being very similar to the €189 (US $207) per patient per year reported previously in a clinical trial with an eHealth program [12]. However, the median expense of €71 (US $78) per patient in telemonitoring contacts calculated here is comparable with these savings. Thus, the median of €211 (US $231) saved per patient from a societal perspective is mainly associated with the improvement in work productivity and not with the benefits in health care costs, as reported in previous studies where expenditure in telemonitoring was not considered. By contrast, when compared with another distance follow-up method such as telephone care, TECCU cost savings were related to both health care and non–health care costs.

Strengths and Limitations

The strengths of this study include the use of validated clinical indexes to assess the effects of the 3 follow-up methods on disease activity and quality of life. We considered national and regional official prices to calculate costs and we chose the societal perspective to perform the economic evaluation in this study, including costs associated with health care and investment in equipment, and costs related to patients' productivity at work and social activity impairment in all 3 groups. Additionally, to characterize the uncertainty of the costs and utilities calculated, we used nonparametric bootstrap estimations, as well as a sensitivity analysis to examine whether the ICERs changed in alternative costing scenarios. Finally, to better reproduce the costs calculated in a real-world setting, the follow-up schedule for all 3 arms in our clinical trial was designed according to the standard clinical practice in our center and based on national and European guidelines, as published elsewhere [20].

This study had a series of limitations. First, quality of life was a secondary outcome, and we only measured it at baseline and at the end of the study to improve adherence to the follow-up schedule. This limitation, associated with the reduced sample size and the higher baseline scores in G_TECCU and G_NT, could hinder the possibility to detect significant differences after 24 weeks. Second, even though we recruited patients consecutively in a referral center, the sample size was relatively small, mainly because we only included patients with complex IBD during the initiation of immunosuppressants or biologic agents, but not those on maintenance therapy. Although we used validated questionnaires to measure the effects and the official rates for Spain, the study of this specific population may compromise the generalization of cost data to other settings. Third, the trial did not consider travel costs related to in-person medical visits, but in any case this would underestimate the cost savings associated with TECCU and telephone care. Fourth, the economic evaluation was limited to the 24 weeks of the study period, as we evaluated patients with complex IBD at the initiation of treatment with immunosuppresants or biologic agents, or both. In this sense, it is possible that costs and effects would change after longer follow-up periods with maintenance therapy, and further studies considering longer time horizons will be necessary.

Conclusion

There is a high probability that the use of the TECCU Web platform for telemonitoring patients with complex IBD produces a greater improvement in disease activity at a lower societal cost, compared with both standard care and telephone care. Considering the increasing burden and costs of managing IBD worldwide, as well as the lack of economic data related to eHealth interventions, our results provide important information regarding the cost-effectiveness of Web telemonitoring for IBD. The use of systems such as TECCU could be a real option to help reorganize the structure of national health systems in the future. However, further studies are still necessary to evaluate the impact of eHealth on quality of life and its cost-effectiveness in larger sample sizes and over longer periods.

Acknowledgments

This work was supported by grants from the Instituto de Salud Carlos III-Fondo de Investigaciones Sanitarias (FIS PI12/00277), cofunded by FEDER (Fondo Europeo de Desarrollo Regional).

We would like to thank the following individuals and institutions for their collaboration in this study: the patients and their families; Ms Diana Muñoz and Margarita Pascual, nurses at the Inflammatory Bowel Disease Unit and Telemedicine Unit of La Fe University Hospital, for the data and safety monitoring; and Ms Victoria Fornes from the Biostatistics Department of Tau Analytics for her assistance with data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH checklist V1.6.1.
[PDF File (Adobe PDF File), 511KB - jmir_v21i9e15505_app1.pdf]

References

1. Burisch J, Jess T, Martinato M, Lakatos PL, ECCO -EpiCom. The burden of inflammatory bowel disease in Europe. J Crohns Colitis 2013 May;7(4):322-337 [FREE Full text] [doi: 10.1016/j.crohns.2013.01.010] [Medline: 23395397]
2. Bähler C, Vavricka SR, Schoepfer AM, Brüngger B, Reich O. Trends in prevalence, mortality, health care utilization and health care costs of Swiss IBD patients: a claims data based study of the years 2010, 2012 and 2014. BMC Gastroenterol 2017 Dec 02;17(1):138 [FREE Full text] [doi: 10.1186/s12876-017-0681-y] [Medline: 29197335]
3. Peery AF, Crockett SD, Murphy CC, Lund JL, Dellon ES, Williams JL, et al. Burden and cost of gastrointestinal, liver, and pancreatic diseases in the United States: update 2018. Gastroenterology 2019 Dec;156(1):254-272.e11. [doi: 10.1053/j.gastro.2018.08.063] [Medline: 30315778]
4. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]
5. Aguas PM, Del Hoyo J, Bebia P, Faubel R, Barrios A, Bastida G, et al. Telemedicine in inflammatory bowel disease: opportunities and approaches. Inflamm Bowel Dis 2015 Feb;21(2):392-399. [doi: 10.1097/MIB.0000000000000241] [Medline: 25437818]
6. Høivik ML, Moum B, Solberg IC, Henriksen M, Cvancarova M, Bernklev T, IBSEN Group. Work disability in inflammatory bowel disease patients 10 years after disease onset: results from the IBSEN Study. Gut 2013 Mar;62(3):368-375. [doi: 10.1136/gutjnl-2012-302311] [Medline: 22717453]
7. Hoivik ML, Moum B, Solberg IC, Cvancarova M, Hoie O, Vatn MH, IBSEN Study Group. Health-related quality of life in patients with ulcerative colitis after a 10-year disease course: results from the IBSEN study. Inflamm Bowel Dis 2012 Aug;18(8):1540-1549. [doi: 10.1002/ibd.21863] [Medline: 21936030]
8. Ondategui S. Informe sobre la enfermedad inflamatoria intestinal: situación actual y retos asistenciales en los últimos años. London, UK: Ernst & Young Global Limited; 2016. URL: https://www.ey.com/es/es/home/ey-informe-sobre-la-enfermedad-inflamatoria-intestinal [accessed 2019-08-31]
9. Carlsen K, Jakobsen C, Houen G, Kallemose T, Paerregaard A, Riis LB, et al. Self-managed eHealth disease monitoring in children and adolescents with inflammatory bowel disease: a randomized controlled trial. Inflamm Bowel Dis 2017 Dec;23(3):357-365. [doi: 10.1097/MIB.0000000000001026] [Medline: 28221247]
10. Carlsen K, Houen G, Jakobsen C, Kallemose T, Paerregaard A, Riis LB, et al. Individualized infliximab treatment guided by patient-managed eHealth in children and adolescents with inflammatory bowel disease. Inflamm Bowel Dis 2017 Dec;23(9):1473-1482. [doi: 10.1097/MIB.0000000000001170] [Medline: 28617758]
11. Heida A, Dijkstra A, Muller KA, Rossen JW, Kindermann A, Kokke F, et al. Efficacy of home telemonitoring versus conventional follow-up: a randomized controlled trial among teenagers with inflammatory bowel disease. J Crohns Colitis 2018 Mar 28;12(4):432-441. [doi: 10.1093/ecco-jcc/jjx169] [Medline: 29228230]
12. Elkjaer M, Shuhaibar M, Burisch J, Bailey Y, Scherfig H, Laugesen B, et al. E-health empowers patients with ulcerative colitis: a randomised controlled trial of the web-guided 'Constant-care' approach. Gut 2010 Dec;59(12):1652-1661. [doi: 10.1136/gut.2010.220160] [Medline: 21071584]
13. de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, Becx MC, Maljaars JP, Cilissen M, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. Lancet 2017 Sep 02;390(10098):959-968. [doi: 10.1016/S0140-6736(17)31327-2] [Medline: 28716313]
14. Johnson MW, Lithgo K, Price T. OC-080 Ibd-Sshamp (Supported, Self help and Management Programme); UK'S first Internet based remote management system for managing stable IBD. Gut 2013 Jun 04;62(Suppl 1):A34.3-A3A35. [doi: 10.1136/gutjnl-2013-304907.079]
15. Cross RK, Cheevers N, Rustgi A, Langenberg P, Finkelstein J. Randomized, controlled trial of home telemanagement in patients with ulcerative colitis (UC HAT). Inflamm Bowel Dis 2012 Jun;18(6):1018-1025 [FREE Full text] [doi: 10.1002/ibd.21795] [Medline: 21688350]

16. Cross RK, Langenberg P, Regueiro M, Schwartz DA, Tracy JK, Collins JF, et al. A randomized controlled trial of telemedicine for patients with inflammatory bowel disease (TELE-IBD). Am J Gastroenterol 2019 Jan 11. [doi: 10.1038/s41395-018-0272-8] [Medline: 30410041]
17. Jackson BD, Gray K, Knowles SR, De Cruz P. EHealth technologies in inflammatory bowel disease: a systematic review. J Crohns Colitis 2016 Sep;10(9):1103-1121. [doi: 10.1093/ecco-jcc/jjw059] [Medline: 26928960]
18. Miller EA. Solving the disjuncture between research and practice: telehealth trends in the 21st century. Health Policy 2007 Jul;82(2):133-141. [doi: 10.1016/j.healthpol.2006.09.011] [Medline: 17046097]
19. Aguas M, Del Hoyo J, Faubel R, Muñoz D, Domínguez D, Bastida G, et al. A web-based telemanagement system for patients with complex inflammatory bowel disease: protocol for a randomized controlled clinical trial. JMIR Res Protoc 2018 Dec 21;7(12):e190 [FREE Full text] [doi: 10.2196/resprot.9639] [Medline: 30578197]
20. Del Hoyo J, Nos P, Faubel R, Muñoz D, Domínguez D, Bastida G, et al. A web-based telemanagement system for improving disease activity and quality of life in patients with complex inflammatory bowel disease: pilot randomized controlled trial. J Med Internet Res 2018 Dec 27;20(11):e11602 [FREE Full text] [doi: 10.2196/11602] [Medline: 30482739]
21. Bergmo TS. How to measure costs and benefits of eHealth interventions: an overview of methods and frameworks. J Med Internet Res 2015;17(11):e254 [FREE Full text] [doi: 10.2196/jmir.4521] [Medline: 26552360]
22. Gomollón F, Dignass A, Annese V, Tilg H, Van Assche G, Lindsay JO, ECCO. 3rd European evidence-based consensus on the diagnosis and management of Crohn's disease 2016, part 1: diagnosis and medical management. J Crohns Colitis 2017 Jan;11(1):3-25. [doi: 10.1093/ecco-jcc/jjw168] [Medline: 27660341]
23. Magro F, Gionchetti P, Eliakim R, Ardizzone S, Armuzzi A, Barreiro-de Acosta M, European Crohn's and Colitis Organisation [ECCO]. Third European evidence-based consensus on diagnosis and management of ulcerative colitis. Part 1: definitions, diagnosis, extra-intestinal manifestations, pregnancy, cancer surveillance, surgery, and ileo-anal pouch disorders. J Crohns Colitis 2017 Jun 01;11(6):649-670. [doi: 10.1093/ecco-jcc/jjx008] [Medline: 28158501]
24. Gomollón F, García-López S, Sicilia B, Gisbert JP, Hinojosa J, en representación del Grupo Espagnol de Trabajode Enfermedad de Crohn y Colitis Ulcerosa or Spanish Group for Working on Crohn's Disease and Ulcerative Colitis (GETECCU). [Therapeutic guidelines on ulcerative colitis: a GRADE methodology based effort of GETECCU]. Gastroenterol Hepatol 2013 Oct;36(8):e1-47. [Medline: 24215088]
25. Instituto Nacional de Estadística. Cálculo de variaciones del Índice de Precios de Consumo (sistema IPC base 2016) [Calculation of the Consumer Price Index percentage changes (CPI system base 2016)]. Madrid, Spain: INE; 2019. URL: http://www.ine.es/varipc/ [accessed 2019-08-29]
26. Disposiciones Generalescomunitat Valenciana. Boletín Oficial del Estado, no 35, sec 1. 2015 Feb 10. URL: https://www.boe.es/boe/dias/2015/02/10/pdfs/BOE-A-2015-1236.pdf [accessed 2019-08-29]
27. Instituto Nacional de Estadística. INEbase Salarios, ingresos, cohesión social. Madrid, Spain: INE; 2019. URL: http://www.ine.es/jaxiT3/Tabla.htm?t=10911 [accessed 2019-09-02]
28. Vergara M, Montserrat A, Casellas F, Villoria A, Suarez D, Maudsley M, et al. A new validation of the Spanish Work Productivity and Activity Impairment Questionnaire-Crohn's disease version. Value Health 2011;14(6):859-861 [FREE Full text] [doi: 10.1016/j.jval.2011.02.1179] [Medline: 21914506]
29. Instituto Nacional de Estadística. Encuesta anual de estructura salarial. Serie 2008-2016. Madrid, Spain: INE; 2019. URL: http://www.ine.es/jaxi/Tabla.htm?path=/t22/p133/cno11/serie/l0/&file=02005.px&L=0 [accessed 2019-09-02]
30. Tan SS, Bouwmans CAM, Rutten FFH, Hakkaart-van Roijen L. Update of the Dutch Manual for Costing in Economic Evaluations. Int J Technol Assess Health Care 2012 Apr;28(2):152-158. [doi: 10.1017/S0266462312000062] [Medline: 22559757]
31. Grossman M. The human capital model. In: Culyer A, Newhouse J, editors. Handbook of Health Economics. Amsterdam, Netherlands: Elsevier; 2000:347-408.
32. Juan J, Estiarte R, Colomé E, Artés M, Jiménez FJ, Alonso J. Burden of illness of Crohn's disease in Spain. Dig Liver Dis 2003 Dec;35(12):853-861. [Medline: 14703880]
33. Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. Lancet 1980 Mar 8;1(8167):514. [Medline: 6102236]
34. Walmsley RS, Ayres RC, Pounder RE, Allan RN. A simple clinical colitis activity index. Gut 1998 Jul;43(1):29-32 [FREE Full text] [Medline: 9771402]
35. Lewis JD, Chuai S, Nessel L, Lichtenstein GR, Aberra FN, Ellenberg JH. Use of the noninvasive components of the Mayo score to assess clinical response in ulcerative colitis. Inflamm Bowel Dis 2008 Dec;14(12):1660-1666 [FREE Full text] [doi: 10.1002/ibd.20520] [Medline: 18623174]
36. Badia X, Roset M, Montserrat S, Herdman M, Segura A. [The Spanish version of EuroQol: a description and its applications. European Quality of Life scale]. Med Clin (Barc) 1999;112 Suppl 1:79-85. [Medline: 10618804]
37. Badia X, Roset M, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. Med Decis Making 2001;21(1):7-16. [doi: 10.1177/0272989X0102100102] [Medline: 11206949]
38. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann 1979;2(3):197-207. [Medline: 10245370]

39. Incremental Cost-Effectiveness Ratio (ICER). Heslington, UK: York Health Economics Consortium; 2017. URL: https://www.yhec.co.uk/glossary/incremental-cost-effectiveness-ratio-icer [accessed 2019-08-29]
40. Honaker J, King G, Blackwell M. Amelia II: a program for missing data. R Package version 1.: Harvard University; 2008. URL: http://gking.harvard.edu/amelia/ [accessed 2019-08-29]
41. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
42. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986 Jan;24(1):67-74. [Medline: 3945130]
43. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions. Health Econ 2004 May;13(5):405-415. [doi: 10.1002/hec.903] [Medline: 15127421]
44. Kidholm K, Ekeland AG, Jensen LK, Rasmussen J, Pedersen CD, Bowes A, et al. A model for assessment of telemedicine applications: MAST. Int J Technol Assess Health Care 2012 Jan;28(1):44-51. [doi: 10.1017/S0266462311000638] [Medline: 22617736]
45. Boehler CEH, de Graaf G, Steuten L, Yang Y, Abadie F. Development of a web-based tool for the assessment of health and economic outcomes of the European Innovation Partnership on Active and Healthy Ageing (EIP on AHA). BMC Med Inform Decis Mak 2015;15 Suppl 3:S4 [FREE Full text] [doi: 10.1186/1472-6947-15-S3-S4] [Medline: 26391559]
46. Vallejo-Torres L, García-Lorenzo B, Serrano-Aguilar P. Estimating a cost-effectiveness threshold for the Spanish NHS. Health Econ 2018 Apr;27(4):746-761. [doi: 10.1002/hec.3633] [Medline: 29282798]
47. Jiang X, Ming W, You JH. The cost-effectiveness of digital health interventions on the management of cardiovascular diseases: systematic review. J Med Internet Res 2019 Jun 17;21(6):e13166 [FREE Full text] [doi: 10.2196/13166] [Medline: 31210136]
48. 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]
49. Pedersen N, Elkjaer M, Duricova D, Burisch J, Dobrzanski C, Andersen NN, et al. eHealth: individualisation of infliximab treatment and disease course via a self-managed web-based solution in Crohn’s disease. Aliment Pharmacol Ther 2012 Nov;36(9):840-849 [FREE Full text] [doi: 10.1111/apt.12043] [Medline: 22971016]
50. Hommel KA, Hente E, Herzer M, Ingerski LM, Denson LA. Telehealth behavioral treatment for medication nonadherence: a pilot and feasibility study. Eur J Gastroenterol Hepatol 2013 Apr;25(4):469-473. [doi: 10.1097/MEG.0b013e32835c2a1b] [Medline: 23325274]

Abbreviations

eHealth: electronic health
EQ-5D: EuroQol 5 dimensions questionnaire
G_control: group receiving standard care with in-person visits
G_NT: group receiving nurse-assisted telephone care
G_TECCU: group receiving remote monitoring
IBD: inflammatory bowel disease
ICER: incremental cost-effectiveness ratio
QALY: quality-adjusted life-year
TECCU: Telemonitorización de la Enfermedad de Crohn y Colitis Ulcerosa (Telemonitoring of Crohn’s Disease and Ulcerative Colitis
WTP: willingness to pay

Edited by G Eysenbach; submitted 20.07.19; peer-reviewed by V Traver Salcedo, P Haubruck, D Koutsouris, A Anastasiou, S Pitoglou; comments to author 07.08.19; revised version received 19.08.19; accepted 21.08.19; published 13.09.19.

Please cite as:
Del Hoyo J, Nos P, Bastida G, Faubel R, Muñoz D, Garrido-Marín A, Valero-Pérez E, Bejar-Serrano S, Aguas M
Telemonitoring of Crohn’s Disease and Ulcerative Colitis (TECCU): Cost-Effectiveness Analysis
J Med Internet Res 2019;21(9):e15505
URL: http://www.jmir.org/2019/9/e15505/
doi:10.2196/15505
PMID:31538948

Original Paper

Fifteen Years' Use of Patient-Reported Outcome Measures at the Group and Patient Levels: Trend Analysis

Niels Henrik I Hjollund[1,2], MD, PhD

[1]AmbuFlex/WestChronic, Occupational Medicine, University Research Clinic, Aarhus University, Herning, Denmark

[2]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

Corresponding Author:
Niels Henrik I Hjollund, MD, PhD
AmbuFlex/WestChronic, Occupational Medicine
University Research Clinic
Aarhus University
Gl Landevej 61
Herning, 7400
Denmark
Phone: 45 78433514
Email: niehjo@rm.dk

Abstract

Background: Since 2004, we have collected patient-reported outcome (PRO) data from several Danish patient populations for use at the group and patient levels.

Objective: The aim of this paper is to highlight trends during the last 15 years with respect to patient inclusion, the methods for collection of PRO data, the processing of the data, and the actual applications and use of the PRO measurements.

Methods: All PRO data have been collected using the AmbuFlex/WestChronic PRO system, which was developed by the author in 2004 and has been continuously updated since. The analysis of trends was based on a generic model applicable for any kind of clinical health data, according to which any application of clinical data may be divided into four processes: patient identification, data collection, data aggregation, and the actual data use. Data for analysis were generated by a specific application in the system and transferred for analysis to the R package.

Results: During the 15-year period, 78,980 patients within 28 different groups of chronic and malignant illnesses have answered 260,433 questionnaires containing a total of 13,538,760 responses. Several marked changes have taken place: (1) the creation of cohorts for clinical epidemiological research purposes has shifted towards cohorts defined by clinical use of PRO data at the patient level; (2) the development of AmbuFlex, where PRO data are used as the entire basis for outpatient follow-up instead of fixed appointments, has undergone exponential growth and the system is currently in use in 47 International Statistical Classification of Diseases and Related Health Problems groups, covering 16,000 patients and 94 departments throughout Denmark; (3) response rates (up to 92%) and low attrition rates have been reached in group level projects, and there are even higher response rates in AmbuFlex where the patients are individually referred; (4) The answering method has shifted, as while in 2005 a total of 66.5% of questionnaires were paper based, this is the case for only 4.3% in 2019; and (5) the approach methods for questionnaires and reminders have changed dramatically from letter, emails, and short message service text messaging to a national, secure electronic mail system through which 93.2% of the communication to patients took place in 2019. The combination of secure email and web-based answering has resulted in a low turnaround time in which half of responses are now received within 5 days.

Conclusions: The demand for clinical use of PRO measurements has increased, driven by a wish among patients as well as clinicians to use PRO to promote better symptom assessment, more patient-centered care, and more efficient use of resources. Important technological changes have occurred, creating new opportunities, and making PRO collection and use cheaper and more feasible. Several legal changes may constitute a barrier for further development as well as a barrier for better utilization of patients' questionnaire data. The current legal restrictions on the joint use of health data imposed by the European Union's General Data Protection Regulation makes no distinction between use and misuse, and steps should be taken to alleviate these restrictions on the joint use of PRO data.

(J Med Internet Res 2019;21(9):e15856) doi:10.2196/15856

KEYWORDS

patient-reported outcome; questionnaires; chronic disease; outpatient follow-up; patient involvement; resource reallocation

Introduction

From the time of Hippocrates, information originating from the patient has been considered indispensable. Today, few diagnoses can be established, and few treatments monitored sufficiently, solely by using paraclinical data without explicit information from the patient. However, until recently, such patient inputs were always shortened and interpreted by a clinician. With the introduction of the term health-related quality of life, systematic measurement was adopted for research in several clinical specialties [1]. The potential of its many applications was further boosted by the US Food and Drug Administration's definition of a patient-reported outcome (PRO) as a measurement based on:

> *Any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else [2].*

The first draft of this term appeared in 2006 [3] and the final version in 2009 [2]. Since 2004, we have collected PRO data (although the term PRO data was not coined at that time) from several Danish patient populations. The aim of this paper is to highlight trends during the last 15 years with respect to patient inclusion, the collection of PRO data, the processing of the data, and the actual applications and use of the PRO measurements.

Methods

Overview

All PRO data have been collected using the AmbuFlex/WestChronic PRO system, which was developed by the author in 2004 and has been continuously updated since. The system has been described in detail elsewhere [4,5]. The first version of the generic PRO system, WestChronic, was developed for mixed-mode (Web and paper) collection of PRO data for research purposes in clinical epidemiological studies with repetitive measurements. Based on experienced feasibility and high response rates, it was decided in 2007 to develop this system into a flexible, multipurpose PRO system. The goal was to use clinical PRO data as the basis for outpatient follow-up in selected patient groups.

AmbuFlex: Telehealth Patient-Reported Outcomes as the Basis for Follow-Up in Chronic Diseases

In AmbuFlex, outpatients report their symptoms from home at regular intervals instead of attending fixed visits at the outpatient clinic. The PRO measures are used to decide whether a patient needs or wishes an outpatient visit, and were developed to promote better symptom monitoring, more patient-centered care, and more efficient use of resources [5]. Specific questionnaires have been developed for each diagnostic group. The AmbuFlex concept consists of three generic elements: PRO data collection, PRO-based automated decision algorithm, and PRO-based graphical overview for clinical decision support.

The AmbuFlex/WestChronic Patient-Reported Outcomes System

The AmbuFlex/WestChronic system supports dynamic mixed-mode data collection with the use of the internet or paper forms, as well as automated communication to the patient and the clinician via personalized letters, emails, text messages, and secure electronic communication. All information regarding implemented projects, items and questionnaires, communication, clinical users, and patients resides in tables in a Structured Query Language database residing in the server park of Region Central Denmark. All administration of projects, questionnaires, users, and patients is supported by the system's software and managed in browser windows.

The system has several integrations (Table 1 and Figure 1). All Danish citizens are assigned a unique 10-digit number (Civil Personal Registration [CPR] number), and continuously updated information on their current postal address and vital status is available from the Danish Civil Registration System [6]. This information is automatically collected online prior to any approach to patients. On-demand printing of questionnaires and letters, as well as scanning of incoming questionnaires with subsequent optical character recognition, is controlled by the system software, and results about all variables end up in result tables for the individual implemented projects in the same database. This occurs irrespective of whether Web or paper forms are used, and all results are instantaneously accessible. WestChronic may implement an arbitrary number of PRO projects with individual protocols, questionnaires, patients, and clinical users. For the patient and clinician, each implemented project appears as a unique PRO project with its own logo, domain, website, accompanying letters, contact information, etc. A new and rewritten version is underway.

Table 1. The AmbuFlex/WestChronic PRO system's online integrations with other systems.

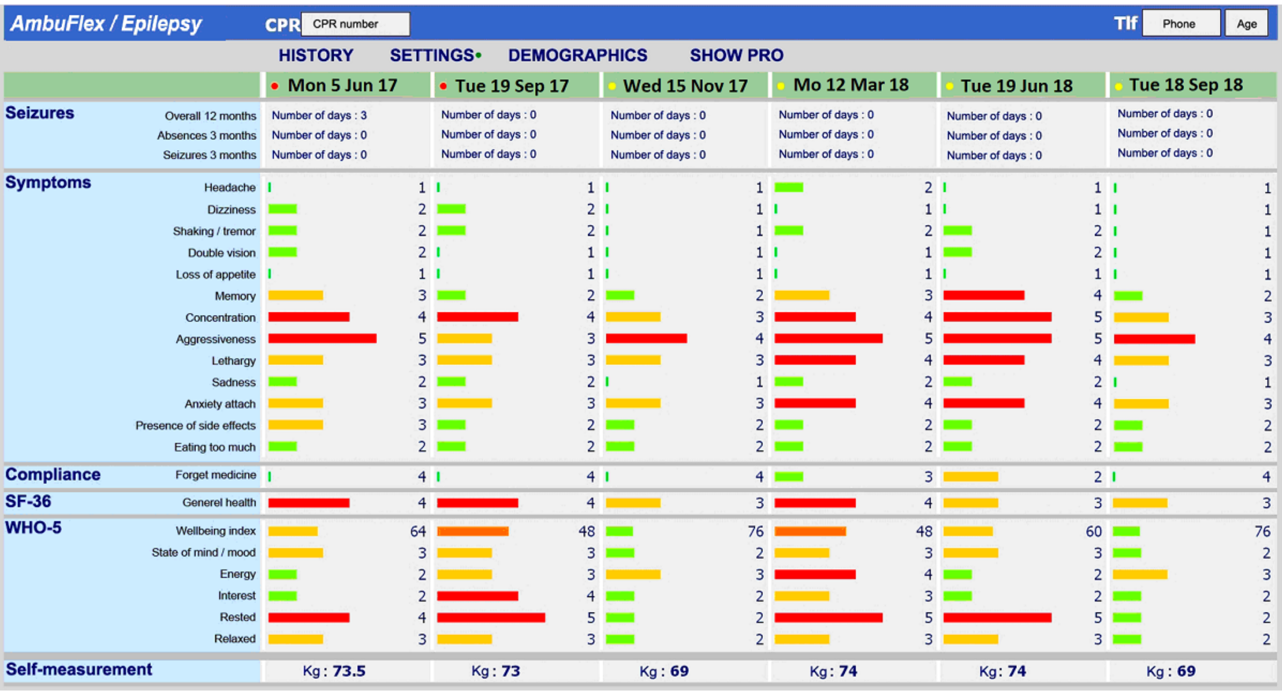
System	Purpose
Electronic Health Record system[a]	Clinicians may access a graphical overview of the patient's PRO[b] measurements in AmbuFlex with a single click from the patient's record in the EHR[c] (Figure 1).
Danish Civil Registration System	Information on current address and vital status, including possible date of death.
The national health portal (sundhed.dk)	Patients may, after secure login at the portal, access the same PRO overview and data as the clinician (Figure 1).
Email	Automated emailing of reminders, etc (obsolete).
Text messaging	Reminders and secure login (two-factor authentication).
Secure electronic mail (e-boks)	Automated mailing of links to questionnaires and reminders.
Paper questionnaire printing and scanning	Automated printing of individualized letters and questionnaires. Automated optical character recognition of received questionnaires.
Health data network (SDN)	A national secure virtual private network connecting hospitals, health data providers, etc.
Single-sign-on	Enables clinicians in other regions to login using their usual credentials.

[a] Available in three of the five Danish Regions. In the other two regions, AmbuFlex appears as a separate system.

[b] PRO: patient-reported outcome.

[c] EHR: electronic health record.

Figure 1. Example of the symptom overview in AmbuFlex/Epilepsy. The bar color and length both indicate the severity of the symptom (translated from Danish).



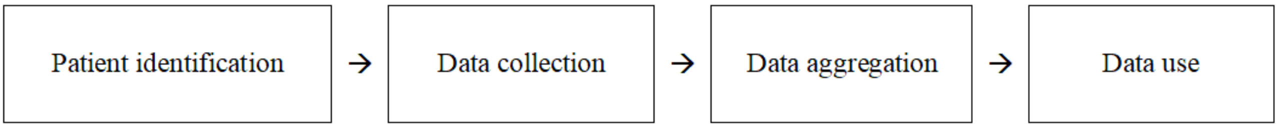
		● Mon 5 Jun 17	● Tue 19 Sep 17	● Wed 15 Nov 17	● Mo 12 Mar 18	● Tue 19 Jun 18	● Tue 18 Sep 18
Seizures	Overall 12 months	Number of days : 3	Number of days : 0	Number of days : 0	Number of days : 0	Number of days : 0	Number of days : 0
	Absences 3 months	Number of days : 0	Number of days : 0	Number of days : 0	Number of days : 0	Number of days : 0	Number of days : 0
	Seizures 3 months	Number of days : 0	Number of days : 0	Number of days : 0	Number of days : 0	Number of days : 0	Number of days : 0
Symptoms	Headache	1	1	1	2	1	1
	Dizziness	2	2	1	1	1	1
	Shaking / tremor	2	2	1	2	2	1
	Double vision	2	1	1	1	2	1
	Loss of appetite	1	1	1	1	1	1
	Memory	3	2	2	3	4	2
	Concentration	4	4	3	4	5	3
	Aggressiveness	5	3	4	5	5	4
	Lethargy	3	3	3	4	4	3
	Sadness	2	2	1	2	2	1
	Anxiety attach	3	3	3	4	4	3
	Presence of side effects	3	2	2	2	2	2
	Eating too much	2	2	2	2	2	2
Compliance	Forget medicine	4	4	4	3	2	4
SF-36	Generel health	4	4	3	4	3	3
WHO-5	Wellbeing index	64	48	76	48	60	76
	State of mind / mood	3	3	2	3	3	2
	Energy	2	3	3	4	2	3
	Interest	2	4	2	3	2	2
	Rested	4	5	2	5	5	2
	Relaxed	3	3	2	3	3	2
Self-measurement		Kg : 73.5	Kg : 73	Kg : 69	Kg : 74	Kg : 74	Kg : 69

Data Analysis

The data for the present paper were generated by an *ad hoc* application in the system and transferred for analysis in the R package (R Foundation for Statistical Computing, Vienna, Austria). Patients who had been referred but not yet answered at least one questionnaire were not included. The data collection in the first project started in September 2004, and during 2004 a total of 301 questionnaires were collected from 241 patients; this period was considered a pilot phase and not included in the analysis.

The analysis of trends was based on a generic model applicable for any clinical health data [7]. Each application of PRO data is divided into four processes: patient identification (eg, registers, consecutive or individual referral), data collection (logistics and organization), data aggregation (at patient or at some specified group level) and the actual data use (Figure 2). In the present paper, the last two processes will be described together. The patient identification process involves the identification of patient(s) from whom data are to be collected, the data collection process involves the actual collection of health data, including logistic procedures, the data aggregation process involves the management and organization of collected data for the data use process, and the data use process involves the use of the health data for the purpose of the specified activity. Each process may be repeated or may take place simultaneously with the previous process (from [7]).

Figure 2. The four processes in the lifespan of patient-related health data (from [7]).



Results

Primary Findings

An overview of selected projects is given in Table 2. In the first cohort studies, emphasis was on response rate, which was promoted by up to three reminders using email and letters and by giving the patient a free choice between Web and paper questionnaires. The maximum response rate was reached in patients with prostate cancer (Table 2), where the first questionnaire was answered by 95.9% of the patients. The PRO-based algorithm was first introduced in a screening program for patients with coronary heart diseases, and the AmbuFlex principle was fully unfolded in the implantation in epilepsy patients (Table 2, Figures 1 and 3).

Figure 4 and Table 3 show the development of the number of projects, patients, and received questionnaires. When considering that 2019 is only partly included, an exponential model of number of received questionnaires per calendar year explains 94% of the variation (Figure 4). At item level, a total of 13,538,760 responses have been received in the period. Until 5 years ago, web- and paper-based questionnaires were used with similar frequencies, but during the last 4 years web-based questionnaires have almost completely taken over. Projects where data was used at group level dominated until 2014, but PRO data used at the patient level is now by far the most frequent (Figure 5). Most projects and about one-third of the questionnaires now have an attached algorithm (AmbuFlex) by which the incoming answers are classified into two or three groups (Table 3 and Figure 3).

Patient Identification

In the first part of the 15-year period, all projects were register-based, epidemiologic cohort studies with defined inclusion criteria, while from 2015 onwards PRO data for clinical use at the patient level has dominated the scene (Table 3 and Figure 5). In the latter case, patients are individually referred by a clinician. Patient cohorts are still created at group level and are used in research for topics like outcome measures, validation, and studies of determinants of referral. In 2019, 16,062 new patients were included, of which 73.8% (11,854/16,062) were referred individually by a clinician while the rest were included as part of a group by means of batch processing.

Patient-Reported Outcome Data Collection

Fundamental changes have taken place with respect to PRO data collection over the course of very few years. Until 2014, paper-based communication was used in more than 90% of the approaches (Figure 6), while electronic communication was predominantly used in epidemiological studies with multiple measurements where about half of the communication was electronic (emailing and web-based answering) (Figures 4 and 6). A minor revolution took place November 1, 2014. From this date, all communication between the Danish authorities (including all public hospitals), and Danish citizens had to use a national, secure electronic mail system, called e-boks. All citizens are provided an e-boks, but while its use was previously voluntary, only people with disabilities can now be granted a dispensation and continue to receive paper letters from places like hospitals. After the AmbuFlex/WestChronic system gained access to e-boks (Table 1), this channel rapidly became the most important communication method to reach patients (Figure 6). These changes in communication methods to and from patients had consequences for turnaround time (ie, the time from asking the patient to answer until the questionnaire is received) (Figure 7). A total of 75% of the answers were received within 17 days in 2005 due to the homogeneous cohorts, but turnaround time increased slowly until 2014. At this time, the Danish Mail decreased their services due to economic problems caused by the steep fall in paper-based mail. The historic low turnaround time in 2019 is caused by a combination of e-boks communication and the web-answering method. Today, the first answers are received within minutes. Use of ordinary, unsecure email ended in 2018. Text messaging is still used for reminders in some projects but is mainly used as a 2-factor authentication tool, where the patient, after entering their credentials (CPR number), receives a 6-digit cipher code as a text message.

Table 2. Characteristics of selected PRO-based projects implemented in the AmbuFlex/WestChronic system.

Studies	Characteristics	Description
Group-level projects		
C50 Breast cancer	PRO[a] data collection: 2004-14 107 items/questionnaire 1008 patients (100% females) 4731 questionnaires (43.1% web) Follow-up[b]: 4.7 (9.1) years Papers: 3 [8-10]	The source population was women referred to mammography at two regional hospitals in Region Central Denmark on clinical suspicion of breast cancer. The women were mailed a baseline questionnaire which was filled in and returned before the date of mammography. All respondents were subsequently interviewed by telephone and invited to join the follow-up study, irrespective of diagnosis. Women diagnosed with breast cancer (n=256; 7.2% of respondents) as well as a sample of women without cancer (n=291) were followed every 3 months for up to 9 years with questionnaires including generic scales on fatigue (MFI-20[c]) [11], depression and anxiety (HADS[d]) [12], and selected items from SF-36[e] [13]. Inclusion in the study ended when a national plan for fast-track diagnosing of cancer was implemented in 2008, with which the design was incompatible since baseline data could no longer be collected before diagnosis. The final cohort size was 60% of the planned size. Register information on treatment and survival were collected from national cancer databases.
I20 Ischemic heart diseases	PRO data collection: 2006-13 103 items/questionnaire (1,323 patients (20.4% females) 13,171 questionnaires (40.3% web) Follow-up: 3.4 (7.0) years Papers: 6 [14-19]	Patients treated with percutaneous coronary intervention at Aarhus University Hospital, which at that time performed this procedure on behalf of the entire Region Central Denmark (1.3 million inhabitants). Patients <80 years old were included based on records in the hospital administration system. Patients were followed up with questionnaires every fourth month for up to 7 years, including HADS [12], MFI-20 [11], IPAQ[f] [20], selected items from SF-36 [13], items related to the psychosocial job-strain model [21], the Seattle angina questionnaire [22], and questions on lifestyle and rehabilitation. Register information on treatment, survival, and sick leave was collected from national databases.
I64 Stroke	PRO data collection: 2009-14 21 items/questionnaire 2618 patients (39.1% females) 9622 questionnaires (25.3% web) Follow-up: 2.1 (5.1) years Papers: 2 [23,24]	Patients with first-time stroke in Region Central Denmark were included prospectively based on online access to a national, disease-specific register. Patients were followed with questionnaires every fourth month for up to 5 years, including HADS [12], MFI-20 [11], WHODAS[g] [25], and selected items from SF-36 [13]. Register information on treatment, survival, and sick leave was collected from national databases.
C61 Prostate cancer	PRO data collection: 2011-17 32 items/questionnaire 13,434 patients (0.0% females) 16,066 questionnaires (15.1% web) Follow-up: 1.7 (3) years	The Danish quality database DaProCa has recorded clinical information nationwide in patients with prostatic cancer since 2010 [26]. In 2011, it was decided that PRO information should also be included to better describe the treatment outcome. Patients were identified in the register and mailed questionnaires 1 year and 3 years after initial diagnosis. The project was the first where an initial response rate above 90% was obtained. The PRO data in the quality databases have, however, only been used sporadically.
Patient-level projects		
I20 Ischemic heart diseases	PRO data collection: 2011-17 14 items/questionnaire 5097 patients (40.5% females) 5121 questionnaires (19.4% web) Papers: 1 [27]	This project was the first to use automatic algorithm-based handling of questionnaires. Patients admitted to Hospital Unit West Jutland for treatment of ischemic heart disease were identified based on ICD-10[h] diagnoses in the business intelligence register in Region Central Denmark. Patients were mailed a questionnaire containing the HADS [12]. An algorithm embedded in the AmbuFlex/WestChronic system processed the incoming answers and printed individualised letters with the results. Patients were advised to contact their GP[i] if depression or anxiety scores were above the established thresholds, which was the case in 30.2% of the responses. An analysis based on register information on consultations with a GP or psychologist revealed that only a few additional cases of depression were detected.

Studies	Characteristics	Description
G40 Epilepsy	6 departments PRO data collection: 2012-present 47 items/questionnaire 6405 patients (50.5% females) 21,296 questionnaires (56.7% web) Follow-up: 3.4 (7.6) years Papers: 6 [5,28-32]	This project was the first project where PRO data were used as the basis for outpatient follow-up (AmbuFlex) and was developed in close cooperation with the Department of Neurology, Aarhus University Hospital. Patients with epilepsy attending neurological outpatient clinics are individually referred to AmbuFlex follow-up. Instead of fixed appointments at the clinic every 3, 6, or 12 months, the patients are prompted to answer a short disease-specific questionnaire developed in cooperation with the clinicians. Based on an automated algorithm, red and yellow flags as well as patient preferences are identified. Patients with flags or a wish for contact are included on the clinicians online to-do list, and the PRO overview is displayed to the clinicians within the Electronic Health Record system (Figure 1). Questionnaires with no flags and no wish for a contact are handled automatically by AmbuFlex and a new questionnaire (eg, 3 months later), is scheduled. Overall, 53% of the PRO-based contacts are handled without further contact to the patient (Figure 3) This was the first of three AmbuFlex projects implemented on a national basis 2013 [33].
C34 Lung cancer	8 departments PRO data collection: 2014-present 52 items/questionnaire 2274 patients (50.1% females) 12,658 questionnaires (100% web) Follow-up: 0.6 (4.6) years	Patients treated for lung cancer were asked by the front-desk staff to fill in the online questionnaire in the waiting area at each follow-up outpatient clinic visit. The intention was to use the PRO information in the consultation a few minutes later. The project was implemented at seven departments throughout Denmark in cooperation with the Danish Cancer Society. Log-files in the AmbuFlex/WestChronic system are kept to document each time patient data are displayed and to whom. On average, only 47% of the questionnaires were viewed by a clinician (Figure 7), with huge differences between hospitals ranging from 14-93% [34].
M05 Rheumatoid arthritis	2 departments PRO data collection: 2014-present 40 items/questionnaire 676 patients (69.7% females) 2785 questionnaires (84.5% web) Follow-up: 1.7 (5.0) years Publications: 1 [35]	The project started as a non-inferiority randomized controlled trial conducted in cooperation with Rheumatologic Department, Aarhus University Hospital, where patients were randomized to PRO-based telehealth or conventional outpatient follow-up. Disease activity was measured by the Danish version of the Flare-RA instrument [36], and the primary outcome was a change in the DAS28[j]. PRO-based tele-health achieved disease control like that of conventional outpatient follow-up. All patients were contacted. The project has continued as an AmbuFlex project with a similar questionnaire, where only patient answers with red or yellow flags are assessed and patients contacted if needed.
C80 Side effects during antineoplastic treatment	2 departments PRO data collection: 2015-present 60 items/questionnaire 7011 patients (57.8% females) 59,202 questionnaires (100% web) Follow-up: 0.3 (4.2) years	In cancer treatment, questions on toxicity symptoms are normally not asked and registered systematically, and the ongoing therapy is therefore not evaluated in accordance with the present state of the patient. In several AmbuFlex projects, PRO-based self-reports are used during the period the patient is receiving chemotherapy in an outpatient setting. PRO data are used to decide if the planned chemotherapy should be postponed or adjusted. The PRO measures used include items from PRO-CTCAE[k] [37], EORTC[l] [38] and PRO measures based on ad hoc developed single red-flag items.
C34 Lung cancer	6 departments PRO data collection: 2018-present 17 items/questionnaire 69 patients (56.7% females) 973 questionnaires (100% web) Follow-up: 0.2 (0.9) years	PRO-based systematic symptom monitoring may improve overall survival in cancer patients who are followed up with after their initial treatment [39,40]. In this Danish multicenter RCT study, we compare standard follow-up with weekly PRO measurements where red flag responses are automatically reported to clinicians for further evaluation.
C80 Cancer, inpatients	1 department PRO data collection: 2017-present 22 items/questionnaire 868 patients (60.9% females) 4849 questionnaires (100% web) Follow-up: 0.1 (1.8) years	Clinical use of PRO measures often includes only outpatients. In this developmental implementation of AmbuFlex, PRO data are used in inpatients to support the dialogue between the patient, the nurse, and the doctor while the patient is hospitalized, and they are used to prioritize patients to be discussed during the daily rounds.

Studies	Characteristics	Description
G40 Epilepsy	2 departments PRO data collection: 2015-present 34 items/questionnaire 182 patients (45.6% females) 449 questionnaires (31.9% web) Follow-up: 2.2 (4.3) years	In all patient groups, a proportion of patients are not capable of answering a questionnaire. Some patients suffering from epilepsy live in institutions or are taken care of by their next of kin. These patients may be at increased risk of having important symptoms left unnoticed by the health care system in connection with normal follow-up, since the person accompanying the patient to the hospital may not be the person who has the most knowledge about the patient. In AmbuFlex/Epilepsy, a specific proxy questionnaire was developed with an algorithm like that used by the other epilepsy patients.

[a] PRO: patient-reported outcome.

[b] Median follow-up with maximum in parenthesis. Based on patients who have answered at least two questionnaires by September 8, 2019.

[c] MFI: Multidimensional fatigue inventory.

[d] HADS: Hospital Anxiety and Depression Scale.

[e] SF: Short Form Health Survey.

[f] IPAQ: International Physical Activity Questionnaire.

[g] WHODAS: World Health Organization disability assessment schedule 2.0.

[h] ICD-10: 10th edition of the International Statistical Classification of Diseases and Related Health Problems.

[i] GP: general practitioner.

[j] DAS: disease activity score.

[k] PRO-CTCAE: patient-reported outcome measure–Common Terminology Criteria for Adverse Events.

[l] EORTC: European Organization for Research and Treatment of Cancer.

Figure 3. Flow chart for outpatients with epilepsy (AmbuFlex). Patients answer the disease-specific questionnaire at fixed intervals (eg, 3 months). In the first step, the answers are processed automatically based on a disease-specific algorithm. Green response: No need or wish for contact (a new questionnaire is scheduled in, eg, 3 months). Yellow response: May need contact (a clinician assesses the PRO overview (Figure 1) and other information to decide whether further contact is needed). Red response: Definite need or wish for contact (the patient is contacted). In total, only 47% of the patients are contacted in each round. PRO: patient-reported outcome.

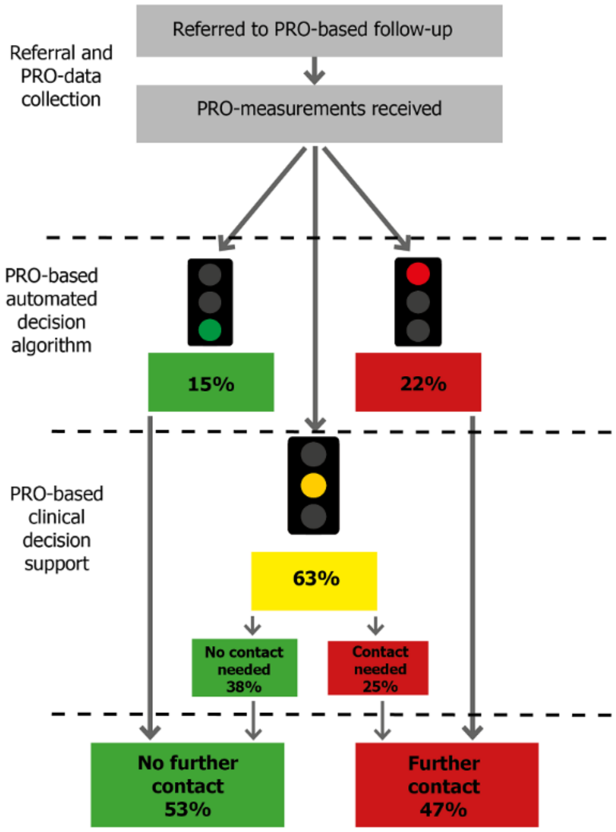


Figure 4. Questionnaires received by the AmbuFlex/WestChronic system 2005-19 by answering method.

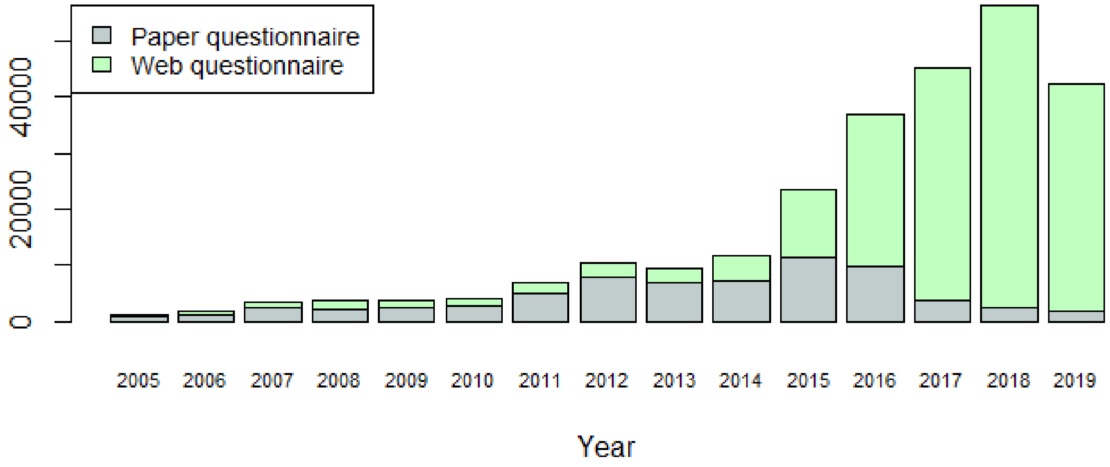


Table 3. Summary of PRO-based projects by year from 2005-2019.

Year	Projects	ICD-10[a] groups	Level of aggregation, %			Departments[b]	Patients[b]	Questionnaires[b]
			Group	Patient, no algorithm	Patient, with algorithm			
2005	1	1	100	0	0	0	542	1200
2006	2	2	100	0	0	0	923	1876
2007	2	2	100	0	0	0	1318	3421
2008	2	2	100	0	0	0	1232	3630
2009	3	4	80	20	0	1	1606	3747
2010	2	4	75	25	0	1	2214	4098
2011	6	6	57	14	29	13	4721	6848
2012	8	7	67	0	33	17	7733	10,434
2013	11	10	58	0	42	10	7174	9346
2014	16	11	35	12	53	23	8415	11,816
2015	17	14	17	22	61	39	16,490	23,444
2016	23	16	13	25	63	49	20,201	36,912
2017	34	20	18	21	62	67	18,721	45,058
2018	44	24	14	27	59	91	21,143	56,178
2019[c]	47	26	15	27	58	97	18,262	42,127
Total	64	28	28	26	46	141	78,980	260,433

[a] ICD-10: 10th edition of the International Statistical Classification of Diseases and Related Health Problems.

[b] Patients and departments may be involved in more than one disease-specific project.

[c] As of September 8, 2019.

Figure 5. Questionnaires received by the AmbuFlex/WestChronic system 2005-19 by type of use.

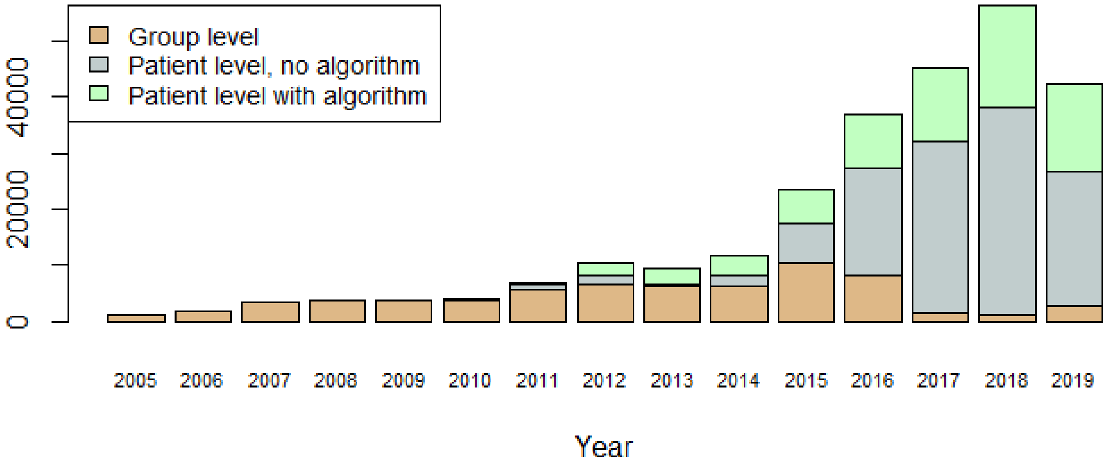


Figure 6. Contact to patients by year, and method for prompting the patients to fill in the questionnaire.

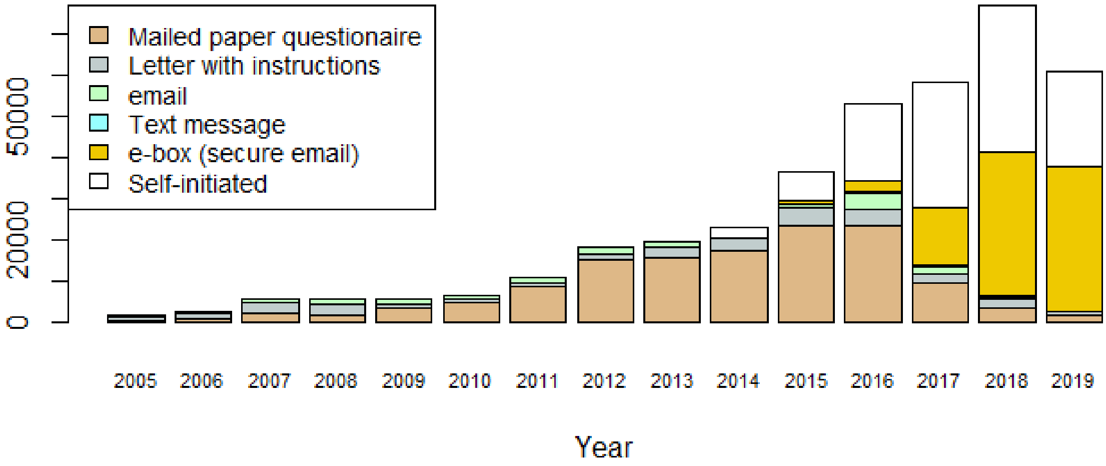
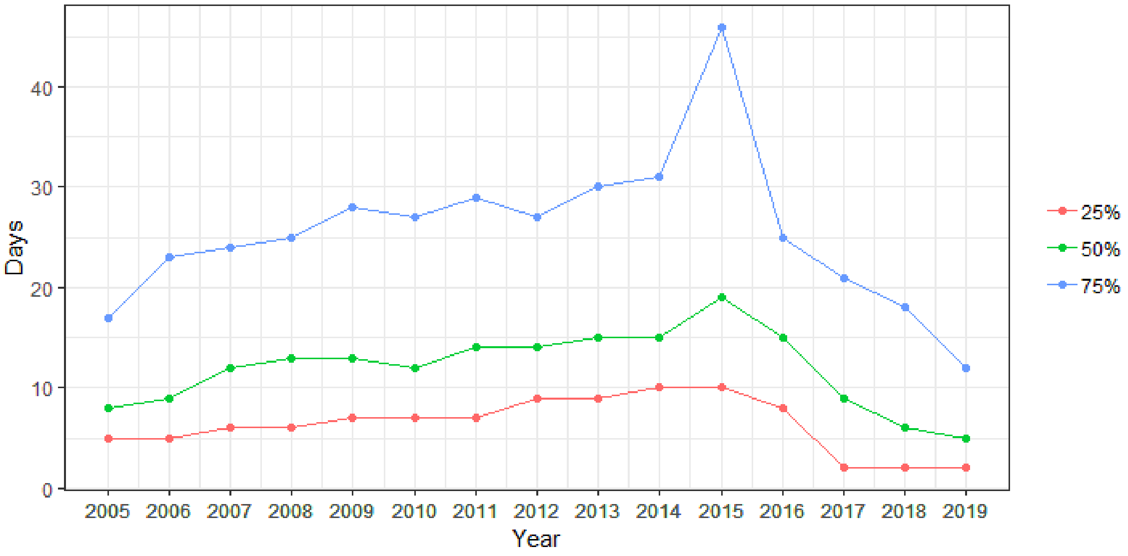


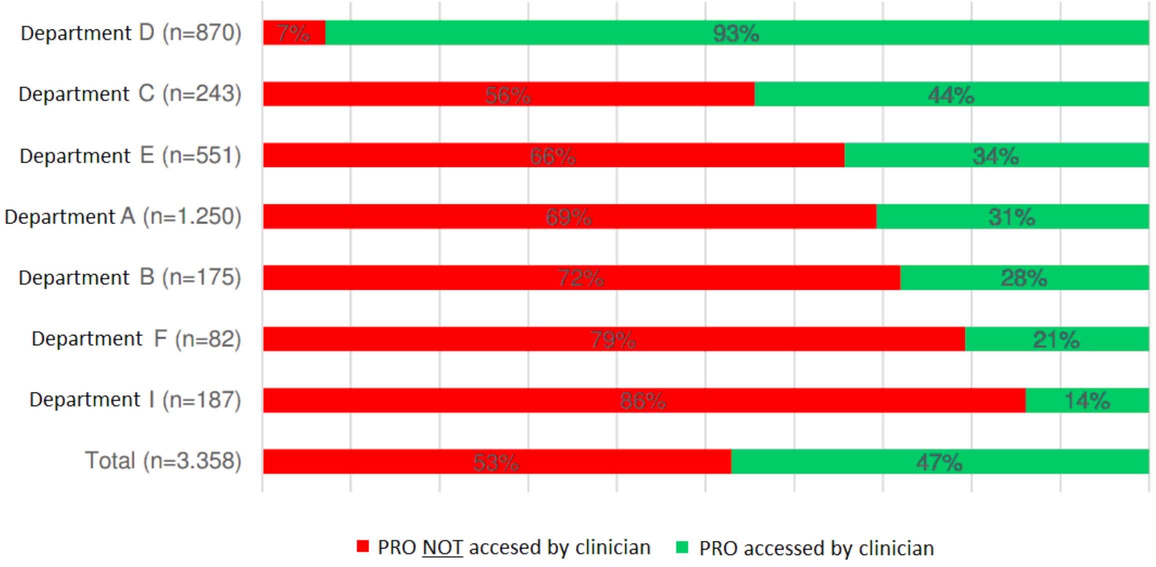
Figure 7. Turnaround time by year. Turnaround time is defined as the number of days from when the request is sent to the patient to when the answer is received.



Patient-Reported Outcome Data Aggregation and Data Use

While data from group-level projects are aggregated and used in the analysis phases of each epidemiological study, data in patient-level projects are, or should be, aggregated and used the date the questionnaire is received; however, this is not always the case. If PRO data are used as an add-on in connection with ordinary organized outpatient follow-up, it is up to each clinician to decide to open the PRO overview or not. In a multicenter project covering eight Danish hospitals, the patients treated for lung cancer answered a PRO questionnaire each time they went to the hospital for a follow-up visit, with the idea being that the PRO data should be used to enforce patient-clinician communication during the consultation (Table 2). For legal reasons, the AmbuFlex/WestChronic system keep automated log files of when and to whom the PRO data are shown. In total, a minority of responses (47%) were seen by a clinician (Figure 8). There were major differences between departments, ranging from 14-93% [34]. Similar figures have been found in other projects where the PRO data are an add-on to the existing outpatient set-up. In AmbuFlex, however, the PRO data are not an add-on but the basis for the follow-up. Each time a questionnaire is received, it is either handled automatically (green response, Figure 3) or included on a list where it remains until a clinician has accessed it and made a clinical decision regarding whether the patient should be contacted or not. Therefore, virtually no AmbuFlex questionnaires remain unnoticed.

Figure 8. The proportion of PRO questionnaires that was viewed by a clinician during the follow-up visit of patients treated for lung cancer in seven Danish oncological departments (translated from Danish [34]). PRO: patient-reported outcome.



Discussion

During the 15-year period, we have experienced an exponential growth in the number of answered questionnaires and several marked changes have taken place. The creation of cohorts for clinical epidemiological research purposes has shifted towards cohorts defined by clinical use of PRO data at the patient level, and PRO-based follow-up (AmbuFlex) has replaced fixed appointments in many patient groups. Overall, the system is currently in use in 47 International Statistical Classification of Diseases and Related Health Problems (ICD)–groups, covering 16,000 patients and 94 departments throughout Denmark. In the first part of the period, a combination of paper- and web-based questionnaires secured response rates above 90% and low attrition rates. Back in 2005 66.5% questionnaires were paper based, but this is only the case for 4.3% in 2019. In AmbuFlex even higher response-rates have been reached, but the numbers are not directly comparable because each patient is individually referred and has given consent to participate. In addition, the approach methods for questionnaires and reminders have changed from letters, emails, and text messaging to a national, secure, electronic mail system through which 93.2% of communication to patients took place in 2019. The combination of secure email and web-based answering has resulted in a low turnaround time, where half of all responses are received within 5 days.

Our experiences probably reflect general changes that have taken place during the 15-year period. First, the demand for PRO measurements has increased, quantitatively as well as qualitatively. PRO measures have been used at the group level for many years, even before the PRO term was coined, when it was termed Quality of Life measures among other things (particularly in clinical trials and observational studies). PRO measures have also been used at the group level to provide evidence for drug and device approval and, in some countries, used to evaluate quality of care and health service provider performance. By contrast, using PRO data systematically at the individual patient level might be relatively new, but it is now by far our predominant activity. This increased demand is driven by a wish among patients, as well as clinicians, to use the potential in PRO data to promote better symptom assessment, more patient-centered care, and more efficient use of resources. In Denmark, hospital administrators and health ministerial civil servants quickly realized PRO measures' potential to prioritize resources to the outpatients who actually want or need clinical attention.

Second, substantial technological changes have occurred in the period, which has created new opportunities. For example, an electronic health record (EHR) system has been implemented that covers the whole Region Central Denmark, and AmbuFlex obtained an early online connection that allowed clinicians to access it because it was a part of the EHR. In addition, the national secure electronic mail system (e-boks) has had a high impact on reducing response time.

Finally, several legal changes have occurred. The rationale for implementation of the General Data Protection Regulation (GDPR) is sound because there is a real risk of patients' information being accessed and used by people for whom it was not initially intended. However, much time is used on details and bureaucratic documentation procedures, which doubly enhance data security. Another legal change is the European Union's medical device regulation. Questionnaires with an attached algorithm, like those used in AmbuFlex, are classified as a medical device and as such they must be certified with a certification (Conformité Européenne) marking. Patient safety is a cornerstone, but since we are dealing with outpatients who are instructed to contact their family doctor or the department directly in case of exacerbation it makes little sense to treat a questionnaire with the same rules as electronic medical equipment. At present, AmbuFlex is granted a dispensation for current projects until the marking is in place, but not allowed to launch new projects. If AmbuFlex were to start today, it is unlikely that we would ever have surfaced.

One issue, which is often overlooked, is whether collected PRO data are used sufficiently [7,41]. Most patients are very careful when filling in their questionnaires, and it should be our obligation to promote as much use of the data as possible, irrespective of if data are originally collected for research, quality assessment, or clinical use at the patient level. In the future, clinical settings will probably be the most important source of PRO data, but also for other purposes like research and quality surveillance, so new ways to conduct complementary data collection will be necessary [7]. The current legal restrictions on the joint use of health data imposed by the GDPR make no distinction between use and misuse, and steps should be taken to alleviate these legal restrictions on the joint use of PRO data.

Acknowledgments

The described developments would have been impossible without contributions from several collaborators throughout the 15 years. Scientific collaborators from AmbuFlex/WestChronic include Sissel Marie Kongsved, Maja Sidelmann Basnov, Elinborg Thorsteinsson, Nina Boje Nielsen, Karin Biering, Louise Pape Larsen, Liv Marit Schougaard, Caroline Trillingsgaard Mejdahl, and Birgith Engelst Grove. External scientific collaborators include Søren Paaske Johnsen, Ole May, Annette Ladefoged de Thurah, Helle Pappot, Per Ramløv Ivarsen, Anette Kargo, Troels Krarup Hansen, Annesofie Lunde Jensen, Halla Skuladottir, and Rasmus Blechingberg Friis. Furthermore, the following IT developers have contributed: Rasmus W Lauritsen, Klaus Hvam Petersen, Jesper Medom Vestergaard, and Kasper Bjerg, together with AmbuFlex' developers and project managers: Anne Jessen, Liv Dørflinger, Anke Elisabeth de Beijer and the rest of the AmbuFlex team members not mentioned above: Mette Vittinghus, Nanna Toxvig Sørensen, Ann-Katrine Kvisgaard Jensen, Laila Bech Olesen, Maria Aagaard Andreasen, Maiken Norup Petersen, Stine Primdahl Rasmussen, Mette Kjær Jensen, Gunnar Hedin Heinesen, Martin Tang Olesen, and Rasmus Andreas Nielsen. Also, more than 15 part-time student assistants have provided valuable help over the years. Several decision makers and administrators have provided important and necessary support for AmbuFlex at different stages. From Region Central Denmark: Kurt Rasmussen, Per Østergaard Jensen, Ida Götke, Henning Vestergaard, Henrik Toft Sørensen, Paul Bartels, Ole Thomsen, Britta Ravn, Halfdan Eika, Jens Friis Bak, and Ole Carstensen, and from the Danish Ministry of Health: Nanna Skovgaard. AmbuFlex would not have been possible without the input from dedicated patients as well as clinicians from the following departments, who have provided essential ideas and feed-back. From Aarhus University Hospital, Department of Neurology: Per Sidenius, Mette Abildgaard Høi, Lene Elbek Overbeck, Anette Møller Clausen, and Birgitte Jensen Hanner, and from Hospital Unit West Jutland, Department of Cardiology: Kirsten Hagger and Department of Oncology: Hanne Linnet.

Conflicts of Interest

None declared.

References

1. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. BMJ 2002 Jun 15;324(7351):1417 [FREE Full text] [doi: 10.1136/bmj.324.7351.1417] [Medline: 12065262]
2. Food and Drug Administration. Patient-reported outcome measures: Use in medicinal product development to support labeling claims. 2009. Guidance for industry URL: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf [accessed 2019-09-24] [WebCite Cache ID 6qqbfANOC]
3. Speight J, Barendse SM. FDA guidance on patient reported outcomes. BMJ 2010 Jun 21;340:c2921. [doi: 10.1136/bmj.c2921] [Medline: 20566597]
4. Hjollund NHI, Larsen LP, Biering K, Johnsen SP, Riiskjær E, Schougaard LM. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. Interact J Med Res 2014 Feb 11;3(1):e5 [FREE Full text] [doi: 10.2196/ijmr.2885] [Medline: 24518281]
5. Schougaard LMV, Larsen LP, Jessen A, Sidenius P, Dorflinger L, de Thurah A, et al. AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Qual Life Res 2016 Mar;25(3):525-534 [FREE Full text] [doi: 10.1007/s11136-015-1207-0] [Medline: 26790427]
6. Pedersen CB. The Danish Civil Registration System. Scand J Public Health 2011 Jul;39(7 Suppl):22-25. [doi: 10.1177/1403494810387965] [Medline: 21775345]
7. Hjollund NHI, Valderas JM, Kyte D, Calvert MJ. Health Data Processes: A Framework for Analyzing and Discussing Efficient Use and Reuse of Health Data With a Focus on Patient-Reported Outcome Measures. J Med Internet Res 2019 May 21;21(5):e12412 [FREE Full text] [doi: 10.2196/12412] [Medline: 31115347]
8. Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. J Med Internet Res 2007 Sep 30;9(3):e25 [FREE Full text] [doi: 10.2196/jmir.9.3.e25] [Medline: 17942387]
9. Basnov M, Kongsved SM, Bech P, Hjollund NH. Reliability of short form-36 in an Internet- and a pen-and-paper version. Inform Health Soc Care 2009 Jan 12;34(1):53-58. [doi: 10.1080/17538150902779527] [Medline: 19306199]
10. Nielsen NB, Zachariae R, Hjollund NH. Perceived risk prior to mammography as an independent predictor of breast cancer. J Psychosom Res 2009 Mar;66(3):245-247. [doi: 10.1016/j.jpsychores.2008.08.006] [Medline: 19232237]
11. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res 1995 Apr;39(3):315-325. [doi: 10.1016/0022-3999(94)00125-o] [Medline: 7636775]
12. Snaith RP. The Hospital Anxiety And Depression Scale. Health Qual Life Outcomes 2003 Aug 01;1:29 [FREE Full text] [doi: 10.1186/1477-7525-1-29] [Medline: 12914662]
13. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
14. Biering K, Bøtker HE, Niemann T, Hjollund NH. Patient-reported health as a prognostic factor for adverse events following percutaneous coronary intervention. Clin Epidemiol 2014;6:61-70 [FREE Full text] [doi: 10.2147/CLEP.S54237] [Medline: 24511242]
15. Biering K, Hjøllund NH, Lund T. Methods in measuring return to work: a comparison of measures of return to work following treatment of coronary heart disease. J Occup Rehabil 2013 Sep 25;23(3):400-405. [doi: 10.1007/s10926-012-9405-x] [Medline: 23184388]
16. Biering K, Nielsen TT, Rasmussen K, Niemann T, Hjollund NH. Return to work after percutaneous coronary intervention: the predictive value of self-reported health compared to clinical measures. PLoS One 2012 Nov 16;7(11):e49268 [FREE Full text] [doi: 10.1371/journal.pone.0049268] [Medline: 23173052]
17. Biering K, Frydenberg M, Hjollund NH. Self-reported health following percutaneous coronary intervention: results from a cohort followed for 3 years with multiple measurements. Clin Epidemiol 2014;6:441-449 [FREE Full text] [doi: 10.2147/CLEP.S65476] [Medline: 25506246]

18. Biering K, Lund T, Andersen JH, Hjollund NH. Effect of Psychosocial Work Environment on Sickness Absence Among Patients Treated for Ischemic Heart Disease. J Occup Rehabil 2015 Dec;25(4):776-782. [doi: 10.1007/s10926-015-9587-0] [Medline: 26077204]
19. Biering K, Andersen JH, Lund T, Hjollund NH. Psychosocial Working Environment and Risk of Adverse Cardiac Events in Patients Treated for Coronary Heart Disease. J Occup Rehabil 2015 Dec;25(4):770-775. [doi: 10.1007/s10926-015-9585-2] [Medline: 26044947]
20. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr 2006 Sep;9(6):755-762. [Medline: 16925881]
21. Karasek R, Baker D, Marxer F, Ahlbom A, Theorell T. Job decision latitude, job demands, and cardiovascular disease: a prospective study of Swedish men. Am J Public Health 1981 Jul;71(7):694-705. [doi: 10.2105/ajph.71.7.694] [Medline: 7246835]
22. Kimble LP, Dunbar SB, Weintraub WS, McGuire DB, Fazio S, De AK, et al. The Seattle angina questionnaire: reliability and validity in women with chronic stable angina. Heart Dis 2002;4(4):206-211 [FREE Full text] [Medline: 12147179]
23. Larsen LP, Biering K, Johnsen SP, Andersen G, Hjollund NH. Self-rated health and return to work after first-time stroke. J Rehabil Med 2016 Apr;48(4):339-345 [FREE Full text] [doi: 10.2340/16501977-2061] [Medline: 26936419]
24. Larsen LP, Johnsen SP, Andersen G, Hjollund NH. Determinants of Self-Rated Health Three Months after Stroke. J Stroke Cerebrovasc Dis 2016 May;25(5):1027-1034. [doi: 10.1016/j.jstrokecerebrovasdis.2015.12.014] [Medline: 26920345]
25. Saltychev M, Katajapuu N, Bärlund E, Laimi K. Psychometric properties of 12-item self-administered World Health Organization disability assessment schedule 2.0 (WHODAS 2.0) among general population and people with non-acute physical causes of disability - systematic review. Disabil Rehabil 2019 Jul 23:1-6. [doi: 10.1080/09638288.2019.1643416] [Medline: 31335215]
26. Regionernes Kliniske Kvalitetsudviklingsprogram (RKKP). The Danish Prostate Cancer Database (In Danish). URL: https://www.sundhed.dk/sundhedsfaglig/kvalitet/kliniske-kvalitetsdatabaser/kraeft/prostata/ [accessed 2019-09-24]
27. Larsen KK, Vestergaard CH, Schougaard LMV, Larsen LP, Jessen A, May O, et al. Contacts to general practice and antidepressant treatment initiation after screening for anxiety and depression in patients with heart disease. Dan Med J 2016 Feb;63(2):A5185. [Medline: 26836792]
28. Schougaard LMV, Mejdahl CT, Petersen KH, Jessen A, de Thurah A, Sidenius P, et al. Effect of patient-initiated versus fixed-interval telePRO-based outpatient follow-up: study protocol for a pragmatic randomised controlled study. BMC Health Serv Res 2017 Jan 26;17(1):83 [FREE Full text] [doi: 10.1186/s12913-017-2015-8] [Medline: 28122609]
29. Schougaard LMV, de Thurah A, Christiansen DH, Sidenius P, Hjollund NH. Patient-reported outcome (PRO) measure-based algorithm for clinical decision support in epilepsy outpatient follow-up: a test-retest reliability study. BMJ Open 2018 Jul 25;8(7):e021337 [FREE Full text] [doi: 10.1136/bmjopen-2017-021337] [Medline: 30049693]
30. Schougaard LMV, de Thurah A, Bech P, Hjollund NH, Christiansen DH. Test-retest reliability and measurement error of the Danish WHO-5 Well-being Index in outpatients with epilepsy. Health Qual Life Outcomes 2018 Sep 06;16(1):175 [FREE Full text] [doi: 10.1186/s12955-018-1001-0] [Medline: 30189867]
31. Mejdahl CT, Schougaard LMV, Hjollund NH, Riiskjær E, Thorne S, Lomborg K. PRO-based follow-up as a means of self-management support - an interpretive description of the patient perspective. J Patient Rep Outcomes 2017;2:38 [FREE Full text] [doi: 10.1186/s41687-018-0067-0] [Medline: 30238083]
32. Mejdahl CT, Schougaard LMV, Hjollund NH, Riiskjær E, Lomborg K. Exploring organisational mechanisms in PRO-based follow-up in routine outpatient care - an interpretive description of the clinician perspective. BMC Health Serv Res 2018 Jul 13;18(1):546 [FREE Full text] [doi: 10.1186/s12913-018-3352-y] [Medline: 30001729]
33. Hjollund NH, Andersen T, Bjerrum A, Skovgaard N, Freil M. 22nd Annual Conference of the International Society for Quality of Life Research. Qual Life Res 2015 Oct;24 Suppl 1:1-191. [doi: 10.1007/s11136-015-1078-4] [Medline: 26395277]
34. The Danish Cancer Society. Clinical application of Patient Reported Outcome Measures. An evaluation. 2016. URL: https://www.cancer.dk/dyn/resources/File/file/9/6659/1505726340/klinisk-anvendelse-af-patient-reported-outcome-measures-prom.pdf [accessed 2019-09-24]
35. de Thurah A, Stengaard-Pedersen K, Axelsen M, Fredberg U, Schougaard LMV, Hjollund NHI, et al. Tele-Health Followup Strategy for Tight Control of Disease Activity in Rheumatoid Arthritis: Results of a Randomized Controlled Trial. Arthritis Care Res (Hoboken) 2018 Mar;70(3):353-360 [FREE Full text] [doi: 10.1002/acr.23280] [Medline: 28511288]
36. de Thurah A, Maribo T, Stengaard-Pedersen K. Patient self-assessment of flare in rheumatoid arthritis: criterion and concurrent validity of the Flare instrument. Clin Rheumatol 2016 Feb;35(2):467-471. [doi: 10.1007/s10067-014-2849-y] [Medline: 25519219]
37. Baeksted C, Pappot H, Nissen A, Hjollund NH, Mitchell SA, Basch E, et al. Feasibility and acceptability of electronic symptom surveillance with clinician feedback using the Patient-Reported Outcomes version of Common Terminology Criteria for Adverse Events (PRO-CTCAE) in Danish prostate cancer patients. J Patient Rep Outcomes 2017;1(1):1 [FREE Full text] [doi: 10.1186/s41687-017-0005-6] [Medline: 29757324]
38. Snyder CF, Blackford AL, Okuyama T, Akechi T, Yamashita H, Toyama T, et al. Using the EORTC-QLQ-C30 in clinical practice for patient management: identifying scores requiring a clinician's attention. Qual Life Res 2013 Dec;22(10):2685-2691 [FREE Full text] [doi: 10.1007/s11136-013-0387-8] [Medline: 23532341]

39. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
40. Denis F, Basch E, Septans A, Bennouna J, Urban T, Dueck AC, et al. Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer. JAMA 2019 Jan 22;321(3):306-307 [FREE Full text] [doi: 10.1001/jama.2018.18085] [Medline: 30667494]
41. Calvert M, Kyte D, Price G, Valderas JM, Hjollund NH. Maximising the impact of patient reported outcome assessment for patients and society. BMJ 2019 Jan 24;364:k5267. [doi: 10.1136/bmj.k5267] [Medline: 30679170]

Abbreviations

CPR: Civil Personal Registration
DAS: disease activity score
EHR: electronic health record
EORTC: European Organization for Research and Treatment of Cancer
GDPR: General Data Protection Regulation
GP: general practicioner
HADS: Hospital Anxiety and Depression Scale
ICD: International Statistical Classification of Diseases and Related Health Problems
IPAQ: International Physical Activity Questionnaire
MFI: multidimensional fatigue inventory
PRO: patient-reported outcome
PRO-CTCAE: patient-reported outcome measure–Common Terminology Criteria for Adverse Events
SF: Short Form Health Survey
WHODAS: World Health Organization disability assessment schedule 2.0

Edited by G Eysenbach; submitted 14.08.19; peer-reviewed by L Ranandeh, M Wallwiener, V Pittet; comments to author 06.09.19; revised version received 09.09.19; accepted 20.09.19; published 30.09.19.

Please cite as:
Hjollund NHI
Fifteen Years' Use of Patient-Reported Outcome Measures at the Group and Patient Levels: Trend Analysis
J Med Internet Res 2019;21(9):e15856
URL: https://www.jmir.org/2019/9/e15856
doi:10.2196/15856
PMID:31573943

Review

The Antecedents and Consequences of Health Care Professional–Patient Online Interactions: Systematic Review

Lili Shang[1], MSCS; Meiyun Zuo[1], PhD; Dan Ma[1], BSc; Qinjun Yu[2], PhD

[1]Research Institute of Smart Senior Care, School of Information, Renmin University of China, Beijing, China
[2]School of Basic Medicine, Peking University Health Science Center, Beijing, China

Corresponding Author:
Meiyun Zuo, PhD
Research Institute of Smart Senior Care
School of Information
Renmin University of China
59 Zhongguancun Street
Haidian District
Beijing, 100872
China
Phone: 86 10 62511569
Email: zuomy@ruc.edu.cn

Abstract

Background: Online health care services effectively supplement traditional medical treatment. The development of online health care services depends on sustained interactions between health care professionals (HCPs) and patients. Therefore, it is necessary to understand the demands and gains of health care stakeholders in HCP-patient online interactions and determine an agenda for future work.

Objective: This study aims to present a systematic review of the antecedents and consequences of HCP-patient online interactions. It seeks to reach a better understanding of why HCPs and patients are willing to interact with each other online and what the consequences of HCP-patient online interactions are for health care stakeholders. Based on this, we intend to identify the gaps in existing studies and make recommendations for future research.

Methods: In accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, a systematic retrieval was carried out from the Web of Science, PubMed, and Scopus electronic databases. The search results were confined to those papers published in English between January 1, 2000 and June 30, 2018. Selected studies were then evaluated for quality; studies that did not meet quality criteria were excluded from further analysis. Findings of the reviewed studies related to our research questions were extracted and synthesized through inductive thematic analysis.

Results: A total of 8440 records were found after the initial search, 28 papers of which were selected for analysis. Accessibility to HCPs, self-management, and unmet needs were the main triggers for patients to participate in online interaction. For HCPs, patient education, career needs, and self-promotion were the major reasons why they took the online approach. There were several aspects of the consequences of HCP-patient online interactions on health care stakeholders. Consequences for patients included patient empowerment, health promotion, and acquisition of uncertain answers. Consequences for HCPs included social and economic returns, lack of control over their role, and gaining more appointments. HCP-patient online interactions also improved communication efficiency in offline settings and helped managers of online health care settings get a better understanding of patients' needs. Health care stakeholders have also encountered ethical and legal issues during online interaction.

Conclusions: Through a systematic review, we sought out the antecedents and consequences of HCP-patient online interactions to understand the triggers for HCPs and patients to participate and the consequences of participating. Potential future research topics are the influences on the chain of online interaction, specifications and principles of privacy design within online health care settings, and roles that sociodemographic and psychological characteristics play. Longitudinal studies and the adoption of text-mining method are worth encouraging. This paper is expected to contribute to the sustained progress of online health care settings.

(J Med Internet Res 2019;21(9):e13940) doi:10.2196/13940

KEYWORDS

health care professional–patient interaction; online health care service; online communication; eHealth; review

Introduction

Health care is closely related to people's life. Previously, interactions between health care professionals (HCPs) and patients occurred mainly in physical hospital settings. Today, eHealth has transformed the pattern of health care delivery with the development of information and communication technology [1]. Online interactions between HCPs and patients are increasingly playing a role in the provision of health care services. As a result, many patients have become dual-path inquirers (online consultation and offline physical access), and HCPs have become dual-path service providers (online and offline). An HCP-patient online interaction here refers to the exchange of health-related information via the internet [2] between HCPs and patients without in-person, face-to-face contact. In the process of interaction, patients can ask questions about health-related matters and HCPs help them by delivering health care advice and support through an online channel [3,4].

With the rise of the internet, online channels for HCPs to interact with patients have gradually emerged, including email, internet portals, social media, and online health communities. Based on the mentioned interactive channels, two forms of online health care interaction are classified, namely written communication and oral communication [5,6]. In the case of online written communication, the interaction between patients and HCPs is based on text messages and does not require the use of interactive instruments concurrently [7]. For online oral communication, it is voice-based and characterized by continuous interaction, allowing professionals and patients to interact uninterruptedly over a period of time. Driven by managers, online health care settings that combine different forms and channels are constructed at a cost-benefit tradeoff to facilitate HCP-patient interaction. Managers also function as decision makers in health care settings, such as formulating patterns of service provision, measuring investment, pricing, and so on. There are three main stakeholders in online health care settings, including normal users (ie, patients or their advocates, collectively called *patients* in this paper), who seek and obtain health care advice; doctors who deliver health care advice and support (*HCPs* in this paper); and managers who make decisions on the operation and development of online health care settings [8].

There is an increasing number of studies referring to HCP-patient online interactions from the perspective of patients or HCPs, such as what triggers patients to consult previously unknown doctors online [9], what determine doctors' reasons for the engagement in online interactions with patients [10], and what consequences the use of online channels has for participants [11]. The advancement of online health care settings is closely associated with sustained interaction between HCPs and patients [12], which is a manifestation of the socialization that indicates the activity level between members [13]. Thus, to define specific action items that help obtain an increased activity level, it is critical to identify what triggers the involvement of patients and HCPs in online interaction. Furthermore, the aim of providing services in the form of online interaction is primarily to improve health care [14]. At present, it engenders some consequences, including pros and cons that need to be further discussed and analyzed. To provide inspiration for the sustainable operation and effectiveness of online health care settings, it is necessary to examine this field thoroughly. Hence, this paper exhibits a systematic review of the antecedents and consequences of HCP-patient online interaction, and tries to answer the following research question: what are the antecedents and consequences of HCP-patient online interactions covered in the literature so far?

Ball and Lillis [1] conducted a literature review on the influence of eHealth on the HCP-patient relationship in 2001 before the rise of Web 2.0, which needs to be updated. After that, reviews related to our study, such as the impact of electronic communication on health care service provision [15], doctors' professional use of the internet and factors that encourage their usage [10], and the effects of social media use [11,16], all focus on specific aspects of this topic. To date, there has been no other comprehensive literature review of the antecedents and consequences of HCP-patient online interaction. It is necessary to re-examine recent studies and conduct a more thorough review. In this paper, we try to identify the antecedents and consequences of HCP-patient online interaction. The antecedent and consequence factors identified in this study could be applied to expound the facilitators and outcomes involved in HCP-patient online interaction. Therefore, we intend to identify the gaps in existing studies and make recommendations for future research. This review is also expected to contribute to the sustained progress of online health care settings.

Methods

Retrieval Strategy

In accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines [17], a systematic retrieval was carried out including the retrieval of the electronic databases Web of Science, PubMed, and Scopus. An additional manual retrieval was performed on the search of references and the identification of studies that may have been missed, as confirmed by experts.

The search terms included all possible keyword combinations from three aspects of population, channels for online interaction, and health care settings. The population covers both sides involved in online interaction, including patients and HCPs, and possible channels for online interaction in all types of health care settings were taken into account. The search terms were limited to those commonly used (see Multimedia Appendix 1 for search strings by each database). The search results were confined to those articles published in English between January 1, 2000 and June 30, 2018, because the earlier studies were mainly exploratory tests [15]. Journal papers and conference papers were retained, and previous reviews relevant to this study were also accepted into our study for their reference value. All study designs were included to find more comprehensive evidence to address our research question.

Selection of the Studies

The selection strategy was analyzed by the research team to reduce the possibility of bias. Two reviewers (LS and DM) worked together to conduct a comprehensive search and then excluded irrelevant literature. For those studies that they were not sure whether to exclude, two reviewers (MZ and QL) were assigned to deliberate it until they reached an agreement. To find the best evidence to address our research question, we defined explicit selection criteria for the inclusion of papers. Papers were included in this review if they studied (1) interactions between HCPs and patients via the internet and (2) the antecedents or consequences of HCP-patient online interaction.

According to the first criterion, the included papers needed to focus on online interaction, namely the bilateral exchange of health-related information via the internet. Therefore, studies focusing on unilateral information acquisition were excluded, such as the usage of physician-rating websites, the search for health information using search engines, and so on. In addition, the two roles of interaction had to be HCPs and patients. Thus, studies addressing peer-peer interaction, such as those just between patients or between HCPs, were not included in this study.

The second criterion indicates that any research that did not involve antecedents or consequences would be excluded, such as only comparing online interactions with offline interactions or focusing on data analysis methods of online interaction. Antecedents here refer to the factors that trigger patients and HCPs to participate in online interaction, whereas the factors impeding patients and HCPs to interact online were not taken into account. Consequences refer to the results brought on by online channel usage to health care stakeholders, including the benefits, risks, opportunities, challenges, and so forth.

Data Extraction, Evaluation, and Analysis

To ensure the reliability of data and analysis, the work in this part was carried out by two reviewers (LS and DM) independently; the inconsistencies were solved by a third reviewer (MZ). We created an information form for selected literature to aggregate relevant research data. Similar to the existing study [10], published information on studies, including the aim of the study, channels for online health care interaction, study design, and characteristics of respondents, were collected as general information.

We used the Critical Appraisal Skills Programme quality assessment tool [18], which consists of 10 items, to evaluate qualitative research and studies using mixed methods. Three broad themes needed to be considered when appraising, including the statement of findings, the validity, and contribution of the research. The assessment tools for quantitative studies in this review were adapted from the US National Heart, Lung, and Blood Institute quality assessment tool [19] and criteria for the evaluation of quantitative research proposed by Tan and Goonawardene [20]. The criteria consist of 14 items and are used to check the clarity of objectives, the selection of samples and methods, the reliability of results, and the outcomes of the research. Moreover, review papers published in peer-reviewed journals were determined to be qualified due to the more rigorous methodology compared with those published without peer review [21]. Finally, studies that did not meet these quality criteria were excluded from further analysis.

We extracted and synthesized findings of the reviewed studies based on our research question related to the antecedents and consequences of HCP-patient online interaction. The data analysis procedure was divided into three stages. In the first stage, all the antecedent and consequence factors in the studies that clearly answered our research question were identified and listed. In the second stage, for those selected studies that did not directly answer our research question, we studied the findings saved in our database to conduct an inductive thematic analysis [22]. Thematic analysis enabled us to synthesize research findings in a transparent manner and facilitate the emergence of new concepts. We coded the text to capture its meaning and grouped them to form descriptive themes by comparing similarities and differences between codes. Then, a synthesis of results that addressed our research question emerged by generating new interpretations based on descriptive themes. In the last stage, antecedent or consequent factors with similar meanings identified in the first two stages formed a synthesis by induction.

Results

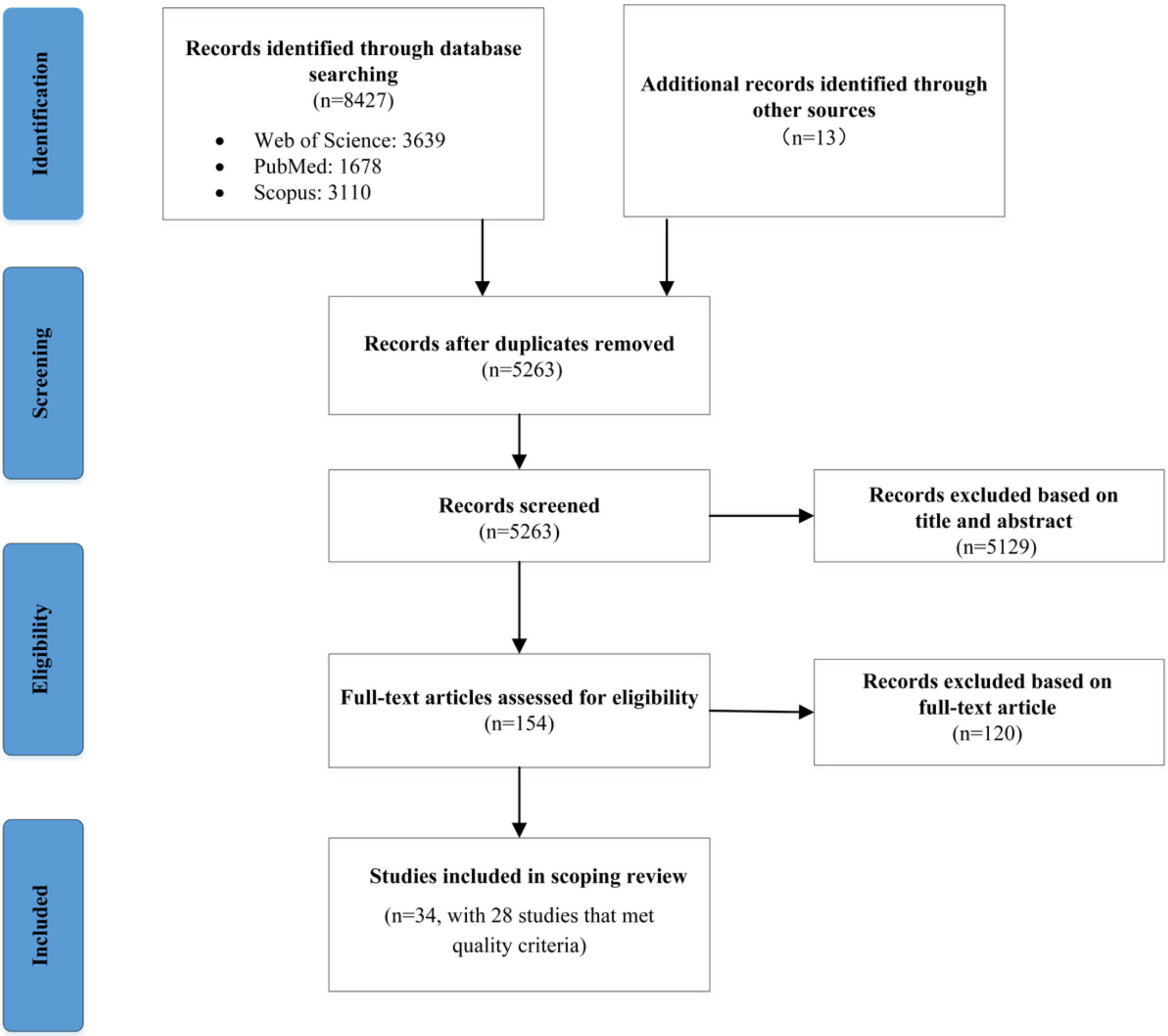
Overview

A total of 8440 records were found after the initial search. All records were exported and sorted, and all duplicates from diverse sources were removed, which reduced the number of records to 5263. Another 5129 papers were excluded after screening titles and abstracts. Of the remaining 154 full-text articles assessed for eligibility, 120 were excluded. Therefore, 34 studies were included in the scoping review. Of the 34 selected studies, 6 were excluded after quality assessment [2,4,16,23-25]. Finally, 28 studies met the quality criteria and were included in the review. Figure 1 is a flow diagram that depicts the selection process and results at each stage. A list of selected studies in chronological order of publication is shown in Multimedia Appendix 2. For the quality assessment of the included studies, see Multimedia Appendix 3.

Characteristics of the Reviewed Studies

Among the 28 included studies, 27 were journal papers and 1 was a conference paper. Research methods in the selected literature were quantitative study (n=15), qualitative study (n=8), mixed methods (n=1), and literature review (n=4). Channels for HCP-patient online interaction used in the targeted literature included email (n=3), online health community (n=14), social media (n=6), internet portal (n=4), desktop videoconferencing equipment (n=1), and other unspecified apps or health service websites (n=6). In most of these studies (n=27), HCPs interacted with patients mainly through written communication based on text messages. Only two studies involved online oral communication. Of the 28 articles, 7 explored the antecedents of online interaction, 11 discussed the consequences, and the rest studied both topics.

Figure 1. Flow diagram for the selection of literature.



The framework constructed from the emerging themes included (1) antecedents of online interactions between patients and HCPs and (2) consequences of online interactions on different health care stakeholders in general. For the antecedents of online interactions between patients and HCPs, we identified themes of antecedents from the perspective of two sides of online interaction: patients and HCPs. For the consequences of online interactions to health care stakeholders, we identified themes related to consequences of online interactions on three sides: patients, HCPs, and managers. Figure 2 is a visualization of the antecedents and consequences of HCP-patient online interactions identified in this paper.

Antecedents of Health Care Professional–Patient Online Interaction

Selected literature examined the factors that triggered patients and HCPs to participate in online interaction. A summary of these antecedents in descending order by the total number of related literature is shown in Table 1.

Antecedents From the Perspective of Patients

The reviewed studies showed that accessibility to HCPs, self-management, and unmet needs are the main factors that trigger patients to participate in online interaction.

Figure 2. A visualization of antecedents and consequences of health care professional–patient online interaction.

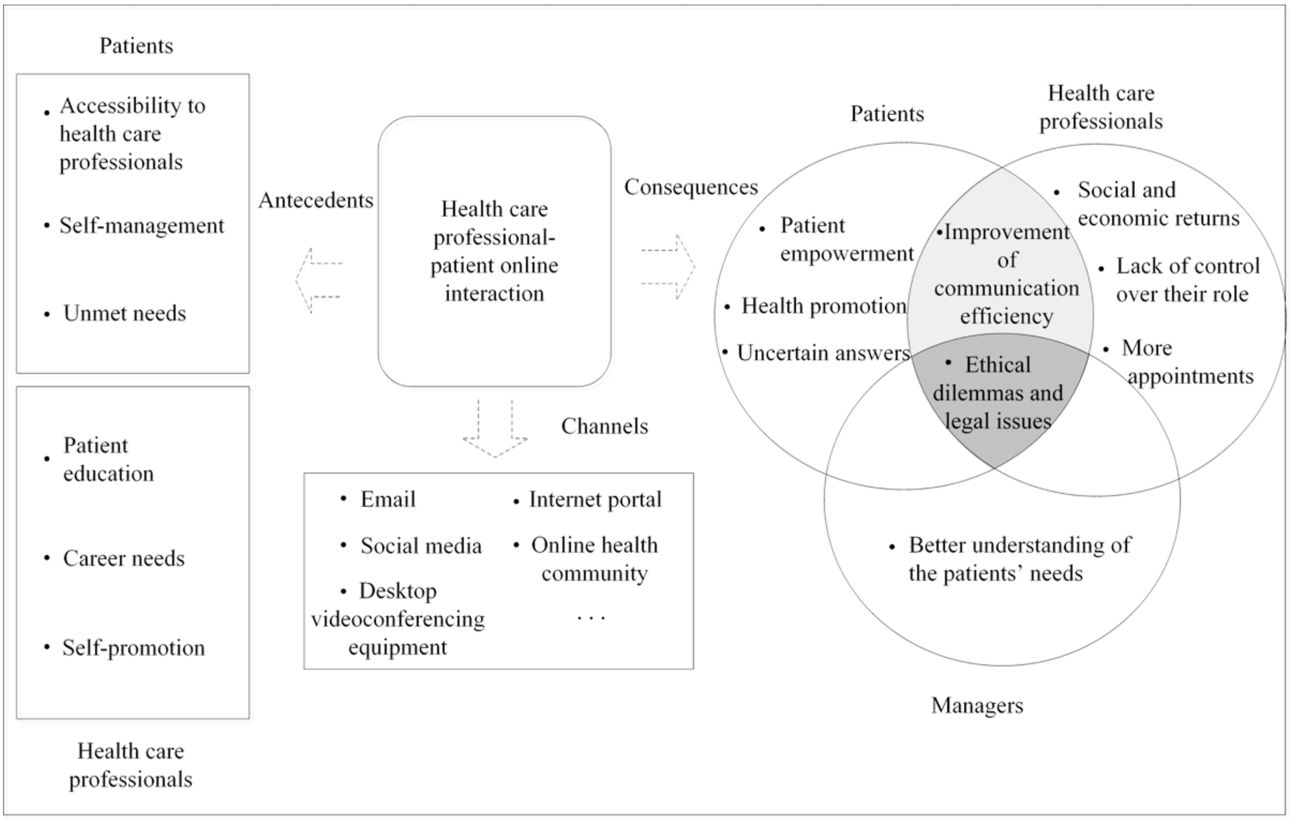


Table 1. Summary of the antecedents of health care professional–patient online interactions from the perspectives of patients and health care professionals.

Perspectives and antecedents	Description	Number of studies
Patients		
Accessibility to health care professionals	Availability for patients to access to health care professionals remotely at any time; support for anonymous access	5
Self-management	Primary evaluation of a medical problem before an offline appointment; preparing for future consultations in physical settings; a more effective way of health conditions self-management	5
Unmet needs	Discontent with health care previously received in offline settings; more alternative choices of health care services than offline settings	5
Health care professionals		
Patient education	An effective way to educate patients, especially patients with chronic diseases, about behavioral changes and drug adherence; a well-suited approach to deepen patients' awareness of a specific health condition	8
Career needs	An essential way to provide health care services to patients in the digital age	5
Self-promotion	Broader self-promotion to others to increase reputation and popularity	3

For many patients in the included studies, a major reason for choosing online consultation was that they had easier access to HCPs. The online channel offers a venue for patients to access doctors—even prestigious doctors—remotely [15] and discuss sensitive issues anonymously [26] at any time [9,27], where patients do not have to wait too long and can save money from not going to physical settings repeatedly [1]. In addition, the need for effective self-management of their conditions drove patients to make an online consultation. Patients, especially those with long-term conditions (eg, patients with diabetes), thought they should be responsible for their condition management, so they independently sought solutions and gained power to manage their own situations [28,29], such as preventive care [15], preparation for future consultations [9], and future treatment trajectory [30] in physical hospitals. Furthermore, the unmet needs of patients through offline channels were essential in determining their online counseling as a source of support. Patients expressed discontent with the health care they previously received offline [9,31] (eg, inadequate information, lack of trust, and hard to understand); therefore, they needed a second opinion from online HCPs [1,26,30].

Antecedents From the Perspective of Health Care Professionals

For HCPs, patient education, career needs, and self-promotion were found to be the main factors that triggered them to participate in online interaction. More specifically, professionals saw patient education as a major driver of the use of online channels as tools [1,27,32-35,37], such as encouraging health behavioral changes and drug adherence, raising public awareness about specific health conditions, and eliminating misunderstandings about certain diseases. Moreover, career needs acted as a trigger for HCPs to provide online health care services [10,32,36-37], which in return provided them with opportunities for future career development. The motivation of self-promotion was also a relatively important reason for participating in online interactions [32,37,38]. For HCPs, they wanted to present themselves to colleagues and patients through online interaction, and they considered it helpful to obtain better career development at their institutions.

Consequences of Health Care Professional–Patient Online Interactions for the Different Health Care Stakeholders

The selected literature indicated that HCP-patient online interactions affected online health care stakeholders, including patients, HCPs, and managers. A summary of these consequences in descending order according to the total number of related literature is shown in Table 2.

Table 2. Summary of the consequences of health care professional–patient online interactions on different health care stakeholders in general.

Stakeholders and consequences	Description	Number of studies
Patients		
Patient empowerment	More empowerment for patients	7
Health promotion	Improvement of the patients' health conditions; increase of patient adherence to treatments	4
Uncertain answers	Unavailability of clear answers due to insufficient patient information held by health care professionals	2
Health care professionals		
Social and economic returns	Acquisition of social returns, such as better reputation and greater popularity; acquisition of economic returns, such as online votes, bonus, likes, and electronic gifts from patients	4
Lack of control over their role	Leading to a lack of private time and life in disorder; service requests beyond one's professional capacity	4
More appointments	Potential opportunities to have more patient appointments	3
Managers		
Better understanding of the patients' needs	Useful insights for online health care service managers to understand the patients' needs, especially in terms of service delivery and pricing strategies	4
Patients, health care professionals		
Improvement of communication efficiency in offline settings	Improved ability of patients describing a specific health problem; increased face-to-face communication skills of professionals on account of the experience of formulating online text-based answers	3
All stakeholders		
Ethical dilemmas and legal issues	Unauthorized dissemination of personal information of health care professionals by patients; professional's uncertainty about the legitimacy whether or not to use online public information about patients and the rationality of online private communication with patients; ambiguity of information authorization for managers	9

Consequences of Online Interactions for Patients

Consequences of HCP-patient online interactions for patients included patient empowerment, health promotion, and uncertain answers. HCP-patient online interactions contributed to patient empowerment [1,26-28,34,39,40]. Online interactions seemed to operate in a different paradigm from offline communication because the professional was no longer at first place in the sequence and the problem was initiated by the patient, ultimately forming a patient-centered pattern [40]. Patients gained a variety of empowerment outcomes after they were provided with online health care service, such as the reinforced capacity to manage their physical conditions and increased acceptance of the disease, enhanced self-efficacy, and it promoted quality of interactions with professionals.

Studies demonstrated the role of HCP-patient online interactions in driving health promotion behaviors of patients, such as ending bad habits, developing a regular schedule, keeping a balanced diet, increasing physical activities, and maintaining a healthy mental state [36,34,41,42]. In particular, online health care interactions can effectually improve patient adherence, such as reminding patients of punctual intake of prescribed medicine, thus contributing to a successful treatment.

In an online health care environment, patients occasionally receive uncertain answers [26,39]. Due to insufficient patient information through online interaction, HCPs find it difficult to provide precise answers to patients and often advise them to make an appointment with an offline doctor. Furthermore, the information provided to professionals who need to respond is mainly text-based, and they seldom have the option to perform medical checks [26]. Therefore, assessing the adequacy or even veracity of the text-based information provided by patients may be impractical [39].

Consequences of Online Interactions for Health Care Professionals

Consequences for HCPs include social and economic returns, lack of control of their role, and gaining more appointments. The participating professionals gained returns from online channels [3,6,33,43]. Two notable ones for participating HCPs were social and economic returns. Specifically, by interacting with patients online, HCPs may obtain social returns such as increased reputation and popularity, and economic returns such as online votes, bonus, likes, and electronic gifts from patients [3].

Studies identified that HCP-patient interactions led to professionals feeling a lack of control of their role [11,15,27,37]. For example, Atanasova et al [37], in an in-depth semistructured interview, found all HCPs in online health communities reported experiences of overload and high burden. Professionals also experienced overcommitment to online interaction. They expressed that the demands of patients were beyond their capacity in the online context, and they were skeptical of the demand-and-supply patterns triggered by these health-related technological innovations [37].

Health care professional–patient online interactions provided professionals with opportunities to get more appointments [6,33,35]. Professionals felt that interactions in online settings enabled them to receive more patients. In addition, studies showed that HCP-patient online interactions could help professionals, especially those with low titles, get more appointments through accumulating more online experience [33].

Consequences of Online Interactions for Managers

Online HCP-patient interactions help managers better understand the needs of patients. Online health care interaction channels have accumulated rich data reflecting the trajectory of users' behavior. The included studies support evidence that data analysis helps managers understand patient needs [3,34,44,45]. For example, Yang et al [44] studied the influence of the process of delivering online services on patient satisfaction with professionals. Results showed that the response speed of professionals and the interaction frequency between professionals and patients—two important variables in the process of delivering online services—positively affected patient satisfaction. Wu and Lu [45] investigated the influence of online service provision and pricing on patient satisfaction. The results indicated HCPs who provide more services have higher patient satisfaction, and the relationship between service price and patient satisfaction is an inverted U shape. The price difference between different services provided by HCPs significantly decreases patient satisfaction.

Other Consequences of Online Interactions for Health Care Stakeholders

Health care professional–patient online interactions can also have an effect on the improvement of communication efficiency in offline settings [26,34,39]. Online interaction, normally used as a basis for further offline discussions with doctors, can help patients better understand and express their health status, problems, treatments, and remedies [34,39]. Developing responses to text-based health care consultations also promotes face-to-face communication skills for professionals [26]. With online interactive experience, they learn more about how to formulate medical terms clearly and how to properly refer to internet resources to provide a wide range of medical information [26].

In the context of online health care interaction, all health care stakeholders have encountered ethical dilemmas and legal issues [11,26,27,32,36-35,46]. For patients, in addition to giving online appraisals to professionals, they may also spread unauthorized personal information about professionals on the internet [46]. For HCPs, they are confused about how to respond to online ethical dilemmas [37]. Moreover, there is no general agreement on the appropriateness of using patient's publicly available online information (eg, patients' private life updates on their Facebook pages) to assist them in treatment. Professionals also argue that the separation of their professional and private lives should be considered. Whether private communication between professionals and patients online should be encouraged remains to be seen. The most common way that professionals deal with undesired private communication is to change the privacy settings of their online interactive apps, followed by ignoring a friend request from their patients [35]. For managers, they are bound to face the threat posed by the disclosure of private information when empowering patients to deliver information about professionals to public [46]. In some cases, profiles of HCPs have been developed and displayed on third-party review sites without their personal participation and confirmation.

Discussion

Principal Findings

This study presents the antecedents and consequences of HCP-patient online interactions covered in the existing literature. In the reviewed studies, the most reported channels of online professional–patient interactions were online health communities and social media, especially in recent years. Interactions between HCPs and patients were based mainly on written communication, whereas a few studies were based on oral communication. The findings of this study provide evidence that the unique advantages of online health care settings over offline ones drive patients and HCPs to participate in online interaction. In addition, HCP-patient online interactions do have positive effects on patients' health care, although it also has limitations. For managers, it provides opportunities and poses challenges for them.

Antecedents that trigger patients to participate in online interactions were identified. The findings of this study show that accessibility to HCPs, self-management, and unmet needs in offline channels are the main factors that trigger patients to participate in online interaction, echoing the views of previous reviews [1,15]. These factors also imply the demands of patients for health care services and have the potential to provide insights into how to design online health settings to meet patients' needs. For example, it is necessary to attract more well-known doctors who are difficult to make an appointment with in offline settings to provide online health care services so that patients can have access to them. Also, managers of online settings could consider defining functions and processes that help patients, especially those with chronic conditions, perform self-management.

Antecedents that encourage HCPs to interact online include both altruistic factors (patient education), which have been determined in an existing review [1], and egoistic factors (career needs and self-promotion), which have been updated in this study. When it comes to altruistic factors, the use of an online channel can be a favorable way to provide patient education for HCPs. Egoistic factors include career needs and self-promotion. Consistent with previous findings [10], HCPs consider the use of online channels as a part of their career development. The newly identified antecedent, self-promotion, suggests HCPs want to take advantage of online approaches to present themselves to colleagues and patients to help their future. Given this, it is necessary for online health care settings to consider incorporating functions or items that facilitate self-presentation of HCPs.

Compared with the consequences of online interactions for patients and HCPs presented in previous reviews, this paper found new positive and negative effects. Positive effects for patients include patient empowerment and health promotion. However, uncertain answers from online HCPs, as a passive consequence, imply patients are often dissatisfied with the vague suggestions from online HCPs, and then revisit their private doctors. The positive consequences of online interactions identified in reviewed studies for HCPs are social and economic returns and more appointments from patients. The negative side is that HCPs feel a lack of control over their role because the appointments and interactions with patients remotely can occur at any time. The improvement of communication efficiency in offline settings was found to have a positive impact on online interactions for both patients and HCPs. In general, compared with previous reviews, the newly identified consequence factors in this review include uncertain answers, social and economic returns, more appointments, and the improvement of communication efficiency in offline settings. Further, we found more positive evidence of the consequences of online interactions than negative ones.

The consequences of online HCP-patient interactions for managers were also examined in the reviewed studies. Specifically, in the process of online interaction, a large amount of potentially useful data are produced, providing opportunities for managers to have a better understanding of patients' needs. Meanwhile, ethical and legal issues are emerging, which pose challenges for managers. In the future, it will be necessary for managers to take advantage of opportunities and address the previously mentioned challenges to enable cost-benefit management.

Limitations of the Review

To obtain more comprehensive evidence of antecedent and consequence factors, this review included diverse and heterogeneous studies with different research analyses and methods. The channels and forms of online interactions are changing, and the motivation of people to participate, and the consequences of participation, may also evolve. However, the heterogeneity of studies is a challenge to reveal the dynamics of the antecedent and consequence factors. If studies adopt a uniform format, comparisons and trends could be more accurately identified.

This review focuses on identifying the antecedents and consequences of online interactions separately, without centering on the influencing chain. It should be noted that factors impeding patients and HCPs from interacting online were not within the scope of our review; therefore, no salient themes were identified in the coding process of the included literature. It would be more comprehensive to incorporate these factors. The entire influencing chain from the antecedents to the online interaction (feelings, thoughts, problems, advantages, and disadvantages) and further on to the consequences would be interesting for a literature review.

Recommendations for Future Studies

Recommendations for future research are from the aspects of potential topics and methods. There are three potential research topics deserving further exploration. First, there is a distinct lack of research on systematically investigating the influencing chain of online interactions that may explain how online health care settings work. It will be meaningful for future original research to uncover the influencing chain of online interaction.

Second, the specifications and principles of privacy design in the context of online health care settings should be considered in future research. With the emergence of online channels, the paradigm of interaction between HCPs and patients has changed, and the guidelines for interaction should also change [46]. However, professionals sometimes still rely on experiences and intuition to interact, and patients may be confused about how to express themselves effectively to get definite answers. Next, most studies have mentioned privacy considerations for health care stakeholders while using these technologies to communicate on social media [47], but only a few feasible solutions have been provided [48]. Consequently, it is necessary to probe the legal and ethical problems in current situations and explore specifications of consultation and service delivery.

Third, the role of sociodemographic and psychological characteristics in existing studies has been underestimated. Demand for online health care interaction of professionals or patients with different characteristics may vary. For example, this type of interaction might be favored by surgeons because of the more effective means of preoperative preparation and the short-term follow-up after surgery it provides [8,49]. In addition, patients with high-risk diseases show higher sensitivity to the process of online interactions [44], and elderly patients' demands for online health care interactions may be different from that of young people [50]. Social presence, the degree to which a person is perceived as “real” and connected to others in the process of communicating through the media [51], and the health-related information-seeking personality, a need for cognition and information more than others when making decisions, may also influence the use of online interaction channels [52]. Therefore, future research could aim to explore how these characteristics play a role in this topic.

The existing studies can be extended from two aspects of research methods. First, more longitudinal studies on HCP-patient interactions should be performed in the future. Previous studies were mainly conducted by cross-sectional design to investigate the effect of online interactions on patients' health conditions, without reflecting the dynamic effects of online interactions [53]. Therefore, long-term observations should be encouraged to investigate the dynamics of the consequences of online interaction. In addition, the text-mining method is worth encouraging to analyze text messages generated by online interaction. Through text mining, studies have characterized communication patterns in the process of health information seeking [54] and identified influential users in online health communities [55], yet there is a lack of analysis of the interactive content itself. Therefore, future research can attempt to carry out text analysis of the content generated in the course of interactions and seek out more precise antecedent or consequence factors of online interactions to help guide the design of more effective patient-centered online health care settings [56].

Acknowledgments

This work was supported in part by Key Projects of Philosophy and Social Sciences Research of Chinese Ministry of Education under grant 19JZD021, the National Natural Science Foundation of China under grant 71771210, and the Beijing Natural Science Foundation under grant 9182008.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings by each database.
[PDF File (Adobe PDF File)101 KB - jmir_v21i9e13940_app1.pdf]

Multimedia Appendix 2
Summary of selected literature.
[PDF File (Adobe PDF File)216 KB - jmir_v21i9e13940_app2.pdf]

Multimedia Appendix 3
Quality assessment of included studies.
[PDF File (Adobe PDF File)231 KB - jmir_v21i9e13940_app3.pdf]

References

1. Ball MJ, Lillis J. E-health: transforming the physician/patient relationship. Int J Med Inform 2001 Apr;61(1):1-10. [Medline: 11248599]
2. Farnan JM, Snyder Sulmasy L, Worster BK, Chaudhry HJ, Rhyne JA, Arora VM. Online medical professionalism: patient and public relationships: policy statement from the American College of Physicians and the Federation of State Medical Boards. Ann Intern Med 2013 Apr 16;158(8):620-627. [doi: 10.7326/0003-4819-158-8-201304160-00100] [Medline: 23579867]
3. Guo S, Guo X, Fang Y, Vogel D. How doctors gain social and economic returns in online health-care communities: a professional capital perspective. J Manage Inform Syst 2017 Aug 17;34(2):487-519. [doi: 10.1080/07421222.2017.1334480]
4. Guo X, Guo S, Vogel D, Li Y. Online healthcare community interaction dynamics. J Manage Sci Eng 2016;1(1):58-74. [doi: 10.3724/SP.J.1383.101004]
5. de Jong CC, Ros WJ, Schrijvers G. The effects on health behavior and health outcomes of internet-based asynchronous communication between health providers and patients with a chronic condition: a systematic review. J Med Internet Res 2014 Jan 16;16(1):e19 [FREE Full text] [doi: 10.2196/jmir.3000] [Medline: 24434570]
6. Wu H, Lu N. Online written consultation, telephone consultation and offline appointment: an examination of the channel effect in online health communities. Int J Med Inform 2017 Nov;107:107-119. [doi: 10.1016/j.ijmedinf.2017.08.009] [Medline: 29029686]
7. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive health communication applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
8. Gruzd A, Haythornthwaite C. Enabling community through social media. J Med Internet Res 2013 Oct;15(10):e248. [doi: 10.2196/jmir.2796] [Medline: 24176835]

9. Umefjord G, Petersson G, Hamberg K. Reasons for consulting a doctor on the internet: web survey of users of an Ask the Doctor service. J Med Internet Res 2003 Oct 22;5(4):e26 [FREE Full text] [doi: 10.2196/jmir.5.4.e26] [Medline: 14713654]
10. Masters K. For what purpose and reasons do doctors use the internet: a systematic review. Int J Med Inform 2008 Jan;77(1):4-16. [doi: 10.1016/j.ijmedinf.2006.10.002] [Medline: 17137833]
11. Gholami-Kordkheili F, Wild V, Strech D. The impact of social media on medical professionalism: a systematic qualitative review of challenges and opportunities. J Med Internet Res 2013 Aug;15(8):e184 [FREE Full text] [doi: 10.2196/jmir.2708] [Medline: 23985172]
12. Young C. Community management that works: how to build and sustain a thriving online health community. J Med Internet Res 2013 Jun 11;15(6):e119 [FREE Full text] [doi: 10.2196/jmir.2501] [Medline: 23759312]
13. Preece J. Sociability and usability in online communities: determining and measuring success. Behav Inform Technol 2010 Nov 08;20(5):347-356. [doi: 10.1080/01449290110084683]
14. Coiera E. Medical informatics. BMJ 1995 May 27;310(6991):1381-1387 [FREE Full text] [doi: 10.1136/bmj.310.6991.1381] [Medline: 7787545]
15. McGeady D, Kujala J, Ilvonen K. The impact of patient-physician web messaging on healthcare service provision. Int J Med Inform 2008 Jan;77(1):17-23. [doi: 10.1016/j.ijmedinf.2006.11.004] [Medline: 17188564]
16. Courtney K. The use of social media in healthcare: organizational, clinical, and patient perspectives. Stud Health Technol Inform 2013;183:244-248. [doi: 10.3233/978-1-61499-203-5-244] [Medline: 23388291]
17. Knobloch K, Yoon U, Vogt PM. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and publication bias. J Craniomaxillofac Surg 2011 Mar;39(2):91-92. [doi: 10.1016/j.jcms.2010.11.001] [Medline: 21145753]
18. Critical Appraisal Skills Programme. 2018. CASP tools & checklists URL: http://www.casp-uk.net/casp-tools-checklists [accessed 2019-04-15]
19. National Heart, Lung, and Blood Institute. 2018. Quality assessment tool for observational cohort cross-sectional studies URL: https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/cohort [accessed 2019-04-15]
20. Tan SS, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. J Med Internet Res 2017 Jan 19;19(1):e9 [FREE Full text] [doi: 10.2196/jmir.5729] [Medline: 28104579]
21. Rochon PA, Bero LA, Bay AM, Gold JL, Dergal JM, Binns MA, et al. Comparison of review articles published in peer-reviewed and throwaway journals. JAMA 2002 Jun 05;287(21):2853-2856. [doi: 10.1001/jama.287.21.2853] [Medline: 12038932]
22. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008;8:45 [FREE Full text] [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
23. Dizon DS, Graham D, Thompson MA, Johnson LJ, Johnston C, Fisch MJ, et al. Practical guidance: the use of social media in oncology practice. J Oncol Pract 2012 Sep;8(5):e114-e124 [FREE Full text] [doi: 10.1200/JOP.2012.000610] [Medline: 23277774]
24. George DR, Rovniak LS, Kraschnewski JL. Dangers and opportunities for social media in medicine. Clin Obstet Gynecol 2013 Sep;56(3):453-462 [FREE Full text] [doi: 10.1097/GRF.0b013e318297dc38] [Medline: 23903375]
25. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520 [FREE Full text] [Medline: 25083128]
26. Björk AB, Hillborg H, Augutis M, Umefjord G. Evolving techniques in text-based medical consultation-physicians' long-term experiences at an Ask the doctor service. Int J Med Inform 2017 Sep;105:83-88. [doi: 10.1016/j.ijmedinf.2017.05.013] [Medline: 28750914]
27. Atanasova S, Kamin T, Petrič G. The benefits and challenges of online professional-patient interaction: comparing views between users and health professional moderators in an online health community. Comput Hum Behav 2018 Jun;83:106-118. [doi: 10.1016/j.chb.2018.01.031]
28. Hewitt-Taylor J, Bond CS. What e-patients want from the doctor-patient relationship: content analysis of posts on discussion boards. J Med Internet Res 2012 Nov 08;14(6):e155 [FREE Full text] [doi: 10.2196/jmir.2068] [Medline: 23137788]
29. Dixon RF, Stahl JE. A randomized trial of virtual visits in a general medicine practice. J Telemed Telecare 2009 Apr;15(3):115-117. [doi: 10.1258/jtt.2009.003003] [Medline: 19364890]
30. Schook RM, Linssen C, Schramel FM, Festen J, Lammers E, Smit EF, et al. Why do patients and caregivers seek answers from the Internet and online lung specialists? A qualitative study. J Med Internet Res 2014 Feb;16(2):e37 [FREE Full text] [doi: 10.2196/jmir.2842] [Medline: 24496139]
31. Lee SY, Hawkins R. Why do patients seek an alternative channel? The effects of unmet needs on patients' health-related Internet use. J Health Commun 2010 Mar;15(2):152-166. [doi: 10.1080/10810730903528033] [Medline: 20390984]
32. Campbell BC, Craig CM. Social media and health: current and future healthcare provider perspectives. J Contemp Med Edu 2014;2(2):128-133. [doi: 10.5455/jcme.20140515123200]
33. Liu X, Guo X, Wu H, Vogel D. Doctor's effort influence on online reputation and popularity. In: Proceedings of Smart Health 2014. 2014 Jul Presented at: ICSH 2014 International Conference on Smart Health; July 10-11, 2014; Beijing p. 111-126 URL: https://www.researchgate.net/profile/Xitong_Guo/publication/

295038968_Doctor%27s_Effort_Influence_on_Online_Reputation_and_Popularity/links/56c680de08ae408dfe4d4e48/Doctors-Effort-Influence-on-Online-Reputation-and-Popularity.pdf [doi: 10.1007/978-3-319-08416-9_11]

34. Haluza D, Jungwirth D. ICT and the future of health care: aspects of doctor-patient communication. Int J Technol Assess Health Care 2014 Sep 02;30(3):298-305. [doi: 10.1017/S0266462314000294] [Medline: 25308693]
35. Daniel F, Jabak S, Sasso R, Chamoun Y, Tamim H. Patient-physician communication in the era of mobile phones and social media apps: cross-sectional observational study on Lebanese physicians' perceptions and attitudes. JMIR Med Inform 2018 Apr 06;6(2):e18 [FREE Full text] [doi: 10.2196/medinform.8895] [Medline: 29625955]
36. Haluza D, Jungwirth D. ICT and the future of health care: aspects of health promotion. Int J Med Inform 2015 Jan;84(1):48-57. [doi: 10.1016/j.ijmedinf.2014.09.005] [Medline: 25293532]
37. Atanasova S, Kamin T, Petrič G. Exploring the benefits and challenges of health professionals' participation in online health communities: Emergence of (dis)empowerment processes and outcomes. Int J Med Inform 2017 Feb;98:13-21. [doi: 10.1016/j.ijmedinf.2016.11.005] [Medline: 28034408]
38. Zhang X, Liu S, Deng Z, Chen X. Knowledge sharing motivations in online health communities: a comparative study of health professionals and normal users. Comput Hum Behav 2017;75:797-810. [doi: 10.1016/j.chb.2017.06.028]
39. Zummo ML. Exploring web-mediated communication: a genre-based linguistic study for new patterns of doctor-patient interaction in online environment. Commun Med 2015;12(2-3):187-198. [Medline: 29048146]
40. Petrič G, Atanasova S, Kamin T. Impact of social processes in online health communities on patient empowerment in relationship with the physician: emergence of functional and dysfunctional empowerment. J Med Internet Res 2017 Mar 13;19(3):e74 [FREE Full text] [doi: 10.2196/jmir.7002] [Medline: 28288953]
41. Wu T, Deng Z, Feng Z, Gaskin DJ, Zhang D, Wang R. The effect of doctor-consumer interaction on social media on consumers' health behaviors: cross-sectional study. J Med Internet Res 2018 Feb 28;20(2):e73 [FREE Full text] [doi: 10.2196/jmir.9003] [Medline: 29490892]
42. Wu T, Deng Z, Zhang D, Buchanan P, Zha D, Wang R. Seeking and using intention of health information from doctors in social media: the effect of doctor-consumer interaction. Int J Med Inform 2018 Jul;115:106-113. [doi: 10.1016/j.ijmedinf.2018.04.009] [Medline: 29779712]
43. Guo S, Guo X, Zhang X, Vogel D. Doctor–patient relationship strength’s impact in an online healthcare community. Inform Technol Dev 2017 Mar 09;24(2):279-300. [doi: 10.1080/02681102.2017.1283287]
44. Yang H, Guo X, Wu T. Exploring the influence of the online physician service delivery process on patient satisfaction. Decis Support Syst 2015 Oct;78:113-121. [doi: 10.1016/j.dss.2015.05.006]
45. Wu H, Lu N. Service provision, pricing, and patient satisfaction in online health communities. Int J Med Inform 2018 Feb;110:77-89. [doi: 10.1016/j.ijmedinf.2017.11.009] [Medline: 29331257]
46. Brown J, Ryan C, Harris A. How doctors view and use social media: a national survey. J Med Internet Res 2014 Dec 02;16(12):e267 [FREE Full text] [doi: 10.2196/jmir.3589] [Medline: 25470407]
47. Gabbard GO, Kassaw KA, Perez-Garcia G. Professional boundaries in the era of the Internet. Acad Psychiatry 2011 May;35(3):168-174. [doi: 10.1176/appi.ap.35.3.168] [Medline: 21602438]
48. Chretien KC, Kind T. Social media and clinical care: ethical, professional, and social implications. Circulation 2013 Mar 02;127(13):1413-1421. [doi: 10.1161/CIRCULATIONAHA.112.128017] [Medline: 23547180]
49. Sculco PK, McLawhorn AS, Fehring KA, De Martino I. The future of social media in orthopedic surgery. Curr Rev Musculoskelet Med 2017 Jun;10(2):278-279 [FREE Full text] [doi: 10.1007/s12178-017-9412-9] [Medline: 28456946]
50. Xie B, Wang M, Feldman R, Zhou L. Exploring older and younger adults' preferences for health information and participation in decision making using the Health Information Wants Questionnaire (HIWQ). Health Expect 2014 Dec;17(6):795-808 [FREE Full text] [doi: 10.1111/j.1369-7625.2012.00804.x] [Medline: 22783800]
51. Short J, Williams E, Christie B. The social psychology of telecommunications. Contemp Sociol 1978 Jan;7(1):32-33. [doi: 10.2307/2065899]
52. Roettl J, Bidmon S, Terlutter R. What predicts patients' willingness to undergo online treatment and pay for online treatment? Results from a web-based survey to investigate the changing patient-physician relationship. J Med Internet Res 2016 Feb 04;18(2):e32 [FREE Full text] [doi: 10.2196/jmir.5244] [Medline: 26846162]
53. Conceição EM, Machado PP, Vaz AR, Pinto-Bastos A, Ramalho S, Silva C, et al. APOLO-Bari, an internet-based program for longitudinal support of bariatric surgery patients: study protocol for a randomized controlled trial. Trials 2016 Mar 01;17(1):114 [FREE Full text] [doi: 10.1186/s13063-016-1246-z] [Medline: 26927479]
54. Chiu MP. An investigation of the questions posted on medical consultation websites. Health Info Libr J 2016 Dec;33(4):283-294. [doi: 10.1111/hir.12153] [Medline: 27380768]
55. Zhao K, Yen J, Greer G, Qiu B, Mitra P, Portier K. Finding influential users of online health communities: a new metric based on sentiment influence. J Am Med Inform Assoc 2014 Oct;21(e2):e212-e218 [FREE Full text] [doi: 10.1136/amiajnl-2013-002282] [Medline: 24449805]
56. Wang X, Zhao K, Street N. Analyzing and predicting user participations in online health communities: a social support perspective. J Med Internet Res 2017 Apr 24;19(4):e130 [FREE Full text] [doi: 10.2196/jmir.6834] [Medline: 28438725]

Abbreviations

HCP: health care professional

Edited by N Zary; submitted 08.03.19; peer-reviewed by H Atherton, S Bidmon; comments to author 12.04.19; revised version received 05.06.19; accepted 31.07.19; published 25.09.19.

Please cite as:
Shang L, Zuo M, Ma D, Yu Q
The Antecedents and Consequences of Health Care Professional–Patient Online Interactions: Systematic Review
J Med Internet Res 2019;21(9):e13940
URL: http://www.jmir.org/2019/9/e13940/
doi:10.2196/13940
PMID:31573908

Original Paper

Digital Health Transformation of Integrated Care in Europe: Overarching Analysis of 17 Integrated Care Programs

Erik Baltaxe[1,2], MD; Thomas Czypionka[3], MD, PhD; Markus Kraus[3], PhD; Miriam Reiss[3], MSc; Jan Erik Askildsen[4], PhD; Renata Grenkovic[5], MSc; Tord Skogedal Lindén[6], PhD; János György Pitter[7], MD, PhD; Maureen Rutten-van Molken[8,9], PhD; Oscar Solans[10], MD; Jonathan Stokes[11], PhD; Verena Struckmann[12], MSc; Josep Roca[1,2], MD, PhD; Isaac Cano[1,2], PhD

[1]Hospital Clinic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Universitat de Barcelona, Barcelona, Spain
[2]Center for Biomedical Network Research in Respiratory Diseases, Madrid, Spain
[3]Institute for Advanced Studies, Vienna, Austria
[4]Department of Economics, University of Bergen, Bergen, Norway
[5]Agency for Quality and Accreditation in Health Care and Social Welfare, Zagreb, Croatia
[6]Department of Social Science, NORCE Norwegian Research Centre, Bergen, Norway
[7]Syreon Research Institute, Budapest, Hungary
[8]School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, Netherlands
[9]Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, Netherlands
[10]Oficina eSalut, Departament de Salut, Generalitat de Catalunya, Barcelona, Spain
[11]Health Organisation, Policy, and Economics, Centre for Primary Care and Health Services Research, University of Manchester, Manchester, United Kingdom
[12]Department of Health Care Management, Technische Universität Berlin, Berlin, Germany

Corresponding Author:
Isaac Cano, PhD
Hospital Clinic de Barcelona
Institut d'Investigacions Biomèdiques August Pi i Sunyer
Universitat de Barcelona
Carrer del Vallarroel 170
Barcelona, 08036
Spain
Phone: 34 932275540
Fax: 34 932275455
Email: ISCANO@clinic.cat

Abstract

Background: Digital health tools comprise a wide range of technologies to support health processes. The potential of these technologies to effectively support health care transformation is widely accepted. However, wide scale implementation is uneven among countries and regions. Identification of common factors facilitating and hampering the implementation process may be useful for future policy recommendations.

Objective: The aim of this study was to analyze the implementation of digital health tools to support health care and social care services, as well as to facilitate the longitudinal assessment of these services, in 17 selected integrated chronic care (ICC) programs from 8 European countries.

Methods: A program analysis based on thick descriptions—including document examinations and semistructured interviews with relevant stakeholders—of ICC programs in Austria, Croatia, Germany, Hungary, the Netherlands, Norway, Spain, and the United Kingdom was performed. A total of 233 stakeholders (ie, professionals, providers, patients, carers, and policymakers) were interviewed from November 2014 to September 2016. The overarching analysis focused on the use of digital health tools and program assessment strategies.

Results: Supporting digital health tools are implemented in all countries, but different levels of maturity were observed among the programs. Only few ICC programs have well-established strategies for a comprehensive longitudinal assessment. There is a strong relationship between maturity of digital health and proper evaluation strategies of integrated care.

Conclusions: Notwithstanding the heterogeneity of the results across countries, most programs aim to evolve toward a digital transformation of integrated care, including implementation of comprehensive assessment strategies. It is widely accepted that the evolution of digital health tools alongside clear policies toward their adoption will facilitate regional uptake and scale-up of services with embedded digital health tools.

(J Med Internet Res 2019;21(9):e14956) doi:10.2196/14956

KEYWORDS

program evaluation; chronic patients; eHealth; elderly; integrated care; social support; telemonitoring; information and communication technology

Introduction

Background

Digital health (eHealth) tools have been proposed to improve access to health care services, enhance care co-ordination and integration, enable self-management, support decision-making, enable monitoring, perform risk analysis, and facilitate proactive interventions [1]. It is within this context that the European Commission has defined eHealth as follows:

> *The use of Information and Communication Technologies in health products, services and processes, combined with organizational change in health care systems and new skills, in order to improve health of citizens, efficiency and productivity in health care delivery, and the economic and social value of health [2]*

The implementation of digital health tools has constituted an area of major research and innovation in the past years [1,4-7]. For example, the impact of electronic health records on health care quality has been explored by Campanella et al who showed improvement in health care quality in terms of guideline adherence and time efficiency while reducing medication errors [8]. In parallel, most countries have recognized the value of patient portals [9-11]. The use of health monitoring devices in the general population is also increasing, with a broad spectrum of sophistication, including examples like integration between artificial intelligence and monitoring in a single device (ie, cardiac rhythm analysis) [12]. Moreover, the use of open-source algorithms for subject-specific as well as population-based risk prediction has been reported [13,14]. Some of these technologies are already in use for the management of chronic multimorbid patients [15-20].

It is currently accepted that eHealth tools can be particularly useful to support Integrated Chronic Care (ICC) programs for patients with multimorbidity, that is, co-occurrence of 2 or more chronic disorders within 1 individual [21]. Insights from the ICARE4EU project [6] concluded that eHealth improves care integration and management processes, but the project identified that inadequate funding mechanisms, poor interoperability, and inadequate technological support represent major barriers for adoption of technologically supported ICC. In fact, it is acknowledged that the takeoff of digital health tools to support ICC is progressing rather slowly. Also, regulatory aspects are still a concern [22-24] to achieve a proper balance between preservation of individual privacy and the need for health data sharing [25], as well as the increasing demand for health data analytics.

The Conceptual Framework

The Sustainable intEgrated care modeLs for multimorbidity: delivery, Financing, and performancE (SELFIE) Horizon 2020 project [26] aims to produce evidence and applicable policy advice on ICC programs for people with multimorbidity. Within the project aims, the SELFIE conceptual framework [27] was developed. It comprises 6 core components of integrated care systems adapted from the World Health Organization, namely: (1) *Service delivery*, (2) *Leadership and governance*, (3) *Workforce*, (4) *Financing*, (5) *Technologies and medical products*, and (6) *Information and research*.

This paper focuses on 2 elements out of the 6 components of the SELFIE conceptual framework [27], which refer to the enabling role of digital health tools (*Technologies and medical products*) and assessment of ICC programs (*Information and research*). For each of these 2 components, the 3 levels of the SELFIE conceptual framework (micro, meso, and macro) were taken into account. The micro level is where the individual with multimorbidity interacts with care professionals and informal caregivers. The meso level relates to the organizational level and the institutional setup of providers. Finally, the macro level includes legislations, governance, policies, and system-wide changes at the national and international level.

For *Technologies and medical products*, the SELFIE framework stresses the need for digital health tools to be widely available and user-friendly to provide robust support to the care processes. At a micro level, the use of technology (eg, electronic medical records [EMR] and patient portals) can be a facilitator of collaborative care if tailored to the needs of the patient with multimorbidity. At a meso level, a shared information system (eg, EMR including shared care plans) among multiple providers and care settings can greatly facilitate communication, person-centeredness, personalized care, and care co-ordination. Finally, at a macro level, nationwide and international policies that foster technological development and innovation most likely would benefit from both implementation and continuous assessment of ICC programs for multimorbidity.

For *Information and research*, the project stresses the successful use of collected data from digital health tools for a 3-fold objective, namely: (1) Population health management; (2) Enhanced subject-specific health care delivery; and (3) Comprehensive assessment of ICC programs. At a micro level, currently collected individual-level data (eg, patient journey record) can effectively be used in the care process for individual risk prediction. At a meso level, shared information systems may further be used for service selection both at individual and group level (eg, triage systems and clinical predictive modelling) to strengthen the evidence base of complex integrated care interventions, as well as to develop indicators particularly relevant for the care of patients with multimorbidity. Alongside data ownership at a meso level, privacy and data protection legislation is an important consideration at macro level.

Aims

It is well accepted that the existence of an important gap between the way in which the role of digital tools is understood and the effective uptake of digital transformation by the different stakeholders in the health care systems at European level. To enable the real implementation and scale-up of the digital health tools with all their potential, we undertook this overarching analysis. Our study aims to synthesize the experiences, views, and opinions (including barriers and enabling factors) of the stakeholders and their impact on the care process, as well as the role and desirable future developments of digital health tools, to foster transformation of health care systems toward sustainability by enhancing management of patients with multimorbidity. A second aim is to characterize the different programs with respect to maturity of their supporting digital health tools and the level of assessment of the ICC program.

Here, we present the results of an overarching analysis of the Thick Descriptions [28] of 17 promising ICC programs selected by the SELFIE project across 8 European countries: Austria, Croatia, Germany, Hungary, the Netherlands, Norway, Spain, and the United Kingdom. Within the overarching analysis, we focus on the different aspects of implementation of digital health tools supporting services and facilitating assessment strategies. This will lead to future directions defining how digital support can contribute to scale-up and evaluation of integrated care services. These services should focus on patient-centered health care provision with dynamic evaluation of technology-enabled integrated care programs, without compromising patient privacy.

Methods

Study Design

To select the 17 ICC programs, each country participating in the SELFIE project [29] applied a search strategy using the findings from an international scoping review, national publications on previous and on-going programs and projects, and consultation with national experts and networks. Details on the process of selection of the programs, as well as the list of the 17 selected programs per country, are reported in Multimedia Appendix 1. The 17 programs were grouped into 4 categories: (1) population health management programs (n=5); (2) frail elderly programs (n=6); (3) programs for individuals at the end-of-life and oncology patients (n=3); and (4) programs for vulnerable individuals who face problems in multiple life domains, like health, housing, and financial problems (n=3).

Procedure and Data Collection

The Thick Description, a qualitative empirical research method, was used in SELFIE to gain a deep understanding of the ICC programs from the different stakeholders' point of view [28,30]. The method undertaken included two different approaches: (1) Study of a variety of documents about each of the 17 ICC programs (ie, official and contractual documents, documents related to past evaluations, and factsheets from each ICC); and (2) interviews conducted with all relevant stakeholders (ie, program managers, initiators, representative of sponsor or payer organizations, health care professionals, informal caregivers, and patients or patient representatives). As described earlier, we concentrate on the two information technology-related dimensions in this paper.

Each partner-country interviewers underwent specific training on how to conduct and analyze the semistructured interviews to ensure uniform procedures. A total of 233 stakeholders were interviewed from November 2014 to September 2016 (see Multimedia Appendix 2 for more detailed information on stakeholder composition per country). The interviews were audio recorded and transcribed verbatim from the audio file by either the interviewer or an independent research transcriber. The resulting transcripts were analyzed by members of the local SELFIE teams using Mayring qualitative content analysis [31]. The quotations which were used in the thick descriptions were edited into readable forms and translated into English. The transcripts were not returned to participants for correction.

When writing this manuscript, we adhered to the COnsolidated criteria for REporting Qualitative research [32]. All information retrieved from the document analysis (including the stakeholders´ interviews) was processed according to the country-specific ethics statement listed under the subheading: *Ethics Approval and Consent to Participate*. The Thick Descriptions of the 17 ICC programs studied can be found on the SELFIE website [33].

Overarching Analysis

The first author did a thematic analysis [34] on the Thick Descriptions of the 2 components of the SELFIE conceptual framework referring to the enabling role of digital health tools (Technologies and medical products) and assessment strategies of the ICC programs (Information and research). He then discussed findings with all other coauthors. For each of these 2 components, this secondary analysis of the Thick Descriptions considered the 3 levels of the SELFIE conceptual framework: micro, meso, and macro. As detailed in Table 1, a 3-level grading system (+ to +++) was developed and used, under the criteria of the coauthors (JR and IC), to score maturity of the 17 ICC programs for each of the 2 components assessed in this research. Finally, the maturity of each of the 2 components of the SELFIE conceptual framework is summarized as the average of the maturity at micro, meso, and macro level.

Table 1. Summary of the maturity grading criteria at micro, meso, and macro level for technologies and medical products and for information and research.

Component	Grading
Technologies and medical products	
Digital health tools (micro)	
Electronic medical records—EMR	+
Personal health records at program level	++
Personal health records at regional level	+++
Organizational interoperability (meso)	
Health information exchange	+
Shared EMR[a]	++
Shared case management systems[b]	+++
Digital transformation policies (macro)	
Only addressing EMR	+
Several initiatives at program level	++
National and regional strategic plans	+++
Information and research	
Evaluation strategies (micro)	
Planned evaluation	+
Partial assessment	++
Full assessment with published papers	+++
Risk assessment (meso)	
Clinical knowledge or rule-based	+
Clinical predictive modelling tools	++
Multilevel predictive modelling tools[c]	+++
Research and innovation policies (macro)	
Incipient initiatives	+
Consolidated programs	++
Strong co-ordination with EU[d] programs	+++

[a] Shared electronic medical records among health care providers.

[b] Shared case management systems among health care providers to support integrated care pathways.

[c] Predictive modelling tools that combine information from various data sources, for example, clinical, population-based, biological, and patient-reported.

[d] EU: European Union.

Ethics Approval and Consent to Participate

Letters of Medical Ethics Approval of study protocols, questionnaires, and informed consent forms were sent and approved by the European Commission as a Deliverable of the SELFIE project.

Austria: Letter from Institute for Advanced Studies (IHS) declaring that ethical approval is not necessary for the evaluation of the two Austrian Integrated Care programs, October 3, 2017.

Croatia: Statement from the Agency for Quality and Accreditation in Health Care and Social Welfare declaring that the two evaluation studies are not within the scope of work of Croatian Central Ethics Committee, August 28, 2017, with reference to Official Gazette No 121/07 and No 25/15.

Germany, Gesundes Kinzigtal: Letter from the Ethical Committee, Technische Universität Berlin, declaring that the research is ethically acceptable (Ref: ST_02_20170620, August 15, 2017).

Germany, Casaplus: Letter from the Ethical Committee, Technische Universität Berlin, declaring that the research is ethically acceptable (Ref: ST_01_20170428, August 4, 2017).

Hungary, OnkoNetwork: The research plan has been authorized under approval No IG/03092–000/2016 by the Director General of Moritz Kaposi General Hospital, based on the positive opinion of the Institutional Research Committee and the responsible person of the Hospital for data protection. The Institutional Ethics Committee of the hospital double-checked the research and publication plan and confirmed that no ethical concerns were emerging related to this research and to the publication of findings (October 10, 2018, Ref: IKEB_IG_04125-000_2018).

Hungary, Palliative Care Consult Service: Letter from the Medical Research Council (Tudomanyos es Kutatasetikai Bizottsag, ETT TUKEB) declaring that the research is granted with Professional-Ethical Approval (Ref: 18632–4/2017/EKU, 24–4-2017).

The Netherlands, Proactive Primary Care Approach for Frail Elderly (U-PROFIT): Letter from the Medical Ethical Committee (MEC) Erasmus Medical Center Rotterdam declaring that the research is exempt from the Medical Research Involving Human Subjects Act (Dutch acronym: WMO; Ref: MEC-2017-402, July 25, 2017).

The Netherlands, Care Chain Frail Elderly (CCFE): Letter from the Medical Ethical Committee (MEC) Erasmus Medical Center Rotterdam declaring that the research is exempt from the Medical Research Involving Human Subjects Act (Dutch acronym: WMO; Ref: MEC-2014.558, December 18, 2014).

The Netherlands, Better Together in Amsterdam North (BSiN): Letter from the Medical Ethical Committee (MEC) of the Free University Medical Centre declaring that the research is exempt from the Medical Research Involving Human Subjects Act (Dutch acronym: WMO; Ref: MEC-2017-121, March 10, 2017).

Norway, Learning Network for Whole, Co-ordinated and Safe Pathways: Letter from the Regional Committees for Medical and Health Research Ethics-West (Komité for medisinsk og helsefaglig forskningsetikk -REK vest) declaring that the research is ethically approved (Ref: 2017/632/REK vest, March 28, 2017).

Norway, Medically Assisted Rehabilitation Bergen: Letter from The Regional Committees for Medical and Health Research Ethics-West (Komité formedisinsk og helsefaglig forskningsetikk -REK vest) declaring that the research is ethically approved (2017/944/REK vest, June 21, 2017).

Spain, Barcelona-Esquerra (AISBE): Letter from Clinic Research Ethical Committee (Comitè Ètic d'Investigació Clinica—CEIC) of the Clinic Hospital of Barcelona (Ref: CIF-G-08431173, Reg. HCB 2017/0451, June 14, 2017).

Spain, Badalona Serveis Assistencials (BSA): Letter from Clinic Research Ethical Committee (Comitè Ètic d'Investigació Clinica—CEIC) of the Clinic Hospital of Barcelona (Ref: CIF-G-08431173, Reg. HCB 2017/0453, June 14, 2017).

United Kingdom: Approval was granted by the National Health Service (NHS) Health Research Authority Research Ethics Committee (SELFIE REF: 16/WM/0295; CLASSIC REF: 14/NW/0206, June 23, 2016).

All participants provided written informed consent before participation.

Results

Overview

Multimedia Appendix 3 provides a high-level description of results for the 2 components of the SELFIE conceptual framework considered in this study. In this table, main features of each ICC program (third and fourth columns) are provided per type of ICC program (first column) and country (second column). Extended results for the 17 ICC programs are reported in the text below. As a summary of the results, Table 2 displays an average of the 3-level maturity grading criteria stated in Table 1 for each of the 17 ICC programs, according to the micro, meso, and macro levels of the SELFIE conceptual framework.

Technologies and Medical Products

The overarching analysis provided the following valuable insights on the implementation and exploitation of digital health tools to support ICC.

Digital Health Tools (Micro)

All 17 ICC programs have at least partial implementation of *EMR* and they are planning to enhance implementation of EMR in the future.

However, *specific personal health records* to enhance patient engagement are not considered in programs like Health Network Tennengau (HNT), Casaplus, OnkoNetwork (ON), Palliative Care Consulting Service (PCSS), Better together in Amsterdam North (BSiN), and Medically Assisted Rehabilitation (MAR). In such programs, digital information exchange between care provider and patient are either not considered or telephone is still the dominant tool for communications. Nevertheless, the use of personal health records has been key to support various telemonitoring services for patient self-management in programs like in GeroS (eKarton). Likewise, South Somerset Symphony Program (SSSP) and Salford Integrated Care Program (SICP), both from the United Kingdom, stress the role of digital health tools (ie, Patients Know Best) to support telemonitoring, albeit suffering from some implementation problems:

> *Tele-dermatology and we're piloting it [...] the GP will take a photograph and email it and get a decision, they're not doing suspected cancers obviously, but rashes. Yeah, we've done it* [IP11_1—SICP]

> *We've also got telehealth, that support. So we've got patients who are on telehealth in their homes, and each morning, the intensivists review the telehealth and see if there's any flags, like, if somebody is on [...] I'm trying to think. If somebody is on some sort of medication that they need to, you know, where fluid balance is an issue, if they've lost six pounds in weight that might flag some medication change. So they get them to weigh themselves, do their blood pressure, and so on. So, telehealth has been hugely supportive, actually, at keeping patients at home* [IP08_2—SSSP]

Table 2. Average maturity levels of the 17 Integrated Chronic Care programs.

Country	Program	Technologies and medical products	Information and research
Austria	HNT[a]	++	+
Austria	SMC[b]	+	+
Germany	Casaplus	++	++
Germany	GK[c]	++	++
Spain	AISBE[d]	+++	++
Spain	BSA[e]	++	++
Croatia	GeroS	++	+
Croatia	PCS[f]	+	+
Hungary	ON[g]	+	+
Hungary	PCCS[h]	+	+
The Netherlands	BSiN[i]	++	++
The Netherlands	CCFE[j]	++	++
The Netherlands	U-PROFIT[k]	++	++
Norway	LN[l]	+	+
Norway	MAR[m]	+	++
United Kingdom	SICP[n]	++	++
United Kingdom	SSSP[o]	++	+

[a] HNT: Health Network Tennengau.

[b] SMC: Sociomedical Centre Liebenau.

[c] GK: Gesundes Kinzigtal.

[d] AISBE: Area Integral de Salut Barcelona-Esquerra.

[e] BSA: Badalona Serveis Assistencials.

[f] PCS: Palliative Care System.

[g] ON: OnkoNetwork.

[h] PCCS: Palliative Care Consulting Service.

[i] BSiN: Better together in Amsterdam North.

[j] CCFE: Care Chain Frail Elderly.

[k] U-PROFIT: Proactive Primary Care Approach for Frail Elderly.

[l] LN: learning network.

[m] MAR: Medically Assisted Rehabilitation.

[n] SICP: Salford Integrated Care Program.

[o] SSSP: South Somerset Symphony Program.

It is of note that the availability of *personal health records at regional level*, which is the case with the programs Area Integral de Salut Barcelona-Esquerra (AISBE) and Badalona Serveis Assistencials (BSA) (La Meva Salut) [35], generates additional potential to foster collaborative work at micro level.

Organizational Interoperability (Meso)

Most of the programs use secure networks for *health information exchange* between hospitals and general practitioners, but with a broad spectrum of maturity. For example, the Casaplus program implemented a specific Web-based platform to support regular communication between case managers and nursing professionals only, but not primary care and the hospital. On the other hand, the health information exchange network used in HNT function only 1 way (Hospital to community). A potentially more mature example can be seen in SICP, which has implemented a single patient record accessible to the professionals of the case management multidisciplinary team and the emergency medicine professionals. Their ultimate goal is for the platform to be accessible by primary, secondary, and community care organizations in the Salford area. In a minority of the programs (eg, ON and PCSS), data transfer across various IT platforms of providers are manually performed by program administrators. All in all, most ICC programs indicate the determinant positive role of the existing regional digital health tools for health information exchange across health care tiers, which facilitates information sharing among heterogeneous providers, as seen in this example from the Proactive Primary Care Approach for Frail Elderly (U-PROFIT) program in the Netherlands:

> *We are working with a vulnerable population, frail in general, and it is important for them to avoid going from one place to another and visiting different service providers and collecting different forms [...] or duplicate papers because you have to present this paper here and this same paper over there [...] I think this is an important progress for the population* [IP04_1—U-PROFIT]

A step forward in terms of organizational interoperability, the computerization of health and social care records via a *shared EMR* among health care tiers, is at the heart of some ICC programs, such as the GeroS program. In line with organizational interoperability, the Care Chain Frail Elderly (CCFE) program focuses on structuring care and stimulating communication between all chain partners in primary care at various access levels with one another, thanks to an additional digital health tool (Care2U) that is used on top of the existing information systems to access the individual care plan and exchange information. However, although there has been much effort by the governments of these countries to have a shared EMR in place, this has not yet been fully successful, mostly due to data privacy issues. Last but not least, the AISBE program aims to consolidate a *shared case management system* [36] on top of the existing regional shared EMR, aiming to support the regional deployment of adaptive case management processes.

Digital Transformation Policies (Macro)

As all 17 ICC programs have at least partial implementation of EMR, all national and regional policies aim to expand the implementation of *EMR* in the future. However, in most program countries, the use and scope of digital health tools depends on *several initiatives at program level*, which serve as pilot sites for the nation and region wide rollout of digital health tools. This is the case, for example, in the Austrian programs (HNT and Sociomedical Centre Liebenau [SMC]), which are part of the *electronic health files*, the most comprehensive eHealth initiative in Austria. Our research has only been able to identify *national and regional strategic plans* for deployment of eHealth in Croatia, Spain, Hungary, the Netherlands, Norway, and the United Kingdom. The largest digital transformation policy being the Whole System Demonstrator pilots in the United Kingdom, which is a strategy proposed by the Department of Health in England to focus on health and social care for people with long-term needs, emphasizing the use of advanced assistive technologies including telehealth and telecare. It has demonstrated a slight reduction in mortality and emergency admission rates but was not demonstrated to be more cost effective than usual care [37,38].

Under the auspices of the Norwegian Directorate of Health, a *Care Journal* has been recently established for all citizens (voluntary); this is an electronic tool comprising selected and important health data that are accessible for the citizen and health personnel for the whole health care sector in Norway (including the 2 programs analyzed in this paper). Another example is the Catalan Health Plan [39], which prioritizes the improvement and transformation of the health system and health care organization through the intensive introduction of emerging digital health technologies.

Information and Research

As summarized in the Multimedia Appendix 3, the overarching analysis provided the following valuable insights on the assessment strategies of the 17 technology-enabled ICC programs.

Evaluation Strategies (Micro)

This research has shown that in some ICC programs, no comprehensive evaluation has been carried out so far (SMC, Palliative Care System, ON, and Learning network), but is *planned* to be performed.

However, most ICC programs have been subject to *partial monitoring and/or preliminary evaluation* (ie, HNT, GeroS, Palliative Care Consulting Service, CCFE, BSiN, MAR, SICP, and SSSP), involving mainly descriptive data analysis over well-defined outcome measures of interest or key performance indicators. Specifically, the SSSP program includes:

> *Number of bed days, average length of stay, 30 day readmission, avoidable emergency admissions, precautionary emergency admissions, patients admitted multiple times, excess bed days, avoidable A&E attendances, confidence to my own health, received enough support to help self-managed long-term conditions, have a written care plan, care plan regularly reviewed, patient access to GP and nurse, online services, GP referrals, mental wellbeing, the Warwick-Edinburgh Mental Wellbeing scale, patient activation measure [PAM], patient satisfaction experience, and number of contacts made.* [IP03_2—SSSP]

Only some programs report *full scientific assessment* (ie, Casaplus [40], Gesundes Kinzigtal [41,42], U-PROFIT [43,44], and AISBE [45-47]). These programs have been evaluated using randomized controlled trials as well as pre-post evaluation with propensity score matching methods, following the Triple Aim outcomes [48,49]:

> *...was the number of hospital admissions reduced? How did they experience the effects of care (the insured person, the environment, the relatives)? Were the per capita costs reduced?* [IP04_1—Casaplus]

Risk Assessment Strategies (Meso)

Patient management purely based on *clinical criteria* (professional training, knowledge, instinct, and experience) *or combined with rules-based clinical management* [50] (thresholds for certain parameters defining pre-established decision criteria) constitutes current health professional practice in most ICC programs.

In contrast, the regular use of *subject-specific predictive modelling* tools for clinical decision support (predictive modelling establishing relationships between sets of variables and outcomes generated using statistical or machine learning tools) is still in its infancy, despite the fact that it seems a natural step toward customization of care to patient's needs. Clinical predictive modelling tools are only reported to be used in some ICC programs (ie, SICP, SSSP, U-PROFIT, Gesundes Kinzigtal, and Casaplus) *for individual risk assessment* (which can be considered within the micro level). Within the U-PROFIT program, available data in the general practitioner EMR system are used by the U-PRIM software to screen frail patients of 60 years and older in every participating practice [51]. SICP and SSSP programs in the United Kingdom use a well-known patient-level risk predictive tool, PARS [52] and the Combined Predictive Model [53], to identify those patients that require the most care and support and to assess the risk of patients having unplanned hospital admissions within a 12-month period. It is used to some extent, but more trust is placed in clinical judgement in many cases:

> *...so we'd looked at some of the higher risk patients that were identified by the Combined Predictive Model and PARS (Patient at Risk Score) exactly the same, because I've done that before for the unscheduled care, and we looked at that; and what you find is the high risk people that are identified by this risk stratification models that are promoted nationally, is that the only data that's easy to count is the hospital data, is the Hospital Episode Statistics of your hospital episodic statistics and stuff.* [IP02_1—SICP]

Similarly, Casaplus uses a clinical predictive modelling tool to identify patients in high risk for hospital admissions within the next 12 months.

The use of clinical predictive modelling tools for *population-based risk assessment* is only reported in BSA and AISBE (Catalonia, ES). Since 2011, the Catalan Health Surveillance system collects detailed information on health care usage for the entire population of Catalonia [54], the region in which AISBE and BSA operate. It includes information on hospitalization, primary care visits, emergency department visits, skilled nursing facilities, palliative care and the mental health services, information on pharmacy prescription and expenditure, and a registry on the billing record also encompassing outpatient visits to specialists, home hospitalization, medical transportation (urgent and nonurgent), ambulatory rehabilitation, respiratory therapies, and dialysis. This information is used for provider payment purposes. Also, external audits are performed periodically to ensure the quality and reliability of the data. The Catalan Health Surveillance System is used to update, on a 6-month basis, the regional population-based health risk assessment tool, (the Adjusted Morbidity Groups) that generates the health risk strata pyramid of the general population of Catalonia [13,14].

Furthermore, AISBE is adopting a holistic approach that fosters inclusion of covariates from multilevel data sources, namely *Multilevel Predictive Modelling*: (1) clinical, (2) informal care; (3) biological research; and (4) outcomes from population-health risk predictive modelling (eg, the Adjusted Morbidity Groups), resulting in enhanced patient-based stratification and optimization of service selection. This approach aims to pave the way toward personalized medicine, provided that access to the multilevel data sources is granted. However, most legal frameworks on data privacy of the 17 ICC programs depend on the ongoing implementation of the European Union Data Protection Directive 95/46/EC to make the concept of multilevel predictive modelling operational.

Research and Innovation Policies (Macro)

A majority of the 17 ICC programs are part of *incipient research and innovation initiatives* constantly being implemented in practice, both bottom-up and top-down, using several, sometimes consecutive, project-budgets but without sustainable structural funding.

However, Croatian, Dutch, German, Norway, and Spanish programs are aligned with *consolidated research and innovation programs* at state and regional level. For example, the Gesundes Kinzigtal program in Germany has been extensively evaluated in terms of prevalence of multimorbidity, polypharmacy, proportion of generic drugs, prevalence of problematic drug prescriptions, prevalence of fractures among patients diagnosed with osteoporosis, quality of services, and overall health care costs [42]. Another example is the Research Council of Norway, commissioned by the Ministry of Health and Care Services to carry out a research-based evaluation of the Co-ordination Reform. The Research Council has conducted a research program tailored at integrated care from 2012 to 2015.

Furthermore, the *strong co-ordination* of most programs from Norway, the Netherlands, Hungary, Spain, Croatia, and Germany *with different European research and innovation initiatives* under the umbrella of H2020 [55], EIP-AHA [56], EIT Health [57], and/or RIS3 [58], as well as other specific research and innovation actions, should contribute to cross-fertilization among health care, research, and innovation.

Discussion

Principal Findings

The overarching analysis allowed us to assess the use of digital health tools to support the care process in the 17 ICC programs on the 2 specific aspects analyzed: *Technologies & Medical* products and *Information & Research*. As most of the ICC programs are pilot experiences in terms of nation and region wide rollout of digital health tools, this analysis was useful to learn from them regarding requirements for a successful large-scale implementation elsewhere.

Acknowledging that the 17 ICC programs are highly heterogeneous regarding the use and impact of digital health tools, the main findings are summarized below.

Electronic Medical Records

Each program studied shows at least partial implementation of EMR, and all of them have plans in place for a future mature implementation of EMR.

Personal Health Records

The use of personal health records to support telemonitoring services for patient self-management is not in place in most ICC programs, for which telephone is still the dominant means of communication.

Health Information Exchange Platforms

Most programs reported on the potential of secure health information exchange across providers to facilitate organizational interoperability for deployment of ICC by facilitating information sharing among heterogeneous providers and avoid generating additional burden of double-registration to health professionals. However, the maturity of implementation is currently rather poor. Moreover, the need for technological tools, on top of health information exchange platforms, supporting collaborative work across health care tiers to foster implementation of shared case management [36,59] was stressed by programs like AISBE.

Digital Transformation Policies

The overarching analysis highlighted the lack of well-defined macro-level policies, with effective operational implementation plans, in the health care systems in which most ICC programs operate. Often, the use and scope of digital health tools depends on local or regional initiatives of individual providers involved in the ICC programs.

Health Data Analytics and Evaluation Strategies

Most programs systematically collect well-defined outcome measures to feed program-specific evaluation strategies, ranging from descriptive data analysis, comparison of trial and control groups, as well as pre- and postmeasuring. Still, some ICC programs recognize barriers for assessment such as a lack of financing, poor research capacity, concerns on data security, and misuse of data. This study clearly shows the need for formulation of structured and comprehensive evaluation and monitoring strategies for ICC including formulation of key performance indicators extensively shared across countries. Moreover, the Quadruple Aim approach [60,61] (ie, the Triple Aim approach plus the health care professionals experience) should serve to standardize the evaluation across European Union sites.

Health Risk Assessment

Just a few of the ICC programs report on the use of clinical predictive modelling tools, and even less ICC programs claim the use of population-based health-risk prediction tools.

Research and Innovation Policies

The majority of the 17 ICC programs (either bottom-up or top-down initiatives) are often based on project-specific budgets without well-defined, operational policies and, consequently, without sustainable structural funding. Implementing the above technological innovations frequently requires hardware and software upgrades. The costs of initial rollout and training of staff also need to be considered and weighed against the likely benefits.

We acknowledge some limitation in our study such as the inherent limitations of the methodological approach adopted. Also, as the study conclusions relate only to programs based in Europe, worldwide representativeness of the study results cannot be assumed.

Comparison With Previous Work and Future Directions

A recent report by the European Union on Integrated Care maturity [62], including the evaluation of Information and eHealth tools, concluded that the level of maturity in Germany, Denmark, Belgium, Italy, Spain, Greece, Sweden, and Iceland scored higher in comparison to their peers in Estonia, the Netherlands, Poland, and Bulgaria. The most mature countries are Denmark, followed by Spain, Germany, and Iceland at the same level. Finally, Sweden, Belgium, Italy, and Greece were in intermediate level. This is line with our findings, except for the 3 case studies in the Netherlands, which in our study were more mature than Dates et al [62] suggests. It is of note that differences between the study by Dates et al [62] and this study might be explained by using different tools as well as the selection of 3 promising cases in the Netherlands which actually use eHealth tools. The European Union on Integrated Care maturity report applies an interactive tool developed by SIROCCO (Scaling Integrated Care in Context) to assess the maturity of ICC programs on a global level using different aspects (one of them being eHealth); whereas in our study, we explicitly focus on different aspects of digital health tools and assessment strategies solely. Also, the previous report is based on a review of the literature on integrated care policies and strategies, whereas our study adds the point of view from the different stakeholders directly involved in generation and implementation of these policies and strategies.

Integrated care programs for chronic patients involve complex interventions for heterogeneous populations; therefore, proper articulation of digital health tools and the different components of the evaluation process are still unmet needs that markedly hinder comparability and scale-up. The overarching analysis of the 17 ICC programs conducted in this study allowed us to identify the following potential areas for future developments:

Refinement of assessment methodologies of large-scale deployment and adoption of ICC programs, likely based on implementation research approaches [63-65], are needed. We understand that assessment should adopt the classical three-dimensional approach including outcomes, processes, and structures [66]. Moreover, usual health outcome variables (ie, mortality, hospital readmissions, etc.) should be ideally expanded [67] considering the Quadruple Aim approach [60,61]. The approach requires the collection of patient-reported outcomes and experience data (PROMS and PREMS) on a regular basis.

The concept of adaptive case management explored in AISBE [36,68] should be made operational. Conventional health information systems rely on the management of clinical episodes with a disease-oriented approach and only very rarely incorporate the required process logics to support continuity of care with a patient-centered approach.

Dynamic health-risk assessment taking into consideration both service commissioning (*population-based health-risk predictive modelling*) and subject-specific service selection involving optimal patient allocation in the health system (*individual health-risk predictive modelling supporting decision support*) should be addressed to improve outcomes [69-71]. Ultimately, the application of holistic strategies for subject-specific risk prediction and stratification that incorporates multilevel determinants of health (eg, socioeconomic, lifestyle, behavioral, clinical, physiological, cellular, and *omics* information) emerges as a high priority goal to properly pave the way toward personalized medicine for complex chronic patients [72]. Enhanced clinical predictive modelling, personalized diagnostic and treatment tools can contribute to the acceleration of transfer of scientific evidence to practice.

Development of pragmatic trials that incorporate real-life evidence from multilevel determinants of health may require implementation strategies, ideally using cloud computing environments, tackling privacy and regulatory constraints [23,72]. Currently, the articulation of the main technical building blocks, that is, multilevel biomedical data integration, tools for clinical predictive modelling in the cloud and High-Performance Computing, as one integrated system is yet a largely unmet potential.

Conclusions

This overarching analysis informs the current implementation status of digital health tools for management of multimorbidity in the 17 promising ICC programs selected in SELFIE. Notwithstanding the heterogeneity of the results, most studied programs are progressively evolving their supporting digital health tools from pilot prototypes to full scale-up at regional and national level. However, the majority of programs have not yet undergone full evaluation and assessment strategies. Future directions which can enable of digital transformation are based on innovation at micro and meso level with full support from the macro level. Some strategic areas that can help toward this end are the following: (1) implementation of research strategies; (2) explore an adaptive case management approach; (3) further developments of health risk assessment; and (4) holistic implementation strategies using future, regulatory compliant, cloud computing environments.

Acknowledgments

The authors gratefully acknowledge the contribution of all other members of the SELFIE consortium and the linked third parties. Membership of the SELFIE consortium can be found on www.selfie2020.eu. The SELFIE project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement #634288. The content of this paper reflects only the SELFIE group’s views and the European Commission is not liable for any use that may be made of the information contained herein.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Essential and additional criteria for preliminary selection of Sustainable intEgrated care modeLs for multimorbidity: delivery, Financing, and performancE (SELFIE) programs.
[PDF File (Adobe PDF File), 223 KB - jmir_v21i8e14956_app1.pdf]

Multimedia Appendix 2
Stakeholders interviews per country.
[PDF File (Adobe PDF File), 241 KB - jmir_v21i8e14956_app2.pdf]

Multimedia Appendix 3
Summary of the overarching analysis of the 17 selected integrated chronic care programs.
[PDF File (Adobe PDF File), 184 KB - jmir_v21i8e14956_app3.pdf]

References

1. Barbabella F, Melchiorre MG, Quattrini S, Papa R, Lamura G. In: Richardson E, van Ginneken E, editors. How Can eHealth Improve Care for People With Multimorbidity in Europe?. Copenhagen, Denmark: European Observatory on Health Systems and Policies; 2017.
2. European Commission. 2012. eHealth Action Plan 2012-2020: Innovative Healthcare for the 21st century URL: https://ec.europa.eu/digital-single-market/en/news/ehealth-action-plan-2012-2020-innovative-healthcare-21st-century [accessed 2018-03-01]
3. World Health Organization. 2019. eHealth Unit URL: https://www.who.int/ehealth/en/ [accessed 2019-02-04]
4. Steinhubl SR, Topol EJ. Digital medicine, on its way to being just plain medicine. NPJ Digit Med 2018;1:20175 [FREE Full text] [doi: 10.1038/s41746-017-0005-1] [Medline: 31304349]
5. Lupiáñez-Villanueva F, Anastasiadou D, Codagnone C, Nuño-Solinís R, Garcia-Zapirain SM. Electronic health use in the European union and the effect of multimorbidity: cross-sectional survey. J Med Internet Res 2018 May 3;20(5):e165 [FREE Full text] [doi: 10.2196/jmir.7299] [Medline: 29724702]

6. Melchiorre MG, Papa R, Rijken M, van Ginneken E, Hujala A, Barbabella F. eHealth in integrated care programs for people with multimorbidity in Europe: insights from the ICARE4EU project. Health Policy 2018 Jan;122(1):53-63 [FREE Full text] [doi: 10.1016/j.healthpol.2017.08.006] [Medline: 28899575]
7. Williams PA, Lovelock B, Cabarrus T, Harvey M. Improving digital hospital transformation: development of an outcomes-based infrastructure maturity assessment framework. JMIR Med Inform 2019 Jan 11;7(1):e12465 [FREE Full text] [doi: 10.2196/12465] [Medline: 30632973]
8. Campanella P, Lovato E, Marone C, Fallacara L, Mancuso A, Ricciardi W, et al. The impact of electronic health records on healthcare quality: a systematic review and meta-analysis. Eur J Public Health 2016 Feb;26(1):60-64. [doi: 10.1093/eurpub/ckv122] [Medline: 26136462]
9. Meskó B, Drobni Z, Bényei E, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth 2017;3:38 [FREE Full text] [doi: 10.21037/mhealth.2017.08.07] [Medline: 29184890]
10. European Commission. 2014. Green Paper on Mobile Health ('mHealth') URL: https://ec.europa.eu/digital-agenda/en/news/green-paper-mobile-health-mhealth [accessed 2014-06-22] [WebCite Cache ID 6QWamz0Y4]
11. Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. Eur Heart J 2016 May 7;37(18):1428-1438 [FREE Full text] [doi: 10.1093/eurheartj/ehv770] [Medline: 26873093]
12. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
13. Monterde D, Vela E, Clèries M, Grupo Colaborativo GMA. [Adjusted morbidity groups: a new multiple morbidity measurement of use in primary care]. Aten Primaria 2016 Dec;48(10):674-682 [FREE Full text] [doi: 10.1016/j.aprim.2016.06.003] [Medline: 27495004]
14. Dueñas-Espín I, Vela E, Pauws S, Bescos C, Cano I, Cleries M, et al. Proposals for enhanced health risk assessment and stratification in an integrated care scenario. BMJ Open 2016 Apr 15;6(4):e010301 [FREE Full text] [doi: 10.1136/bmjopen-2015-010301] [Medline: 27084274]
15. Merck SF. Chronic disease and mobile technology: an innovative tool for clinicians. Nurs Forum 2017 Oct;52(4):298-305. [doi: 10.1111/nuf.12202] [Medline: 28407257]
16. Cho JH, Lee HC, Lim DJ, Kwon HS, Yoon KH. Mobile communication using a mobile phone with a glucometer for glucose control in type 2 patients with diabetes: as effective as an internet-based glucose monitoring system. J Telemed Telecare 2009;15(2):77-82. [doi: 10.1258/jtt.2008.080412] [Medline: 19246607]
17. Logan AG, McIsaac WJ, Tisler A, Irvine MJ, Saunders A, Dunai A, et al. Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. Am J Hypertens 2007 Sep;20(9):942-948. [doi: 10.1016/j.amjhyper.2007.03.020] [Medline: 17765133]
18. 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]
19. Kim YJ, Rhee SY, Byun JK, Park SY, Hong SM, Chin SO, et al. A smartphone application significantly improved diabetes self-care activities with high user satisfaction. Diabetes Metab J 2015 Jun;39(3):207-217 [FREE Full text] [doi: 10.4093/dmj.2015.39.3.207] [Medline: 26124991]
20. Zulman DM, Jenchura EC, Cohen DM, Lewis ET, Houston TK, Asch SM. How can ehealth technology address challenges related to multimorbidity? Perspectives from patients with multiple chronic conditions. J Gen Intern Med 2015 Aug;30(8):1063-1070 [FREE Full text] [doi: 10.1007/s11606-015-3222-9] [Medline: 25691239]
21. National Institute for Health and Care Excellence. 2016. Multimorbidity: Clinical Assessment and Management URL: https://www.nice.org.uk/guidance/ng56/resources/multimorbidity-clinical-assessment-and-management-pdf-1837516654789 [accessed 2018-01-29]
22. Vayena E, Haeusermann T, Adjekum A, Blasimme A. Digital health: meeting the ethical and policy challenges. Swiss Med Wkly 2018;148:w14571 [FREE Full text] [doi: 10.4414/smw.2018.14571] [Medline: 29376547]
23. Renner AT, Bobek J, Habl C, Anja J. European Commission. 2016. Study on Big Data in Public Health, Telemedicine and Healthcare URL: https://ec.europa.eu/digital-single-market/en/news/study-big-data-public-health-telemedicine-and-healthcare [accessed 2019-08-12]
24. Adjekum A, Blasimme A, Vayena E. Elements of trust in digital health systems: scoping review. J Med Internet Res 2018 Dec 13;20(12):e11254 [FREE Full text] [doi: 10.2196/11254] [Medline: 30545807]
25. Schiza EC, Kyprianou TC, Petkov N, Schizas CN. Proposal for an ehealth based ecosystem serving national healthcare. IEEE J Biomed Health Inform 2019 May;23(3):1346-1357. [doi: 10.1109/JBHI.2018.2834230] [Medline: 29993757]
26. SELFIE: Sustainable Integrated Care Models for Multi-Morbidity: Delivery, Financing and Performance. URL: http://www.selfie2020.eu/ [accessed 2018-04-02]
27. Leijten LF, Struckmann V, van Ginneken E, Czypionka T, Kraus M, Reiss M, SELFIE Consortium. The SELFIE framework for integrated care for multi-morbidity: development and description. Health Policy 2018 Jan;122(1):12-22 [FREE Full text] [doi: 10.1016/j.healthpol.2017.06.002] [Medline: 28668222]
28. Selfie 2020. 2018. Work Package 2: Executive summary URL: https://www.selfie2020.eu/wp-content/uploads/2016/12/SELFIE_WP2_Executive-summary.pdf [accessed 2017-12-03]

29. Struckmann V, Leijten FR, van Ginneken E, Kraus M, Reiss M, Spranger A, SELFIE Consortium. Relevant models and elements of integrated care for multi-morbidity: results of a scoping review. Health Policy 2018 Jan;122(1):23-35 [FREE Full text] [doi: 10.1016/j.healthpol.2017.08.008] [Medline: 29031933]
30. Pitter JG, Csanádi M, Szigeti A, Lukács G, Kovács A, Moizs M, et al. Planning, implementation and operation of a personalized patient management system for subjects with first suspect of cancer (OnkoNetwork): system description based on a qualitative study. BMC Health Serv Res 2019 Feb 21;19(1):131 [FREE Full text] [doi: 10.1186/s12913-019-3957-9] [Medline: 30791917]
31. Mayring P. Qualitative Inhaltsanalyse: Grundlagen und Techniken. Twelfth Edition. Weinheim, Germany: Beltz; 2015.
32. 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]
33. Selfie 2020. 2018. SELFIE Case Studies URL: https://www.selfie2020.eu/publications/ [accessed 2018-03-01]
34. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, CA: Sage Publications; 1998.
35. Modol JR. Navigating towards self-care: the Catalan public patient portal. In: Aanestad M, Grisot M, Hanseth O, Vassilakopoulou P, editors. Information Infrastructures within European Health Care. Cham, Germany: Springer; 2017.
36. Cano I, Alonso A, Hernandez C, Burgos F, Barberan-Garcia A, Roldan J, et al. An adaptive case management system to support integrated care services: lessons learned from the NEXES project. J Biomed Inform 2015 Jun;55:11-22 [FREE Full text] [doi: 10.1016/j.jbi.2015.02.011] [Medline: 25796455]
37. Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Hirani S, Whole System Demonstrator Evaluation Team. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. Br Med J 2012 Jun 21;344:e3874 [FREE Full text] [doi: 10.1136/bmj.e3874] [Medline: 22723612]
38. Henderson C, Knapp M, Fernández JL, Beecham J, Hirani SP, Cartwright M, Whole System Demonstrator Evaluation Team. Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. Br Med J 2013 Mar 20;346:f1035 [FREE Full text] [doi: 10.1136/bmj.f1035] [Medline: 23520339]
39. Inici. Departament de Salut. 2016. Pla de Salut de Catalunya 2016-2020 URL: http://salutweb.gencat.cat/web/.content/_departament/pla-de-salut/Pla-de-salut-2016-2020/documents/Pla_salut_Catalunya_2016_2020.pdf [accessed 2017-11-03]
40. DocPlayer. 2008. MedicalContact AG. 365 Tage Casaplus. Ein Präventionsprogramm feiert Geburtstag URL: https://docplayer.org/37867445-Medicalcontact-ag-365-tage-casaplus-ein-praeventionsprogramm-feiert-geburtstag.html [accessed 2018-11-04]
41. Schulte T, Pimperl A. Pursuing the triple aim: a three-dimensional outcome evaluation of the integrated care project 'gesundes kinzigtal'? Int J Integr Care 2013 Oct 23;13(5):-. [doi: 10.5334/ijic.1217]
42. Hildebrandt H, Hermann C, Knittel R, Richter-Reichhelm M, Siegel A, Witzenrath W. Gesundes kinzigtal integrated care: improving population health by a shared health gain approach and a shared savings contract. Int J Integr Care 2010;10(2):e046. [doi: 10.5334/ijic.539] [Medline: 20689772]
43. Bleijenberg N, Drubbel I, Schuurmans MJ, Dam HT, Zuithoff NP, Numans ME, et al. Effectiveness of a proactive primary care program on preserving daily functioning of older people: a cluster randomized controlled trial. J Am Geriatr Soc 2016 Sep;64(9):1779-1788. [doi: 10.1111/jgs.14325] [Medline: 27459236]
44. Bleijenberg N, Drubbel I, Neslo RE, Schuurmans MJ, ten Dam VH, Numans ME, et al. Cost-effectiveness of a proactive primary care program for frail older people: a cluster-randomized controlled trial. J Am Med Dir Assoc 2017 Dec 1;18(12):1029-36.e3 [FREE Full text] [doi: 10.1016/j.jamda.2017.06.023] [Medline: 28801235]
45. Hernández C, Aibar J, Seijas N, Puig I, Alonso A, Garcia-Aymerich J, et al. Implementation of home hospitalization and early discharge as an integrated care service: a ten years pragmatic assessment. Int J Integr Care 2018 May 16;18(2):12 [FREE Full text] [doi: 10.5334/ijic.3431] [Medline: 30127696]
46. Hernández C, Alonso A, Garcia-Aymerich J, Grimsmo A, Vontetsianos T, Cuyàs FG, et al. Integrated care services: lessons learned from the deployment of the NEXES project. Int J Integr Care 2015;15:e006 [FREE Full text] [doi: 10.5334/ijic.2018] [Medline: 26034465]
47. Barberan-Garcia A, Vogiatzis I, Solberg HS, Vilaró J, Rodríguez DA, Garåsen HM, NEXES Consortium. Effects and barriers to deployment of telehealth wellness programs for chronic patients across 3 European countries. Respir Med 2014 Apr;108(4):628-637 [FREE Full text] [doi: 10.1016/j.rmed.2013.12.006] [Medline: 24451438]
48. South Carolina Hospital Association. 2012. A Guide to Measuring the Triple Aim: Population Health, Experience of Care, and Per Capita Cost URL: https://www.scha.org/files/ihiguidetomeasuringtripleaimwhitepaper2012.pdf [accessed 2017-11-05]
49. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff (Millwood) 2008;27(3):759-769. [doi: 10.1377/hlthaff.27.3.759] [Medline: 18474969]
50. Margolis CZ. Uses of clinical algorithms. J Am Med Assoc 1983 Feb 4;249(5):627-632. [doi: 10.1001/jama.1983.03330290049028] [Medline: 6336813]
51. Drubbel I. Frailty Screening in Older Patients in Primary Care Using Routine Care Data. The Netherlands: Utrecht University; 2013.

52. Billings J, Dixon J, Mijanovich T, Wennberg D. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. Br Med J 2006 Aug 12;333(7563):327 [FREE Full text] [doi: 10.1136/bmj.38870.657917.AE] [Medline: 16815882]
53. Wennberg D, Siegel M, Darin B, Billings J, Dixon J, Curry N. The King's Fund. 2006. Combined Predictive Model: Final Report URL: https://www.kingsfund.org.uk/sites/default/files/Combined_Predictive_Model_Final_Report.pdf [accessed 2018-12-04]
54. Farré N, Vela E, Clèries M, Bustins M, Cainzos-Achirica M, Enjuanes C, et al. Medical resource use and expenditure in patients with chronic heart failure: a population-based analysis of 88 195 patients. Eur J Heart Fail 2016 Sep;18(9):1132-1140 [FREE Full text] [doi: 10.1002/ejhf.549] [Medline: 27108481]
55. European Commission. Horizon 2020 URL: https://ec.europa.eu/programmes/horizon2020/en [accessed 2017-06-01]
56. European Commission. Innovation Union: Aims of the Innovation Union, State of Progress and Related Policy URL: http://ec.europa.eu/research/innovation-union/index_en.cfm?section=active-healthy-ageing [accessed 2017-06-19] [WebCite Cache ID 6QRp7pXLY]
57. EIT Health. URL: https://www.eithealth.eu/ [accessed 2018-01-06]
58. Smart Specialisation Platform - Europa EU. RIS3 Guide URL: http://s3platform.jrc.ec.europa.eu/s3-guide [accessed 2018-05-17]
59. Herrmann C, Kurz M. Adaptive Case Management: Supporting Knowledge Intensive Processes with IT Systems. In: Proceedings of the International Conference on Subject-Oriented Business Process Management. 2011 Presented at: S-BPM ONE'11; September 29-30, 2011; Ingolstadt, Germany p. 80-97. [doi: 10.1007/978-3-642-23471-2_6]
60. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]
61. West CP. Physician well-being: expanding the triple aim. J Gen Intern Med 2016 May;31(5):458-459 [FREE Full text] [doi: 10.1007/s11606-016-3641-2] [Medline: 26921157]
62. Dates M, Lennox-Chhugani N, Sant H, Pereira A, Tedeschi M. European Commission. 2018. Health System Performance Assessment – Integrated Care Assessment (20157303 HSPA) URL: https://ec.europa.eu/health/sites/health/files/systems_performance_assessment/docs/2018_integratedcareassessment_en.pdf [accessed 2018-09-14]
63. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Implementation research: what it is and how to do it. Br Med J 2013 Nov 20;347:f6753. [doi: 10.1136/bmj.f6753] [Medline: 24259324]
64. Pinnock H, Barwick M, Carpenter CR, Eldridge S, Grandes G, Griffiths CJ, StaRI Group. Standards for Reporting Implementation Studies (StaRI): explanation and elaboration document. BMJ Open 2017 Apr 3;7(4):e013318 [FREE Full text] [doi: 10.1136/bmjopen-2016-013318] [Medline: 28373250]
65. Peters DH, Tran NT, Adam T. Implementation Research in Health: A Practical Guide. Geneva, Switzerland: World Health Organization; 2013.
66. Donabedian A. The quality of care. How can it be assessed? J Am Med Assoc 1988;260(12):1743-1748. [doi: 10.1001/jama.260.12.1743] [Medline: 3045356]
67. Sikka R, Morath JM, Leape L. The Quadruple Aim: care, health, cost and meaning in work. BMJ Qual Saf 2015 Oct;24(10):608-610. [doi: 10.1136/bmjqs-2015-004160] [Medline: 26038586]
68. Hauder M, Pigat S, Matthes F. Research Challenges in Adaptive Case Management: A Literature Review. In: Proceedings of the 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations. 2014 Presented at: EDOCW'14; September 1-2, 2014; Ulm, Germany. [doi: 10.1109/EDOCW.2014.24]
69. Hu Z, Hao S, Jin B, Shin AY, Zhu C, Huang M, et al. Online prediction of health care utilization in the next six months based on electronic health record information: a cohort and validation study. J Med Internet Res 2015 Sep 22;17(9):e219 [FREE Full text] [doi: 10.2196/jmir.4976] [Medline: 26395541]
70. Cho I, Boo E, Chung E, Bates DW, Dykes P. Novel approach to inpatient fall risk prediction and its cross-site validation using time-variant data. J Med Internet Res 2019 Feb 19;21(2):e11505 [FREE Full text] [doi: 10.2196/11505] [Medline: 30777849]
71. Hao S, Jin B, Shin AY, Zhao Y, Zhu C, Li Z, et al. Risk prediction of emergency department revisit 30 days post discharge: a prospective study. PLoS One 2014;9(11):e112944 [FREE Full text] [doi: 10.1371/journal.pone.0112944] [Medline: 25393305]
72. Cano I, Tenyi A, Vela E, Miralles F, Roca J. Perspectives on big data applications of health information. Curr Opin Syst Biol 2017 Jun;3:36-42. [doi: 10.1016/j.coisb.2017.04.012]

Abbreviations

AISBE: Area Integral de Salut Barcelona-Esquerra
BSA: Badalona Serveis Assistencials
BSiN: Better together in Amsterdam North
CCFE: Care Chain Frail Elderly
EMR: Electronic Medical Record

HNT: Health Network Tennengau (Gesundheitsnetzwerk Tennengau)
ICC: Integrated Chronic Care
MAR: Medically Assisted Rehabilitation
ON: OnkoNetwork
SELFIE: Sustainable intEgrated care modeLs for multimorbidity: delivery, Financing, and performancE
SICP: Salford Integrated Care Program
SMC: Sociomedical Centre Liebenau (Sozialmedizinisches Zentrum Liebenau)
SSSP: South Somerset Symphony Program
U-PROFIT: Proactive Primary Care Approach for Frail Elderly

Edited by G Eysenbach; submitted 06.06.19; peer-reviewed by KC Wong, I Mircheva, K Goniewicz; comments to author 21.07.19; revised version received 30.07.19; accepted 30.07.19; published 26.09.19.

Please cite as:
Baltaxe E, Czypionka T, Kraus M, Reiss M, Askildsen JE, Grenkovic R, Lindén TS, Pitter JG, Rutten-van Molken M, Solans O, Stokes J, Struckmann V, Roca J, Cano I
Digital Health Transformation of Integrated Care in Europe: Overarching Analysis of 17 Integrated Care Programs
J Med Internet Res 2019;21(9):e14956
URL: https://www.jmir.org/2019/9/e14956
doi:10.2196/14956
PMID:31573914

Original Paper

The Use of and Experiences With Telelactation Among Rural Breastfeeding Mothers: Secondary Analysis of a Randomized Controlled Trial

Kandice Kapinos[1], PhD; Virginia Kotzias[1], MPP; Debra Bogen[2], MD; Kristin Ray[2], MD; Jill Demirci[3], PhD, RN, IBCLC; Mary Ann Rigas[4], MD; Lori Uscher-Pines[1], PhD, MSc

[1]RAND Corporation, Arlington, VA, United States
[2]University of Pittsburgh School of Medicine, Pittsburgh, PA, United States
[3]University of Pittsburgh School of Nursing, Pittsburgh, PA, United States
[4]UPMC Cole Hospital, Coudersport, PA, United States

Corresponding Author:
Kandice Kapinos, PhD
RAND Corporation
1200 South Hayes Street
Arlington, VA, 22202
United States
Phone: 1 703 413 1100 ext 5361
Email: kkapinos@rand.org

Abstract

Background: Telelactation services connect breastfeeding mothers to remotely located lactation consultants through audio-visual technology and can increase access to professional breastfeeding support in rural areas.

Objective: The objective of this study was to identify maternal characteristics associated with the demand for and use of telelactation and to describe visit characteristics.

Methods: We conducted a descriptive study within the context of a randomized controlled trial. Participant survey data and vendor electronic medical record data were used to assess video call characteristics like timing, duration, topics discussed, and participant satisfaction. Recruitment occurred from 2016-2018 at a rural critical access hospital in Pennsylvania. The 102 women enrolled in the study were given access to unlimited, on-demand video calls with lactation consultants through a mobile phone app and were tracked for 12 weeks following their postpartum hospitalization.

Results: A total of 94 participants out of 102 recruits (92%) participated in the final, 12-week survey assessment were included in the analysis. Of those, 47 (50%) participants reported participating in one or more video calls, and 31 (33%) completed one or more calls that included a substantive discussion of a breastfeeding challenge. Participants who used telelactation (21/31, 68%; P=.02) were more likely to be working at 12 weeks postpartum compared to others (26/63, 41%), were less likely (12/31, 39%; P=.02) to have prior breastfeeding experience on average compared to nonusers (41/63, 65%), and were less likely to have breastfed exclusively (16/31, 52%; P<.001) prior to hospital discharge compared to mothers who didn’t use telelactation services (51/63, 81%). Most video calls (58/83, 70%) occurred during the infant’s first month of life and 41% (34/83) occurred outside of business hours. The most common challenges discussed included: breast pain, soreness, and infection (25/83, 30%), use of nipple shields (21/83, 25%), and latch or positioning (17/83, 24%). Most telelactation users (43/47, 91%) expressed satisfaction with the help received.

Conclusions: Telelactation is an innovation in the delivery of professional breastfeeding support. This research documents both demand for and positive experiences with telelactation in an underserved population.

Trial Registration: ClinicalTrials.gov NCT02870413; https://clinicaltrials.gov/ct2/show/NCT02870413

(J Med Internet Res 2019;21(9):e13967) doi:10.2196/13967

KEYWORDS

telehealth; telemedicine; breastfeeding; lactation; lactation support; rural health

Introduction

Breastfeeding has multiple benefits for mothers, infants, and society, helping to protect against a variety of short- and long-term health conditions and reducing associated healthcare costs [1-4]. Unfortunately, the majority of women in the United States stop breastfeeding before it is recommended, with only 36% of infants receiving any breastmilk at 12 months of age [5]. There are also significant disparities in breastfeeding rates [6], with rural and low-income mothers less likely to initiate and maintain breastfeeding, thus putting them at greater risk for a range of negative health outcomes [7,8].

One strategy to improve breastfeeding rates, particularly among underserved populations, is to increase access to professional breastfeeding support by International Board-Certified Lactation Consultants (IBCLCs). In fact, the Surgeon General's Call to Action to Support Breastfeeding identifies increasing access to IBCLCs as a policy priority [9] because research has shown that they improve breastfeeding duration and exclusivity (ie, proportion of breastmilk in an infant's diet) [10-13]. Although a range of professionals provide breastfeeding support and address breastfeeding complications, IBCLCs are exclusively focused on breastfeeding in contrast to other medical professionals who must balance a range of competing demands during each visit [14,15]. In rural communities, poor access to IBCLCS may contribute to breastfeeding disparities [16-20].

Telelactation is one tool that can be leveraged to increase access to IBCLCs in rural settings [21]. Telelactation services connect breastfeeding mothers to remotely located IBCLCs through audio-visual technology. Several companies, including Medela, Lansinoh, American Well, and Pacify Health, currently offer telelactation video visits to breastfeeding women through their mobile apps [21,22], but the current published literature on this particular form of mobile telelactation is limited to one qualitative study on mothers' experiences [22]. Furthermore, there is no research on who uses telelactation, or the content, duration, and timing of these visits.

In the context of a pilot randomized controlled trial on telelactation in a rural population, we set out to address this gap in the literature. We aimed to use survey and electronic medical record (EMR) call log data to identify maternal characteristics associated with demand for and use of telelactation, to describe telelactation visit characteristics, and to describe experiences with visits.

Methods

Overview

From October 2016 to May 2018, we recruited rural women into a randomized controlled trial of telelactation. Participants were enrolled during their postpartum hospitalization at a critical access hospital in North-Central Pennsylvania and tracked for 12 weeks following enrollment. The hospital is located in a Health Professional Shortage Area (HPSA) within a county designated as 100% rural by the 2010 United States Census. At the time of the study, it had approximately 200 births per year and did not have any IBCLCs on staff or practicing within 40 miles of the hospital. At the time recruitment began, breastfeeding initiation rates among the hospital's patients (74%) were below the national average (80%) [23].

Women were eligible to participate if they were >18 years old, had an infant with a gestational age of > 35 weeks, had initiated breastfeeding and planned to continue after hospital discharge, and did not have a condition incompatible with breastfeeding (eg, HIV positive) [24]. Following informed consent, participants were enrolled into one of two study arms: telelactation (intervention) or usual care (control). Participants randomized to the telelactation arm were given an orientation to Pacify Health's telelactation app by hospital nurses involved in study recruitment. The nurses showed participants how to download the app on their personal device, provided a coupon code for free, unlimited video visits, and helped participants conduct a test call. After this orientation, participants could request unlimited, on demand video calls with IBCLCs through the app. The telelactation app used in the trial was designed for video calls to occur within seconds of visit initiation by the mother. In addition, it is Health Insurance Portability and Accountability Act (HIPAA)–compliant and involves a large network of geographically dispersed IBCLCs available to take video calls 24 hours a day. Participants randomized to the control arm received care as usual (ie, no additional lactation support beyond what is typically received in the hospital and in outpatient settings). These participants were not included in this analysis because our aim was to describe use of, and experiences with, telelactation among those given access to it. Control arm participants did not have access to these services.

Data and Sample

We used two data sources to assess use of and experiences with telelactation: participant survey data and EMR call log data supplied by the telelactation vendor.

Survey Data

We fielded online surveys to all study participants at baseline (enrollment in the hospital) and at 2, 4, and 12 weeks postpartum. All participants received questions on demographics (baseline survey), breastfeeding practices (all surveys), breastfeeding plans (baseline survey), work plans and status (baseline and 12-week surveys), and breastfeeding challenges (2-week and 12-week surveys). However, participants enrolled in the telelactation arm (n=102) received an additional module on experiences with the telelactation app at 2, 4, and 12 weeks. This module included questions on their use of the telelactation app, challenges with using the app, satisfaction with the app, and recommendations for improvement. For this study, we restricted our sample to telelactation arm participants who completed the final 12-week survey (n=94), which represents 92% of the subjects in this arm.

Electronic Medical Record Data

The telelactation vendor provided a log of all the video calls completed by study participants during the study period. For each video call with clinical content (n=83), the vendor supplied patient identification, call duration, date and time of call, and reason for call or issues discussed. We merged the individual-level survey data with the EMR call log data to examine how telelactation arm participants who used the app differed from those who did not.

Measures

We applied two definitions of telelactation use. First, we defined telelactation use based on vendor EMR data. Any participant with one or more documented video calls that addressed a breastfeeding challenge was classified as a telelactation user for the analyses comparing users to nonusers and characterizing visits. Second, for analyses on satisfaction and experiences with the app, we relied on participants' self-reported telelactation use. For this measure, participants who reported participating in one or more video calls during the study period in the 12-week survey, including for test or orientation calls that did not specifically address a breastfeeding issue, were classified as telelactation users. We applied the definition based on EMR data for most analyses because it is likely more accurate (ie, not subject to recall bias and social desirability bias) and captures meaningful use of the app. However, we also applied the definition based on participants' self-reported use in some analyses to represent the perspectives of all participants who engaged with the app in any manner. Furthermore, it is possible that select participants accessed IBCLC support on the app in a way that was not recorded by the vendor (eg, through an account of a healthcare provider involved in the study), and we wanted to capture the experiences of those participants.

We included two measures of participants' overall assessment of telelactation as reported in the final 12-week survey. First, participants were asked to rate, "How helpful was the help you received through the app?" (very helpful, helpful, neither helpful nor unhelpful, somewhat helpful, or not at all helpful). Second, they were asked to indicate, "How satisfied were you with the help you received through the app?" (very satisfied, satisfied, neither satisfied nor unsatisfied, somewhat satisfied, and not at all satisfied). For both measures, we collapsed the two positive categories (ie, very satisfied and satisfied) and three neutral or negative categories (eg, neither satisfied nor unsatisfied, somewhat satisfied, and not at all satisfied) to create binary variables: helpful (yes or no) and satisfied (yes or no). Finally, using vendor data on video call date and time, we constructed a measure of visit timing. All calls that occurred between 6:01 PM-7:59 AM were defined as occurring outside of regular business hours.

Statistical Approach

First, we examined whether participants who used telelactation differed from those who did not on multiple, self-reported, sociodemographic characteristics by using Chi-square tests and *t* tests for categorical and continuous variables, respectively. Second, we assessed telelactation visit characteristics. We calculated sample statistics and two-tailed *P* values to determine whether means across groups were statistically different using a *P* value of <.05. This study was reviewed and approved by the RAND Corporation's Institutional Review Board (IRB). The recruitment hospital deferred to RAND's IRB.

Results

Participants' Use of Telelactation

Of the 102 women randomized to the telelactation arm within the trial, 94 (92%) completed the final, 12-week survey. Of these, 47 (50%) self-reported participating in one or more video calls during the study period. Using EMR data, we identified 31 (33%) participants who completed one or more calls that included a substantive discussion of a breastfeeding challenge. Among these 31 participants who discussed at least one substantive challenge through the app, 14 (45%) had one call only, 8 (26%) had 2 calls, and 9 (29%) had 3 or more calls.

Among the 94 participants, 68 (72%) reported experiencing one or more breastfeeding challenges during the study period. Of the 26 women who did not report breastfeeding challenges, none of them participated in a substantive discussion about a breastfeeding challenge via the app. As such, the rate of uptake for substantive video calls among women who reported challenges (ie, reported a potential need for lactation support) was 45% (31/68).

In Table 1, we present descriptive statistics on the full sample of participants in the telelactation arm that are also stratified by telelactation user versus nonuser. Overall, the mean age of the maternal participants was 26 years (SD 5) at baseline. In addition, 44 (47%) had a high school degree or less, 45 (48%) had private health insurance during pregnancy, 39 (41%) were first-time mothers and slightly more than half (49/94, 52%) planned to work for pay within 12 months of delivery.

Table 1. Characteristics of telelactation users versus nonusers (N=94).

Characteristics	Nonusers (n=63), n (%)	Telelactation users (n=31), n (%)	*P* value
Maternal sociodemographics			
Marital status (married)	33 (52)	17 (55)	.82
Race (Caucasian)	60 (95)	30 (97)	.73
High school degree or less	31 (49)	13 (42)	.51
Smartphone ownership at baseline (yes)	61 (97)	30 (97)	.99
Private insurance during pregnancy	29 (46)	16 (52)	.62
Public insurance during pregnancy	31 (49)	14 (45)	.72
Childbirth and breastfeeding characteristics and plans			
Plan to work in baby's first year (yes)	29 (46)	20 (65)	.09
Working by 12 weeks	26 (41)	21 (68)	.02
First time mom (yes)	22 (35)	17 (55)	.07
Caesarean section delivery	28 (44)	16 (52)	.52
Baby sex (boy)	39 (62)	17 (55)	.52
Baby delivered > 37 weeks	54 (86)	27 (87)	.86
Prior breastfeeding experience (yes)	41 (65)	12 (39)	.02
Plans to breastfeed exclusively for 12+ weeks	54 (86)	26 (84)	.66
Breastfed exclusively while in hospital	51 (81)	16 (52)	<.001
Maternal health and risk factors			
Smoking at 12 weeks	12 (19)	6 (19)	.97
Depression (prepregnancy)	15 (24)	6 (19)	.63
Anxiety (prepregnancy)	21 (33)	6 (19)	.16
Obesity (prepregnancy)	5 (8)	4 (13)	.45
Hypertension (prepregnancy)	7 (11)	1 (3)	.20

Compared to participants who had one or more telelactation video calls that included discussion of a breastfeeding challenge, those who did not engage in calls were similar on most sociodemographic, maternal health, childbirth, and breastfeeding characteristics. However, participants who used telelactation were more likely to be working at 12 weeks postpartum (21/31, 68%; *P*=.02) compared to mothers who did not use telelactation (26/63, 41%) and less likely to have prior breastfeeding experience (12/31, 39%; *P*=.02) on average compared to nonusers (41/63, 65%). They were also less likely to have breastfed exclusively prior to discharge from the hospital (16/31, 52%; *P*<.001) relative to mothers who did not use telelactation (51/63, 81%), as assessed at the baseline survey during the postpartum hospitalization.

Characteristics of Video Calls

According to vendor EMR data, 31 participants in the telelactation arm completed a total of 83 video calls that addressed one or more breastfeeding challenges. Characteristics of these video calls are presented in Table 2. The average video call duration was 7 minutes and 19 seconds (SD 5.5 mins). Most calls (58/83, 70%) occurred during the infant's first month of life; however, 19% (16/83) of video calls took place nine or more weeks after the infant's birth. In total, 41% (34/83) of video calls occurred outside of regular business hours (defined as any time between 8 AM and 6 PM, excluding weekends), and 59% (49/83) addressed more than one breastfeeding challenge or concern. The most common challenges discussed on video calls included: breast pain, soreness, and infection (25/83, 30% of calls), use of nipple shields (21/83, 25%), latch or positioning (17/83, 24%), milk supply and production (14/83, 17%), and the use of a breast pump (14/83, 17%).

Table 2. Characteristics of video calls.

Characteristic	Value, n (%)
Age of infant when call occurred	
<7 days	37 (45)
2-4 Weeks	21 (25)
5-8 Weeks	11 (13)
9-12 Weeks	6 (7)
13+ Weeks	10 (12)
Time of day of call	
8 AM-12 PM	23 (28)
12 PM-4 PM	29 (35)
4 PM-8 PM	22 (27)
8 PM-8 AM	9 (11)
During business hours (yes)	49 (59)
Topics discussed on calls	
Breast pain, soreness, infection	25 (30)
Use of nipple shields	21 (25)
Latch or positioning	17 (24)
Milk supply and production	14 (17)
Breast pump use	14 (17)
Infant condition or health	13 (16)
Infant weight or whether getting enough milk	10 (12)
Mother’s return to work	9 (11)
Mother’s illness or medications interacting with breastfeeding	8 (10)
Mother’s dietary restrictions	7 (8)
Infant uninterested in eating	4 (5)

Experiences With and Attitudes About Telelactation Among Self-Reported Telelactation Users

According to survey data from the final 12-week assessment, 47/94 (50%) of participants who were given access to telelactation did not participate in any video calls. Leading reasons for not participating in any video calls for breastfeeding support during the study included: not experiencing any breastfeeding problems (26% of nonusers), not being comfortable with video calls (11%), no longer breastfeeding or stopped breastfeeding during the study (8%) and being too busy (7%).

At the 12-week assessment, the majority of telelactation users indicated that the breastfeeding assistance they received via the app was helpful (41/47, 87%) and expressed satisfaction with the help received (43/47, 91%). Of the 47 telelactation users, 9 (19%) had one or more recommendations for improving the app. Of these participants, four suggested that the app should allow mothers to text with IBCLCs and two users requested an option to do audio only calls. Each of the following recommendations was offered by one user, who suggested that the app should allow users: 1) to request a particular IBCLC; 2) to connect to breastfeeding peers; 3) to access blogs and other written resources on breastfeeding; or 4) to automatically schedule monthly calls that the mother does not need to initiate.

Over the course of the study, 10 participants reported attempting to use the app without success. We included these participants along with self-reported telelactation users (n=57) to identify any technical difficulties encountered in requesting or participating in video calls. A total of three participants (7%) reported difficulties in finding or connecting to wireless internet. No other technical problems (eg, dropped calls, poor sound quality, long wait for IBCLCs) were identified by more than one participant. Of all the participants, 36/47 (78%) did not report experiencing any technical difficulties.

Discussion

This study revealed relatively robust demand for telelactation, particularly in the first weeks after delivery, and user satisfaction with telelactation services. Video calls tended to be much shorter in duration (7 minutes on average) than a scheduled, in-person visit with an IBCLC. Also, video calls often took place outside of business hours, suggesting that participants were taking advantage of the scheduling flexibility supported by this delivery model. Telelactation users were relatively similar to nonusers;

however, users were less likely to be breastfeeding exclusively prior to discharge from the hospital and also less likely to have prior breastfeeding experience. These findings suggest that telelactation services may be particularly attractive to mothers who lack the confidence and skills that come with prior breastfeeding experience as well as those experiencing early difficulties at the time of initiation.

Forty-five percent of participants with self-reported breastfeeding challenges and 31% of all eligible participants requested and participated in substantive telelactation visits that addressed breastfeeding issues. In addition, 50% reported engaging with the app in some manner. This rate of uptake is dramatically higher than observed for other telehealth interventions offered to a population, and for telephone-based breastfeeding support (uptake rate of 0.5-24%) [25-28]. We attribute the relatively high demand for telelactation in the context of this study to several factors. First, according to baseline survey data on breastfeeding plans, the population of mothers that participated in the trial were committed to long duration, exclusive breastfeeding, and relatedly, to overcoming any challenges that arose. This was the case in part because women who did not intend to breastfeed following hospital discharge were excluded from the trial. At the time of enrollment, 84% (26/31) of telelactation arm participants indicated that they planned to breastfeed exclusively for 12 or more weeks, as compared to 57% of pregnant women nationally who reported plans to breastfeed exclusively for the first few weeks [29]. Second, telelactation services were introduced to potential users by trusted healthcare providers (ie, nurses caring for participants in person during the postpartum hospitalization), and participants were given an orientation to the app at the time of enrollment that often included a test call.

While we could not identify any research that describes the content of telelactation visits via personal devices, numerous studies both within and outside of the United States have assessed telephone breastfeeding support visits [30-35]. The top three issues discussed in video calls in this study included: breast pain, soreness, or infection, use of nipple shields, and latch or positioning. No prior analyses of telephone visits identified nipple shield use as a common topic, however, main reasons for visits seemed to vary depending on the patient population. For example, common issues discussed by 908 callers to a helpline in the United Kingdom included difficulties with positioning and concerns about inadequate milk supply [30]. In contrast, common issues identified in an analysis of 1969 calls to an inner-city hospital breastfeeding support line in the United States included questions about obtaining and using a breast pump, and breast issues [31]. Numerous questions about nipple shields in this study may reflect the local practices of healthcare professionals at the recruitment hospital during the postpartum hospitalization.

This work builds upon our prior qualitative work that showed that telelactation is acceptable and feasible for rural mothers. While select participants in both analyses recommended the addition of audio only and text message visit options, one participant in the current analysis also indicated a desire for regularly scheduled video calls that the breastfeeding mother did not have to initiate. At this point, this form of telelactation requires the mother to identify and seek help for a breastfeeding issue. Prior research has differentiated between proactive (prescheduled at regular intervals) and reactive (as demanded by the mother) breastfeeding support and suggested that proactive support can lead to greater engagement and impact [36]. Future telelactation models can consider how best to deliver services that are responsive to urgent needs but also require less activation from mothers who could benefit from professional advice.

The primary limitation of this study is that we recruited participants at one study site that serves a population of rural, predominantly Caucasian mothers in Pennsylvania. As such, it is unclear how patterns of telelactation use may differ in other communities with different breastfeeding support services. Nonetheless, the study was conducted in the context of a randomized controlled trial and is the first of its kind to use quantitative methods to explore use of and experiences with telelactation.

Telelactation is an innovation in the delivery of professional breastfeeding support. Although our study focused on a rural population, these services may increase convenience and reduce costs associated with seeking in-person breastfeeding support in urban settings as well. Although additional research should document the impacts of these services on breastfeeding outcomes and healthcare costs within more diverse populations, this research documents demand for and positive experiences with telelactation in an underserved population.

Acknowledgments

This research was funded by the Health Resources and Services Administration (HRSA) grant #R40MC294510100.

Conflicts of Interest

None declared.

References

1. Stuebe AM, Jegier BJ, Schwarz EB, Green BD, Reinhold AG, Colaizy TT, et al. An Online Calculator to Estimate the Impact of Changes in Breastfeeding Rates on Population Health and Costs. Breastfeed Med 2017 Dec;12(10):645-658. [doi: 10.1089/bfm.2017.0083] [Medline: 28906133]
2. Stuebe AM, Rich-Edwards JW, Willett WC, Manson JE, Michels KB. Duration of lactation and incidence of type 2 diabetes. JAMA 2005 Nov 23;294(20):2601-2610. [doi: 10.1001/jama.294.20.2601] [Medline: 16304074]

3. Olson CM, Strawderman MS, Hinton PS, Pearson TA. Gestational weight gain and postpartum behaviors associated with weight change from early pregnancy to 1 y postpartum. Int J Obes Relat Metab Disord 2003 Jan;27(1):117-127. [doi: 10.1038/sj.ijo.0802156] [Medline: 12532163]
4. Østbye T, Peterson BL, Krause KM, Swamy GK, Lovelady CA. Predictors of postpartum weight change among overweight and obese women: results from the Active Mothers Postpartum study. J Womens Health (Larchmt) 2012 Feb;21(2):215-222 [FREE Full text] [doi: 10.1089/jwh.2011.2947] [Medline: 22092110]
5. Centers for Disease Control and Prevention. 2019. Breastfeeding Report Card 2018 URL: https://www.cdc.gov/breastfeeding/data/reportcard.htm [accessed 2019-08-22]
6. Khoury A, Moazzem S, Jarjoura C, Carothers C, Hinton A. Breast-feeding initiation in low-income women: Role of attitudes, support, and perceived control. Womens Health Issues 2005;15(2):64-72. [doi: 10.1016/j.whi.2004.09.003] [Medline: 15767196]
7. Bartick M, Jegier B, Green B, Schwarz E, Reinhold A, Stuebe A. Disparities in Breastfeeding: Impact on Maternal and Child Health Outcomes and Costs. J Pediatr 2017 Feb;181:49-55.e6. [doi: 10.1016/j.jpeds.2016.10.028] [Medline: 27837954]
8. Chapman D, Pérez-Escamilla R. Breastfeeding among minority women: moving from risk factors to interventions. Adv Nutr 2012 Jan;3(1):95-104 [FREE Full text] [doi: 10.3945/an.111.001016] [Medline: 22332107]
9. US Department of Health and Human Services. US Department of Health and Human Services: Office of the Surgeon General; 2011. The Surgeon General's Call to Action to Support Breastfeeding URL: https://www.hhs.gov/surgeongeneral/reports-and-publications/breastfeeding/index.html [accessed 2019-08-22]
10. Morris C, Gutowski J. The Effect of an International Board Certified Lactation Consultant in the Pediatric Primary Care Setting on Breastfeeding Duration and Exclusivity During the First Year of Life. Clin Lactation 2015 Aug 01;6(3):109-116. [doi: 10.1891/2158-0782.6.3.109]
11. Renfrew MJ, McCormick FM, Wade A, Quinn B, Dowswell T. Support for healthy breastfeeding mothers with healthy term babies. Cochrane Database Syst Rev 2012 May 16(5):CD001141 [FREE Full text] [doi: 10.1002/14651858.CD001141.pub4] [Medline: 22592675]
12. Witt A, Smith S, Mason M, Flocke S. Integrating routine lactation consultant support into a pediatric practice. Breastfeed Med 2012 Feb;7(1):38-42 [FREE Full text] [doi: 10.1089/bfm.2011.0003] [Medline: 21657890]
13. Bonuck K, Stuebe A, Barnett J, Labbok M, Fletcher J, Bernstein P. Effect of primary care intervention on breastfeeding duration and intensity. Am J Public Health 2014 Feb;104 Suppl 1:S119-S127. [doi: 10.2105/AJPH.2013.301360] [Medline: 24354834]
14. Freed G, Clark S, Sorenson J, Lohr J, Cefalo R, Curtis P. National assessment of physicians' breast-feeding knowledge, attitudes, training, and experience. JAMA 1995 Feb 08;273(6):472-476. [doi: 10.1001/jama.1995.03520300046035] [Medline: 7837365]
15. Feldman-Winter L, Barone L, Milcarek B, Hunter K, Meek J, Morton J, et al. Residency curriculum improves breastfeeding care. Pediatrics 2010 Aug;126(2):289-297. [doi: 10.1542/peds.2009-3250] [Medline: 20603262]
16. Flower F, Willoughby M, Cadigan R, Perrin E, Randolph G, Family Life Project Investigative Team. Understanding breastfeeding initiation and continuation in rural communities: a combined qualitative/quantitative approach. Matern Child Health J 2008 May;12(3):402-414 [FREE Full text] [doi: 10.1007/s10995-007-0248-6] [Medline: 17636458]
17. Wiener R, Wiener M. Breastfeeding prevalence and distribution in the USA and Appalachia by rural and urban setting. Rural Remote Health 2011;11(2):1713 [FREE Full text] [Medline: 21534710]
18. Flood J. Breastfeeding Supports and Services in Rural Hawaii: Perspectives of Community Healthcare Workers. Nurs Res Pract 2017;2017:6041462 [FREE Full text] [doi: 10.1155/2017/6041462] [Medline: 28168053]
19. Dumphy D, Thompson J, Clark M. A Breastfeeding Quality Improvement Project in Rural Primary Care. J Hum Lact 2016 Nov;32(4):633-641. [doi: 10.1177/0890334416662240] [Medline: 27502513]
20. Ray K, Demirci J, Uscher-Pines L, Bogen D. Geographic Access to International Board-Certified Lactation Consultants in Pennsylvania. J Hum Lact 2019 Feb;35(1):90-99. [doi: 10.1177/0890334418768458] [Medline: 29969344]
21. Uscher-Pines L, Mehrotra A, Bogen D. The emergence and promise of telelactation. Am J Obstet Gynecol 2017 Aug;217(2):176-178.e1. [doi: 10.1016/j.ajog.2017.04.043] [Medline: 28483569]
22. Demirci J, Kotzias V, Bogen D, Ray K, Uscher-Pines L. Telelactation Via Mobile App: Perspectives of Rural Mothers, Their Care Providers, and Lactation Consultants. Telemed J E Health 2018 Sep 13:1-18. [doi: 10.1089/tmj.2018.0113] [Medline: 30212280]
23. Centers for Disease Control and Prevention. 2015. Breastfeeding Report Card 2014 URL: https://www.cdc.gov/breastfeeding/pdf/2014breastfeedingreportcard.pdf [accessed 2019-08-22]
24. Eidelman AI. Breastfeeding and the use of human milk: an analysis of the American Academy of Pediatrics 2012 Breastfeeding Policy Statement. Breastfeed Med 2012 Oct;7(5):323-324. [doi: 10.1089/bfm.2012.0067] [Medline: 22946888]
25. Uscher-Pines L, Mehrotra A. Analysis of Teladoc use seems to indicate expanded access to care for patients without prior connection to a provider. Health Aff (Millwood) 2014 Feb;33(2):258-264. [doi: 10.1377/hlthaff.2013.0989] [Medline: 24493769]

26. Osman H, Chaaya M, El Zein L, Naassan G, Wick L. What do first-time mothers worry about? A study of usage patterns and content of calls made to a postpartum support telephone hotline. BMC Public Health 2010 Oct 15;10:611 [FREE Full text] [doi: 10.1186/1471-2458-10-611] [Medline: 20946690]
27. Hoddinott P, Craig L, Maclennan G, Boyers D, Vale L, NHS Grampianthe University of Aberdeen FEST project team. Process evaluation for the FEeding Support Team (FEST) randomised controlled feasibility trial of proactive and reactive telephone support for breastfeeding women living in disadvantaged areas. BMJ Open 2012;2(2):e001039 [FREE Full text] [doi: 10.1136/bmjopen-2012-001039] [Medline: 22535794]
28. Rae M, Cox C. Kaiser Family Foundation. 2018 Oct 3. More employers are paying for telemedicine, but enrollee take-up has been relatively low URL: https://www.healthsystemtracker.org/brief/more-employers-are-paying-for-telemedicine-but-enrollee-take-up-has-been-relatively-low/ [accessed 2019-08-22]
29. Centers for Disease Control and Prevention. 2015. Infant Feeding Practices Study II and Its Year Six Follow-Up URL: http://www.cdc.gov/breastfeeding/data/ifps/index.htm [accessed 2019-08-22]
30. Thomson G, Crossland N, Dykes F, Sutton C. UK Breastfeeding Helpline support: An investigation of influences upon satisfaction. BMC Pregnancy Childbirth 2012 Dec 13;12:150 [FREE Full text] [doi: 10.1186/1471-2393-12-150] [Medline: 23234293]
31. Chamberlain L, Merewood A, Malone K, Cimo S, Philipp B. Calls to an inner-city hospital breastfeeding telephone support line. J Hum Lact 2005 Feb;21(1):53-58. [doi: 10.1177/0890334404272512] [Medline: 15681637]
32. Mullen S, Marshall A, Warren M. Statewide Breastfeeding Hotline Use Among Tennessee WIC Participants. J Nutr Educ Behav 2017;49(7 Suppl 2):S192-S196.e1. [doi: 10.1016/j.jneb.2017.04.024] [Medline: 28689557]
33. Gallegos D, Cromack C, Thorpe K. Can a phone call make a difference? Breastfeeding self-efficacy and nurse responses to mother's calls for help. J Child Health Care 2018 Sep;22(3):433-446. [doi: 10.1177/1367493518757066] [Medline: 29514511]
34. Parrilla-Rodríguez AM, Dávila Torres R, Gorrín-Peralta JJ. Profile of calls to a breastfeeding clinic information and help telephone line. P R Health Sci J 2001 Dec;20(4):377-381. [Medline: 11845670]
35. Wang S, Chen C, Chen C. Related factors in using a free breastfeeding hotline service in Taiwan. J Clin Nurs 2008 Apr;17(7):949-956. [doi: 10.1111/j.1365-2702.2007.02111.x] [Medline: 18321292]
36. Ericson J, Flacking R, Udo C. Mothers' experiences of a telephone based breastfeeding support intervention after discharge from neonatal intensive care units: a mixed-method study. Int Breastfeed J 2017 Dec;12:50 [FREE Full text] [doi: 10.1186/s13006-017-0142-9] [Medline: 29270208]

Abbreviations

EMR: electronic medical record
HIPAA: Health Insurance Portability and Accountability Act
HPSA: Health Professional Shortage Area
HRSA: Health Resources and Services Administration
IBCLC: International Board-certified Lactation Consultant
IRB: Institutional Review Board

Edited by G Eysenbach; submitted 08.03.19; peer-reviewed by G Signorelli, E van der Velde; comments to author 26.06.19; revised version received 03.07.19; accepted 05.07.19; published 03.09.19.

Please cite as:
Kapinos K, Kotzias V, Bogen D, Ray K, Demirci J, Rigas MA, Uscher-Pines L
The Use of and Experiences With Telelactation Among Rural Breastfeeding Mothers: Secondary Analysis of a Randomized Controlled Trial
J Med Internet Res 2019;21(9):e13967
URL: https://www.jmir.org/2019/9/e13967/
doi:10.2196/13967
PMID:31482848

Original Paper

Application of a Blockchain Platform to Manage and Secure Personal Genomic Data: A Case Study of LifeCODE.ai in China

Xiao-Ling Jin[1], PhD; Miao Zhang[1], MD; Zhongyun Zhou[2], PhD; Xiaoyu Yu[1,3], PhD

[1]Management School, Shanghai University, Shanghai, China
[2]Department of Management Science and Engineering, School of Economics and Management, Tongji University, Shanghai, China
[3]SHU Center for Innovation and Entrepreneurship, Shanghai University, Shanghai, China

Corresponding Author:
Zhongyun Zhou, PhD
Department of Management Science and Engineering, School of Economics and Management, Tongji University
1410 Tongji Building (Block A), 1500 Siping Road, Tongji University
Shanghai,
China
Phone: 86 02165981443
Fax: 86 02165981443
Email: philzhou@tongji.edu.cn

Abstract

Background: The rapid development of genetic and genomic technologies, such as next-generation sequencing and genome editing, has made disease treatment much more precise and effective. The technologies' value can only be realized by the aggregation and analysis of people's genomic and health data. However, the collection and sharing of genomic data has many obstacles, including low data quality, information islands, tampering distortions, missing records, leaking of private data, and gray data transactions.

Objective: This study aimed to prove that emerging blockchain technology provides a solution for the protection and management of sensitive personal genomic data because of its decentralization, traceability, encryption algorithms, and antitampering features.

Methods: This paper describes the case of a blockchain-based genomic big data platform, LifeCODE.ai, to illustrate the means by which blockchain enables the storage and management of genomic data from the perspectives of data ownership, data sharing, and data security.

Results: Blockchain opens up new avenues for dealing with data ownership, data sharing, and data security issues in genomic big data platforms and realizes the psychological empowerment of individuals in the platform.

Conclusions: The blockchain platform provides new possibilities for the management and security of genetic data and can help realize the psychological empowerment of individuals in the process, and consequently, the effects of data self-governance, incentive-sharing, and security improvement can be achieved. However, there are still some problems in the blockchain that have not been solved, and which require continuous in-depth research and innovation in the future.

(J Med Internet Res 2019;21(9):e13587) doi:10.2196/13587

KEYWORDS

genomics big data platform; blockchain; data ownership; data sharing; data security; digital health

Introduction

Background

The rapid development of next-generation sequencing and genome editing has dramatically reduced the cost of single genome-wide testing. This technology has facilitated the global adoption and wide application of gene sequencing, a technique that once existed only in cutting-edge research laboratories, in medical and health care institutions such as hospitals and clinics around the world. Its rise reflects the increasing interest among both researchers and practitioners in the interdependence between gene data analysis and medical treatment. Moreover, it has been demonstrated that gene sequencing has many applications, such as enhancing chronic disease prediction, new drug development, and precision medicine [1-4].

Despite the benefits of gene sequencing, the storage, transmission, and management of genomic big data in a secure fashion presents a great challenge. This is a critical issue for the development of gene sequencing and has thus been of interest to both researchers and practitioners. To address this issue, an increasing number of companies and institutions are using advanced computational tools, including big data analysis and cloud computing, to construct genomics big data platforms to integrate personal genomics data [5,6]. These platforms offer support services to a variety of user groups (eg, patients, medical institutions, and insurance companies) in domains such as data archiving, computational analysis, knowledge search, management authorization, and association visualization, and they have been leveraged for chronic disease prediction, new drug research and development (R&D), and precision medicine [7].

Challenges of Genomics Big Data Platforms

The core function of genomics big data platforms is to collect personal genetic data for application and sharing. However, the collection and sharing of genetic data is often challenged by problems such as low data quality, information islands, tampering distortions, missing records, leaking of individual privacy, and gray data transactions. The key challenges that hinder the development of genomics big data platforms can be classified into 3 facets: data ownership, data sharing mechanisms, and data security [8-11].

Unclear Ownership of Data

It is difficult to define data ownership due to its replicability and virtual nature. As no laws or regulations provide detailed guidelines, the ownership of personal health data remains controversial, and various stakeholders have different opinions on data ownership issues, especially in China. It is thus imperative to correctly determine the ownership of medical data.

Insufficient Data Sharing

Data can produce value with interoperability and sharing, and data analysis is an important way to promote the development of health care and medicine [12]. However, personal data owners are reluctant to share genomic data without effective sharing incentives and without a clear value exchange [13,14]. As a result, the benign self-running sharing mechanism design of the data platform has great importance.

Privacy Leakage and Insecurity

Human DNA contains extremely sensitive and private information [15], so individuals’ privacy may be jeopardized when genomic and medical data are spread or shared. Health data have traditionally been stored by a single organization in a centralized fashion, but the data may be taken without the users’ consent for unintended purposes, resulting in a threat to the data owners [16,17]. Hence, privacy protection and security are fundamental to the collection and application of genomic big data.

Application of Blockchain

Blockchain is an emerging technology that has attracted increasing attention from both researchers and practitioners. Briefly, a blockchain is a publicly distributed ledger that seals blocks with timestamps and encrypted hash links in a secure and immutable manner [18,19]. It enables transaction processes without the need for a trusted third party because of its traceability, smart contracts, resistance to tampering, decentralization, and encryption algorithm [20]. First, a blockchain uses timestamp technology to achieve data traceability and verifiability, which means that it offers a secure and transparent method to track the ownership of assets before, during, and after any transaction [21]. Second, the open-source sharing protocol built into blockchain enables simultaneous data logging and storage by all participants, which ensures that the details of the recorded transactions cannot be retroactively changed without the full agreement of the network [22,23]. Finally, a blockchain’s architecture and governance are decentralized, which makes it highly fault-tolerant and resistant to attacks and collusion. Moreover, a blockchain can protect users’ privacy with the use of cryptographic hash functions and asymmetric encryption to enable users to encrypt data with their own private key [24-26].

As such, a blockchain provides a new method of solving the problems encountered in the construction of genomics big data platforms. Specifically, it can empower a genomics big data platform in at least 3 ways.

Traceability Empowers Data Ownership

Traceability is the ability to verify an item’s history, location, or application by means of documented recorded identification [27]. Blockchain uses timestamp technology to make data traceable and verifiable, so it completely records the entire process from data generation to final storage [28]. Thus, each piece of data on the blockchain can be determined to be owned by the individual data producer at the time of generation, and this data ownership is verifiable based on the record.

Antitampering and Smart Contracts Empower Data Sharing

A smart contract contains a set of rules that help the parties to the smart contract interact automatically [29]. No record on the blockchain can be tampered with by any individual node [29]. The smart contract code facilitates, verifies, and enforces the negotiation or performance of an agreement or transaction, and the antitampering feature helps implement access and protect the data’s originality, thus ensuring the operation of smart contracts [30-32]. Therefore, the application of these 2 characteristics of blockchain can improve the data sharing mechanism, which then stimulates the platform’s participants to share data for extrinsic benefits.

Decentralization and Encryption Algorithms Empower Data Security and Privacy Protection

Decentralized storage helps to reduce security risks, increase trust, and manage data [33]. An encryption algorithm is a component for electronic data transport security. The blockchain encryption algorithm encrypts the data and strongly guarantees data security and privacy [34]. For example, Sun et al [35] proposed the use of blockchain technology to protect intellectual property and data security and privacy in internet education.

Objective and Research Questions

The research on blockchain in the field of health care focuses primarily on technical aspects such as the algorithm model, feasible solutions, and structural design. For example, Yue et al [36] designed an app architecture that enables patients to own, control, and share their data easily and securely without violating their privacy by using blockchain-based platform. Zhang et al [37] implemented a blockchain-based decentralized app (DApp) to address interoperability challenges in smart health care. On the basis of blockchain developed for telecare medical information systems, Ji et al [38] proposed a novel blockchain-based multilevel privacy-preserving location-sharing scheme to realize location sharing while preserving privacy. Few studies, however, have focused on revealing how to apply blockchain technology to empower genetic engineering. With this motivation, this study aims to address the following research question:

> *RQ: How does blockchain technology empower the storage and management of genomics data in a genomics big data platform?*

To answer this question, we report a case study of a blockchain-based genomics big data platform in China, LifeCODE.ai, to explore the effectiveness of blockchain in the management and security of personal genomic data.

Methods

Case Study Method

To address the proposed research question, we used case study as the research method because of the following reasons. First, the application of a blockchain platform to manage and secure personal genomic data is a complicated process, the adoption process concerns the *how* question, and the case study method is suitable for a process-based analysis [39]. In addition, the case study research method has a practical nature and can increase the effectiveness of the evidence [40]. Finally, the case study method has a strong descriptive nature, conveys a clear analysis of events and a strong sense of reality, and is easily understood by readers [41-43].

Given the research question, the selection of a case was subject to certain conditions. First, the case organization needed to be a business in biomedical health care and/or medicine. Second, the case organization was required to apply a blockchain platform to address genetic data management and security issues. Third, the case organization had to be willing to accept repeated surveys.

On the basis of these conditions, we selected LifeCODE.ai, a genomics big data platform based on blockchain technology, as the object of this study. LifeCODE.ai was first released at the TechCrunch International Innovation Summit 2018. It uses a decentralized consensus approach to construct a blockchain platform to aggregate distributed health data. With the ultimate goal of improving overall health outcomes via genomics research, LifeCODE.ai comprises a secure, decentralized, and visible personal health data center that removes the boundaries of health information between hospitals, pharmaceutical R&D institutions, doctors, patients, and individuals. At the same time, LifeCODE.ai released to individual users the blockchain-based DApp, Laiyin Health, which also applies blockchain for data self-governance implementation, token mechanism design, and encryption algorithm application. Figure 1 shows the interfacial design of this DApp. Unlike traditional apps whose backend code is run on centralized servers, the backend code of a DApp is run on a decentralized peer-to-peer network [44]. LifeCODE.ai is available on the official website [45].

On-site data were then collected via 2 interviews of the company that developed and manages LifeCODE.ai, its departmental managers were interviewed, and the operation of LifeCODE.ai and the application of blockchain in 2018 were discussed. The interviewees included corporate executives and department managers at LifeCODE.ai, and each interview lasted approximately 60 to 120 min. Each interview was digitally recorded and transcribed. At the same time, we systematically collected secondary data about this platform from a variety of sources, including newspapers, magazines, books, and the internet.

Figure 1. Laiyin Health app interface.

Features and Services

LifeCODE.ai and its synchronous online client DApp apply blockchain technology in an innovative manner to establish the legality and effectiveness of large-scale health data collection while protecting user privacy and creating data interaction value. As such, they truly realize the propertization and assetization of large health data. The main features of LifeCODE.ai are as follows:

- It ensures that health data are owned only by the individuals who upload their own health data and can never be changed without the owner’s acknowledgment.
- It establishes a trading mechanism for data exchange via tokens; these tokens are issued by the LifeCODE Databank, which was designed in accordance with Ethereum’s ERC-20 protocol, with a maximum issuance of 3,000,000,000 pieces. Medical and health institutions that wish to use health data must obtain the owners’ consent and pay a certain fee (ie, tokens) to *buy* the data. This positive mechanism can greatly encourage users to share their health data.
- The health data in LifeCODE.ai are stored on decentralized distributed nodes. All data and data transactions in the blockchain network are encrypted and traceable. The LifeCODE.ai platform also has a unique encryption algorithm and consensus mechanism that enable the resistance to tampering and traceability of health data, thus guaranteeing the privacy of the users’ health data.

LifeCODE.ai advocates blockchain as a service, which makes the use of blockchain as easy as surfing the internet. LifeCODE.ai mainly provides the following 3 product services:

- Smart contracts: It decouples decentralized smart contracts based on blockchain technology.
- Bookkeeping function: It uses distributed bookkeeping to address the lack of trust. Without a central organization, all parties have the same book to ensure that an open and transparent transaction process.
- Computational processing functions: It coordinates platform resources, processes gene sequencing data, and enables the data to be used.

Platform Architecture

LifeCODE.ai integrates genomic data and phenotypic health data with the highest level of data quality and privacy to create a platform for data exchange, interoperability, and a wide range of services for all participants. The platform’s architecture is roughly composed of 4 layers (as shown in Figure 2): an infrastructure layer, a data layer, a blockchain layer, and an interface layer. The interface layer contains the application programming interface (API) and related business services for customers. The blockchain layer forms the platform’s core. In addition to traceability, antitampering, decentralization, smart contracts, and encryption algorithms, it also has a key management mechanism and a CAM-brain Engine. The data layer includes searchable encryption and trusted data storage that provide further protection for data security. The infrastructure layer contains basic services, among others, and a security module that comprises a network of access control, basic security service, and so forth. The blockchain layer, data layer, and infrastructure layer are used together to execute secure storage, complex algorithms, and other functions at the back end, thus providing stable services in the interface layer at the front end.

Figure 2. LifeCODE.ai platform architecture (adapted from LifeCODE.ai White Paper).

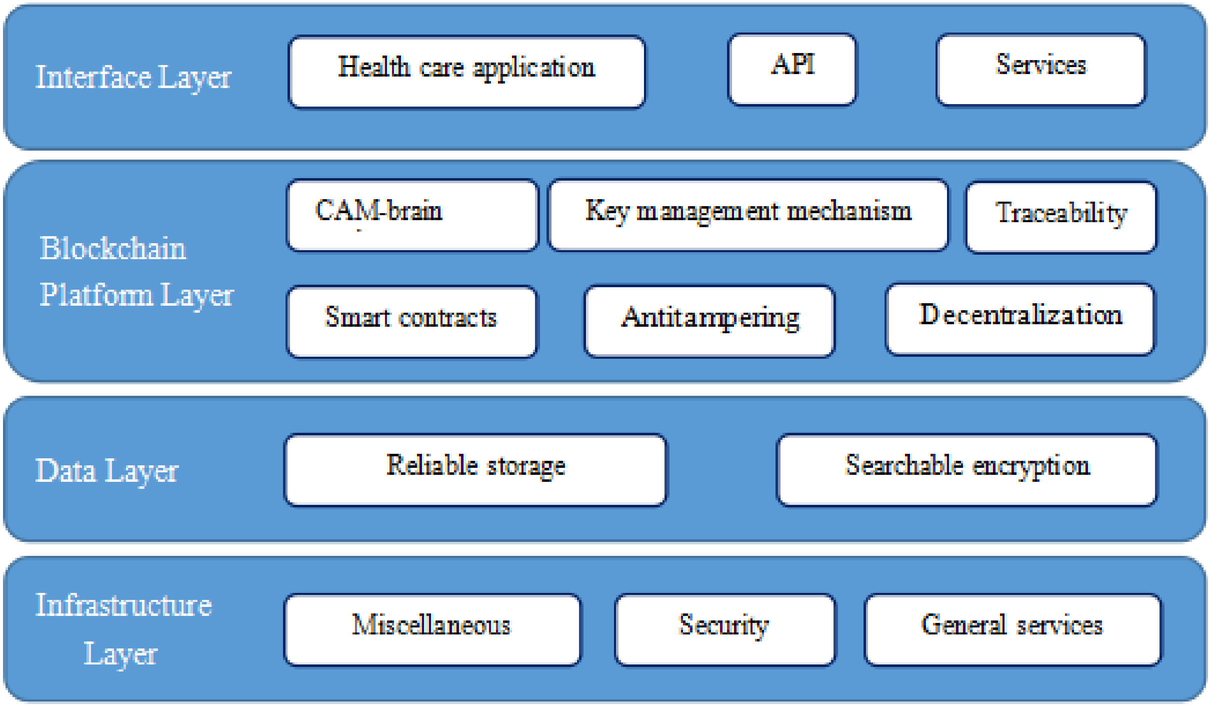


Table 1. Blockchain empowerment in genomics big data platform.

Empowered facets	Blockchain features	LifeCODE.ai solution	References
Data ownership	Traceability	LifeCODE.ai guarantees that the data belong only to individuals who upload their own health data. Only these individuals hold the private key for data management.	[12,13], [26,27]
Data sharing	Smart contracts, Antitampering	LifeCODE.ai establishes a set of conversion mechanisms between health data and assets. If platform participants wish to share or use data, they must follow the asset-based operating mechanism (ie, token mechanism).	[14,15], [28-31]
Data security	Decentralization, Encryption algorithm	LifeCODE.ai utilizes the features of blockchain technology to form privacy protection mechanism and a database with reliable storage, which guarantees the data security.	[16,17], [32-35]

Case Analysis

The purpose of this study was to investigate how blockchain technology can empower the storage and management of genomics data in genomics big data platforms. The case evidence was analyzed across 3 facets as shown in Table 1: determining data ownership, establishing a data-sharing mechanism, and ensuring data security and privacy.

Determining Data Ownership

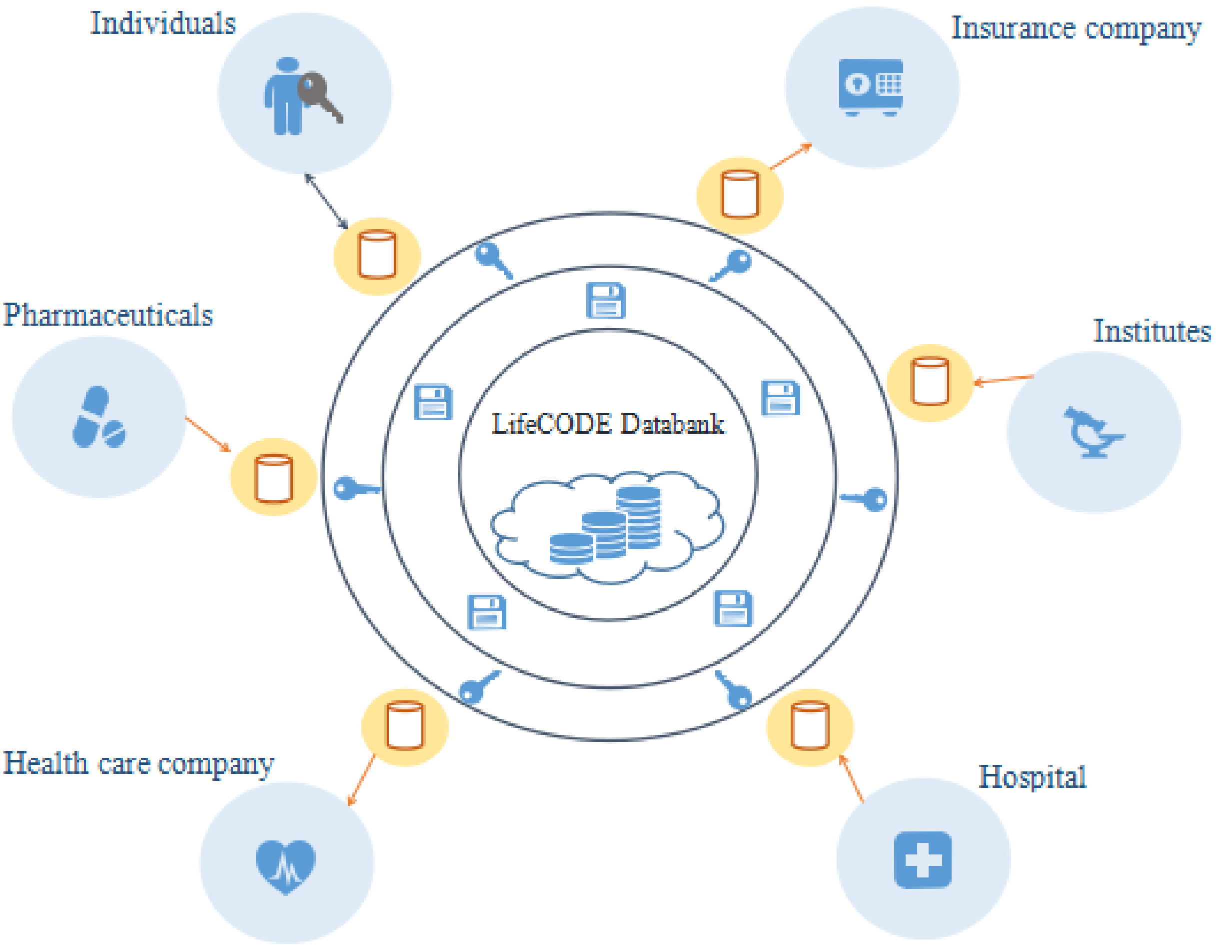
In LifeCODE.ai, the ownership of gene data absolutely belongs to the individual. Typically, once an individual uploads his or her personal genomic and phenotypic health data to the platform, the data are encrypted immediately and fully by default, and the business participants can access the data only with the owner's permission. This model makes the platform just like a bank that stores and manages data. In the traditional case, although the gene data comprise individuals' data, they were often managed by institutions such as hospitals. However, in the LifeCODE.ai blockchain, the rights are returned to the individuals, and each person's data are stored on the platform and remain within his or her control such as a deposit in a bank. This decentralization of rights will give individuals peace of mind.

The technology to implement the determinacy of data ownership in individuals is blockchain. Specifically, in this process, the public and private key mechanism in the blockchain defines the data producer as the data owner and fully encrypts the data, thereby performing rights division and data protection. This private key mechanism allows data holders to securely store, access, and share their data via encryption keys under their own authorization. At the same time, traceability allows the blockchain to strictly monitor any access to the data and record it in LifeCODE blockchain to make all data transactions traceable. Thus, monitoring each processing record of the data is an operation authorized by the data owner, and the rights are locked in the individual's hands, which will promote data self-governance. However, the autonomous nature of the blockchain plays down the requirement of state supervision. In the case that the supervision cannot be reached, some profit-seeking markets will lead the application of blockchain technology to the illegal field, thus causing certain risks.

Establishing a Data-Sharing Mechanism

LifeCODE.ai uses a token mechanism to enable data trading and sharing in a platform-based closed-loop ecosystem. Individuals, hospitals, research institutions, insurance companies, health care companies, and pharmaceutical companies are the major players in the LifeCODE ecosystem (as shown in Figure 3). The individuals are the data owners and hold the private key for data management, and other participants can communicate and trade with them to acquire or apply their data. For instance, if a research institution wishes to obtain 100 pieces of data from diabetes patients for drug R&D, they can search keywords such as *diabetes* in the platform. They can then obtain anonymous contact information of the relevant individuals, thereby initiating a data application and making them a token offer. The individuals can either approve the deal or reject it. The gained token through data transaction can be used by the individuals to obtain services from medical institutions and insurance institutions. This token exchange mechanism may enhance individuals' positive motivations to share their personal genomics data.

Figure 3. LifeCODE.ai operating mechanism.



To use the token mechanism to achieve data sharing, LifeCODE.ai not only uses the blockchain's smart contract and antitampering features, it also effectively combines other technologies. In the process of data sharing, the smart contract code facilitates, verifies, and enforces the negotiation or performance of an agreement or transaction. The antitampering feature helps implement access and protect the originality of the data, thus ensuring the operation of smart contracts. The token mechanism involves 4 processes.

First, the data owner's data are imported into the genomic ordered relational (GOR) architecture, which is a system specifically designed for genomic data storage, query, search, indexing, and many other analyses. The GOR architecture and the GOR pipe, the Clinical Sequence Analyzer (API), and Risk Engine API it provides can analyze the data to derive potential correlations between the genetic data and clinical phenotypes or diseases.

Second, according to the requirements, the data are fully or partially encrypted and imported into a searchable encryption database that was developed on the basis of asymmetric encryption algorithms, tag-based fingerprint extraction, and homomorphic encryption. Here, the data generate a corresponding label as an index to search by keyword.

Third, LifeCODE.ai aggregates, classifies, integrates, and indexes the phenotype data. As these data are so variable, LifeCODE.ai uses multiple approaches to manage and integrate them: (1) HL7-compatible specifications, currently under design, will be used to build general exchangeable and extensible data structures for both high-quality phenotype data and that from less-developed systems; (2) data from trusted sources verified by the permissioned LifeCODE blockchain are aggregated, indexed, and linked by logic-based and experience-based protocols and adapters; and (3) a master patient index mechanism is used to link an individual's phenotypic data to an indexed keychain.

Finally, during the data transaction process, the secure transmission of data is the last layer of technical support for the sharing mechanism. The LifeCODE blockchain is a permissioned blockchain based on JP Morgan Quorum platform, an enterprise-focused version of Ethereum. Permissioned blockchain maintain an access control layer that allows certain actions to be performed only by certain identifiable participants. These blockchains differ from public and private blockchains.

However, the inability to tamper with the blockchain is both an advantage and a disadvantage. Due to the antitampering nature, once the incorrect account is transferred, it cannot be modified. The loss of the key will cause irreparable damage as well. In addition, transaction data will always be generated in the data-sharing project. As the body of transaction data continues to increase, the time required for the node to record the transaction data is also prolonged; the performance cannot be improved, and the efficiency will continue to decrease. Capacity will thus become a big problem for blockchain.

Ensuring Data Security and Privacy

The importance of genetic data security privacy cannot be overstated, so LifeCODE.ai adopts a combination of software and hardware based on blockchain to ensure data security. Data security and privacy can be summarized as 2 aspects: decentralized data storage and the data-sharing key mechanism. It is conceivable that reliable security can increase people's sense of security or trust.

First, LifeCODE.ai guarantees the highest-level protection of data authenticity, privacy, and permanent data storage. All owner-authorized health data and associated data exchanges are securely recorded and stored to prevent data leakage, abuse, and loss. Specifically, LifeCODE.ai leverages the decentralization of the blockchain to offer reliable storage (object storage service, OSS/Hadoop Distributed File System, HDFS/network file system, NFS) of various types of health data in a secure cloud environment, matched with various forms of data-persistence implementations. OSS provides highly extensible and robust storage for genomic data. The HDFS is incorporated by LifeCODE.ai for high-performance storage, such as structured health data and complex mapping and correlation. The NFS and local storage are used by worker nodes to run analysis services and the API and to provide low-latency high-throughput data access.

Moreover, LifeCODE.ai exports the health data to an intermediate searchable-encrypted data repository for data analysis and research purposes, and only there can the data be searched. At the same time, the privacy-sensitive portion of the data is encrypted in the same way as the original data, and only data that cannot cause damage to the data owner can be searched. The LifeCODE.ai platform then cryptographically conceals the genomic assets of a transaction by applying *zero knowledge proofs*, which means that party A can prove to party B that he knows the specific information without revealing the information; thus, party A is the prover, and party B is the verifier. This feature is especially useful in cryptography because it provides an additional layer of privacy protection for the prover, and in the case of LifeCODE.ai, it protects the data owner. The LifeCODE.ai blockchain then supports private transactions through Cambrian, a peer-to-peer encrypted message exchange that secures data while communicating directly with network participants.

Nevertheless, this kind of security is not absolutely safe. The core of blockchain technology is cryptography, and its focus is the hash function. Many hash functions have been designed and widely used, but they generally have a short life span. Algorithms that are considered safe are often not successfully used for long periods of time, and new and more secure algorithms are designed successively. In addition, some imperfections remain in the computer languages upon which the blockchain relies. There will be a lag in the integration with new technologies that will affect the blockchain technology system. Unfortunately, the Ethereum blockchain has reportedly been hacked multiple times, resulting in the loss of currency.

Results

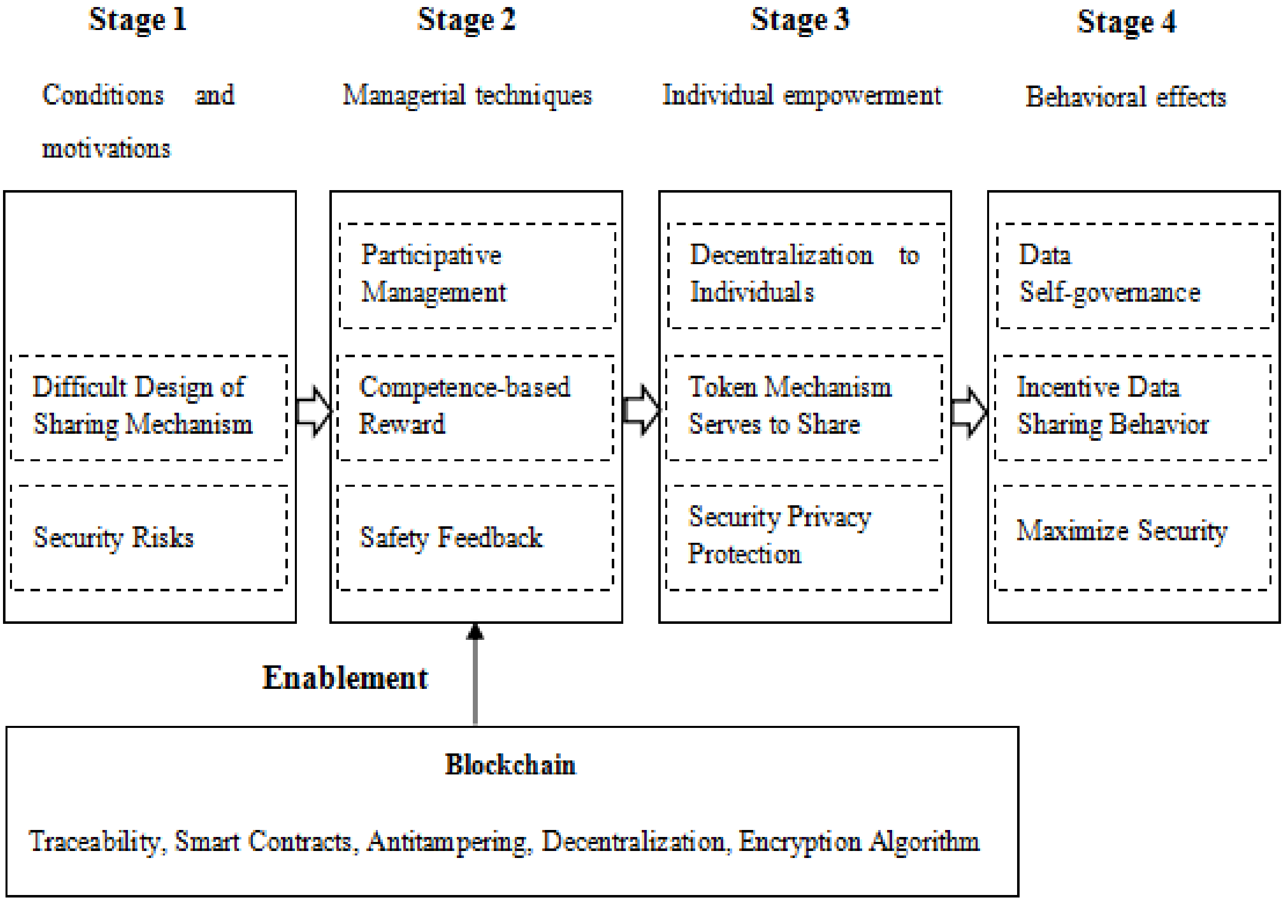
Blockchain as a Service

The blockchain-based genomics big data platform, LifeCODE.ai, is supposed to provide relatively safe and trustworthy data storage and management services to genomic stakeholders. As such, the platform can empower them to enhance precise medicine, personalized treatment, and new drug R&D. Specifically, it provides Blockchain as a Service, a new type of cloud service based on blockchain technology, for the genomic stakeholders including both institutions and individuals.

Even though LifeCODE.ai has been established less than a year, its tools have been successfully leveraged in several cases. For example, LifeCODE.ai provides blockchain data encryption and secure storage services to the National Center for Clinical Medical Research on Geriatric Diseases in China for storing and managing gene data of the patients with Parkinson and dyskinesia diseases. The research and analysis of this dataset can help to overcome the clinical research difficulties of Parkinson and dyskinesia diseases and promote the popularization and application of genetic testing technology in clinical diagnosis and treatment. However, LifeCODE.ai has not yet gathered enough users, and the limitations of the blockchain have not been well resolved. The future development of LifeCODE.ai still needs time verification.

To better understand the empowerment process of blockchain for individual's genomic data exchange behavior, we draw on the research framework of psychological empowerment process. This theoretical framework was proposed by Conger and Kanungo (1988) to explain the process of empowerment in 5 stages including the psychological state of empowering experience as well as its antecedent conditions and behavioral consequences [46]. This model has been widely adopted to understand how to change the psychology of subordinates through empowerment behavior, thereby promoting behavior effects [47,48]. For example, this model was adopted to investigate how information system empowers users' attitudes and promotes behavior effects [49]. On the basis of Conger and Kanungo's empowerment process model and the case analysis of LifeCODE.ai, we proposed a framework to illustrate the empowerment process of blockchain for individual's data exchange in a genomic big data platform. As there is no subordinate relationship in the big data platform, this study removed stage 3 of Conger and Kanungo's framework. As can be seen in Figure 4, blockchain empowerment process can be divided into 4 stages as mentioned below.

Figure 4. The empowerment process of blockchain for individual’s genomic data management.



Diagnosis of Conditions and Motivations

As mentioned above, 3 obstacles hinder the process of establishing genomic big data platforms including unclear data ownership, difficult design of sharing mechanism, and security risk.

Use of Managerial Techniques

Due to blockchain’s special features (ie, traceability, smart contracts, antitampering, decentralization, and encryption algorithm), individuals in the blockchain platform have the ability to participate in the data management process. They can also obtain token rewards from data transaction. The token rewards can be viewed as a kind of competence-based reward. Moreover, the application of blockchain technology can guarantee data security and privacy to a greater extent.

Empowerment of Individuals

Blockchain technology provides new ideas for the establishment of a genomic big data platform. The application of blockchain technology assigns data ownership to the individuals and delegates power to them. The token mechanism can serve the data-sharing process, and the security and privacy of the data are better protected.

Leading to Behavioral Consequences

Ultimately, the above aspects will result in changes in the behavior of the individuals in the platform. Personal ownership of data will facilitate the autonomy of personal data. The token mechanism will encourage more active sharing between individuals and institutions. Besides, reliable protection of security and privacy will encourage individuals to use this platform more actively.

Discussion

Conclusions and Implications

We conducted a case analysis to show how blockchain can empower the storage and management of personal genomics data from the perspectives of data ownership, the data-sharing mechanism design, and data security. We found that blockchain can allocate data ownership to individuals because of its traceability feature, so that individuals can control and monitor all the activities of their genomic data. This process leads to decentralization of data ownership and facilitates the autonomy of personal data. Meanwhile, the blockchain can use its smart contract and antitampering features to design a token-sharing mechanism promoting more active data sharing between individuals and organizations. Last but not least, the distributed storage and secret key mechanism, resulting from blockchain’s decentralization feature and cryptographic feature, respectively, provide greater guarantee for the security and privacy of the genetic data. This security feedback will increase confidence in the personal application platform.

In conclusion, the blockchain platform provides new possibilities for the management and security of genetic data and can help realize the psychological empowerment of individuals in the process, and consequently, the effects of data self-governance, incentive-sharing, and security improvement can be achieved. As concepts of data ownership vary among cultures and laws vary around the world, it is worth noting that the question regarding data ownership needs further exploration. The discussion in this paper is based on the Chinese background, so it is more reasonable to hand over the ownership of data to the data producer (which is individuals) without the specific guidance of the law. Undeniably, the blockchain is far from perfect. Regulatory issues, capacity and efficiency issues, and hash cryptographic security issues remain challenges in the current phase of blockchain.

Limitations and Directions for Future Research

This study has several limitations that call for additional research. First, we selected only 1 blockchain-based genomics big data platform in China (ie, LifeCODE.ai) as our research object. A future study should compare LifeCODE.ai with similar platforms to enhance generalizability. Second, because LifeCODE.ai has been in use for less than a year, successful cases are still limited. Its client DApp (Lai-Tribe) has only around 5000 registered users. It takes time to prove whether users and medical institutions will be willing to store and share their genomics and health data in this blockchain-based platform. Future studies are needed to investigate the attitudes of individual users and medical institutions toward blockchain-based data storage and management services provided by LifeCODE.ai and similar entities. Third, at the current stage of blockchain technology, some problems remain to be solved, such as capacity issues and regulatory issues. Therefore, the study of blockchain in gene data management and security assurance requires long-term follow-up and in-depth research with multiple cases and multiple angles.

Acknowledgments

This work was supported by the National Science Foundation of China (Grant Nos 71872112 and 71871162) and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2018).

Conflicts of Interest

None declared.

References

1. Sharp SA, Weedon MN, Hagopian WA, Oram RA. Clinical and research uses of genetic risk scores in type 1 diabetes. Curr Opin Genet Dev 2018 Jun;50:96-102 [FREE Full text] [doi: 10.1016/j.gde.2018.03.009] [Medline: 29702327]
2. Ogishima S. [Human genome data and drug development]. Gan To Kagaku Ryoho 2018 Apr;45(4):597-600. [Medline: 29650811]
3. Cai YD, Huang T. Accelerating precision medicine through genetic and genomic big data analysis. Biochim Biophys Acta Mol Basis Dis 2018 Jun;1864(6 Pt B):2215-2217 [FREE Full text] [doi: 10.1016/j.bbadis.2018.03.012] [Medline: 29548968]
4. Low SK, Zembutsu H, Nakamura Y. Breast cancer: the translation of big genomic data to cancer precision medicine. Cancer Sci 2018 Mar;109(3):497-506. [doi: 10.1111/cas.13463] [Medline: 29215763]
5. Interuniversity Institute of Bioinformatics in Brussels: IBsquare. 2016. BRiDGEIris: BRussels Big Data Platform for Sharing and Discovery in Clinical GEnomics URL: https://ibsquare.be/drupal/research/projects/bridgeiris-brussels-big-data-platform-sharing-and-discovery-clinical-genomics [accessed 2019-01-30]
6. Siwicki B. Healthcare IT News. 2018. New Genomics Analytics Platform From Databricks Aims to Speed Discovery of New Treatments URL: https://www.healthcareitnews.com/news/new-genomics-analytics-platform-databricks-aims-speed-discovery-new-treatments [accessed 2019-01-30]
7. Bhuvaneshwar K, Belouali A, Singh V, Johnson RM, Song L, Alaoui A, et al. G-DOC Plus - an integrative bioinformatics platform for precision medicine. BMC Bioinformatics 2016 Apr 30;17(1):193 [FREE Full text] [doi: 10.1186/s12859-016-1010-0] [Medline: 27130330]
8. Li JQ. Open Medical Big Data and Open Consent and Their Impact on Privacy. In: Proceedings of the International Congress on Big Data. 2017 Presented at: BigData Congress'17; June 25-30, 2017; Honolulu, HI, USA. [doi: 10.1109/BigDataCongress.2017.76]
9. Pizzolante R, Castiglione A, Carpentreri B, de Santis A, Palmeiri F, Castiglione A. On the protection of consumer genomic data in the internet of living things. Comput Secur 2018;74:384-400 [FREE Full text] [doi: 10.1016/j.cose.2017.06.003]
10. Tao JL, Chen SY. Research on Personal Privacy Protection in Medical Big Data. In: Proceedings of the 2nd International Conference on Artificial Intelligence and Engineering Applications. 2017 Presented at: AIEA'17; September 23-24, 2017; Guilin,China p. 1033-1039. [doi: 10.12783/dtcse/aiea2017/15044]
11. Kulynych J, Greely HT. Clinical genomics, big data, and electronic medical records: reconciling patient rights with research when privacy and science collide. J Law Biosci 2017 Apr;4(1):94-132 [FREE Full text] [doi: 10.1093/jlb/lsw061] [Medline: 28852559]

12. Balsari S, Fortenko A, Blaya JA, Gropper A, Jayaram M, Matthan R, et al. Reimagining health data exchange: an application programming interface-enabled roadmap for India. J Med Internet Res 2018 Jul 13;20(7):e10725 [FREE Full text] [doi: 10.2196/10725] [Medline: 30006325]
13. Trinidad SB, Fullerton SM, Bares JM, Jarvik GP, Larson EB, Burke W. Genomic research and wide data sharing: views of prospective participants. Genet Med 2010 Aug;12(8):486-495 [FREE Full text] [doi: 10.1097/GIM.0b013e3181e38f9e] [Medline: 20535021]
14. WARC. 2017. Consumers Need Benefits to Share Data URL: https://www.warc.com/NewsAndOpinion/News/Consumers_need_benefits_to_share_data/38117, [accessed 2019-02-01]
15. Lin P. Forbes Magazine. 2017. Blockchain: The Missing Link Between Genomics and Privacy? URL: https://www.forbes.com/sites/patricklin/2017/05/08/blockchain-the-missing-link-between-genomics-and-privacy/#5464c33e4b77, [accessed 2019-01-14]
16. Chen A. The Verge. 2018. Why a DNA Data Breach is Much Worse Than a Credit Card Leak URL: https://www.theverge.com/2018/6/6/17435166/myheritage-dna-breach-genetic-privacy-bioethics, [accessed 2019-01-31]
17. Tiller J, Otlowski M, Lacaze P. Should Australia ban the use of genetic test results in life insurance? Front Public Health 2017;5:330 [FREE Full text] [doi: 10.3389/fpubh.2017.00330] [Medline: 29322039]
18. Zhang P, White J, Schmidt DC, Lenz G, Rosenbloom ST. FHIRChain: applying blockchain to securely and scalably share clinical data. Comput Struct Biotechnol J 2018;16:267-278 [FREE Full text] [doi: 10.1016/j.csbj.2018.07.004] [Medline: 30108685]
19. Singh S, Singh N. Blockchain: Future of Financial and Cyber Security. In: Proceedings of the 2nd International Conference on Contemporary Computing and Informatics. 2016 Dec Presented at: IC3I'16; December 14-17, 2016; Noida, India p. 463-467. [doi: 10.1109/IC3I.2016.7918009]
20. Holbl M, Kompara M, Kamlsalic A, Zlatolas LN. A systematic review of the use of blockchain in healthcare. Symmetry-Basel 2018;10:470 [FREE Full text] [doi: 10.3390/sym10100470]
21. Chris G, Alos D. Academia.edu. 2016. Increasing Supply Chain Assurance via the Blockchain URL: https://www.academia.edu/19623721/Increasing_Supply_Chain_Assurance_via_the_Blockchain
22. Paul G, Sarkar P, Mukherjee S. Towards a More Democratic Mining in Bitcoins. In: Proceedings of the International Conference on Information Systems Security. 2014 Presented at: ICISS'14; December 16-17, 2014; Hong Kong p. 185-203. [doi: 10.1007/978-3-319-13841-1_11]
23. Wang L, Liu Y. Exploring Miner Evolution in Bitcoin Network. In: Proceedings of the International Conference on Passive and Active Network Measurement. 2015 Presented at: PAM'15; March 19-20, 2015; New York, NY, USA p. 290-302. [doi: 10.1007/978-3-319-15509-8_22]
24. Zyskind G, Nathan O, Pentland AS. Decentralizing Privacy: Using Blockchain to Protect Personal Data. In: Proceedings of Security and Privacy Workshops. 2015 Presented at: SPW'15; May 21-22, 2015; San Jose, CA, USA p. 180-184. [doi: 10.1109/SPW.2015.27]
25. Kosba A, Miller A, Shi E, Wen Z, Papamanthou C. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts. In: Proceedings of the Symposium on Security and Privacy. 2016 Presented at: SP'16; May 22-26, 2016; San Jose, CA, USA p. 839-858. [doi: 10.1109/SP.2016.55]
26. Meiklejohn S, Orlandi C. Privacy-Enhancing Overlays in Bitcoin. In: Proceedings of the International Conference on Financial Cryptography and Data Security. 2015 Presented at: FC'15; January 26-30, 2015; San Juan, Puerto Rico p. 127-141. [doi: 10.1007/978-3-662-48051-9_10]
27. Kim HM, Laskowski M. Toward an ontology-driven blockchain design for supply-chain provenance. Intell Syst Account Financ Manag 2018;25(1):18-27 [FREE Full text] [doi: 10.1002/isaf.1424]
28. Lim SY, Fotsing PT, Almasri A, Musa O, Kiah ML, Ang TF, et al. Blockchain technology the identity management and authentication service disruptor: a survey. Int J Adv Sci Eng Inf Technol 2018;8:1735-1745. [doi: 10.18517/ijaseit.8.4-2.6838]
29. Christidis K, Devetsikiotis M. Blockchains and smart contracts for the internet of things. IEEE Access 2016;4:2292-2303. [doi: 10.1109/ACCESS.2016.2566339]
30. Li Y, Huang JQ, Shengzhi QZ, Ruijin W. Big Data Model of Security Sharing Based on Blockchain. In: Proceedings of the 3rd International Conference on Big Data Computing and Communications. 2017 Presented at: BIGCOM'17; August 10-11, 2017; Chengdu, China p. 117-121. [doi: 10.1109/BIGCOM.2017.31]
31. Shabani M. Blockchain-based platforms for genomic data sharing: a de-centralized approach in response to the governance problems? J Am Med Inform Assoc 2019 Jan 1;26(1):76-80. [doi: 10.1093/jamia/ocy149] [Medline: 30496430]
32. Xia Q, Sifah EB, Asamoah KO, Gao J, Du X, Guizani M. MeDShare: trust-less medical data sharing among cloud service providers via blockchain. IEEE Access 2017;5:14757-14767. [doi: 10.1109/ACCESS.2017.2730843]
33. NKCC White Paper. URL: http://www.nk2627.com/white_paper_en.html#chap5, [accessed 2019-01-28]
34. Kuo TT, Kim HE, Ohno-Machado L. Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 2017 Nov 01;24(6):1211-1220 [FREE Full text] [doi: 10.1093/jamia/ocx068] [Medline: 29016974]
35. Sun H, Wang XY, Wang XG. Application of blockchain technology in online education. Int J Emerg Technol Learn 2018 Oct 26;13(10):252-259. [doi: 10.3991/ijet.v13i10.9455]

36. Yue X, Wang HJ, Jin DW, Li M, Jiang W. Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. J Med Syst 2016 Oct;40(10):218. [doi: 10.1007/s10916-016-0574-6] [Medline: 27565509]
37. Zhang P, White J, Schmidt DC, Lenz G. arXiv. 2017. Applying Software Patterns to Address Interoperability in Blockchain-Based Healthcare Apps URL: https://arxiv.org/abs/1706.03700
38. Ji YX, Zhang JW, Ma JF, Yang C, Yao X. BMPLS: blockchain-based multi-level privacy-preserving location sharing scheme for telecare medical information systems. J Med Syst 2018 Jun 30;42(8):147. [doi: 10.1007/s10916-018-0998-2] [Medline: 29961160]
39. Gummesson E. Qualitative Methods in Management Research. Thousand Oaks, California: Sage Publications; 2000.
40. Darke P, Shanks G, Broadbent M. Successfully completing case study research: combining rigour, relevance and pragmatism. Inf Syst 1998;8(4):273-289 [FREE Full text] [doi: 10.1046/j.1365-2575.1998.00040.x]
41. Barbieri S, Vettore G, Pietrantonio V, Snenghi R, Tredese A, Bergamini M, et al. Pedestrian inattention blindness while playing Pokémon Go as an emerging health-risk behavior: a case report. J Med Internet Res 2017 Apr 1;19(4):e86 [FREE Full text] [doi: 10.2196/jmir.6596] [Medline: 28365563]
42. Barnett S, Jones SC, Caton T, Iverson D, Bennett S, Robinson L. Implementing a virtual community of practice for family physician training: a mixed-methods case study. J Med Internet Res 2014 Mar 12;16(3):e83 [FREE Full text] [doi: 10.2196/jmir.3083] [Medline: 24622292]
43. Shaw J, Shaw S, Wherton J, Hughes G, Greenhalgh T. Studying scale-up and spread as social practice: theoretical introduction and empirical case study. J Med Internet Res 2017 Jul 7;19(7):e244 [FREE Full text] [doi: 10.2196/jmir.7482] [Medline: 28687532]
44. Ethereum Stack Exchange. What is a DApp? URL: https://ethereum.stackexchange.com/questions/383/what-is-a-dapp, [accessed 2019-01-31]
45. LifeCODE.ai. URL: https://www.lifecode.org.cn/databank/app/#/ [accessed 2019-08-07]
46. Conger JA, Kanungo RN. The empowerment process: integrating theory and practice. Acad Manag Rev 1988;13(3):471-482 [FREE Full text] [doi: 10.2307/258093]
47. Kirkman BL, Rosen B, Tesluk PE, Gibson CB. The impact of team empowerment on virtual team performance: the moderating role of face-to-face interaction. Acad Manag J 2004;47(2):175-192 [FREE Full text] [doi: 10.2307/20159571]
48. Zhang X, Bartol KM. Linking empowering leadership and employee creativity: the influence of psychological empowerment, intrinsic motivation, and creative process engagement. Acad Manag J 2010;53(1):107-128 [FREE Full text] [doi: 10.5465/amj.2010.48037118]
49. Bonias D, Bartram T, Leggat SG, Stanton P. Does psychological empowerment mediate the relationship between high performance work systems and patient care quality in hospitals? Asia Pac J Hum Resour 2010;48(3):319-337 [FREE Full text] [doi: 10.1177/1038411110381667]

Abbreviations

API: application programming interface
DApp: decentralized app
GOR: genomic ordered relational
HDFS: Hadoop Distributed File System
NFS: network file system
OSS: object storage service
R&D: research and development

Edited by P Zhang, K Clauson; submitted 01.02.19; peer-reviewed by P Lynch, M Hölbl, A Kamisalic, J Liu; comments to author 27.04.19; revised version received 26.06.19; accepted 19.07.19; published 10.09.19.

Please cite as:
Jin XL, Zhang M, Zhou Z, Yu X
Application of a Blockchain Platform to Manage and Secure Personal Genomic Data: A Case Study of LifeCODE.ai in China
J Med Internet Res 2019;21(9):e13587
URL: http://www.jmir.org/2019/9/e13587/
doi:10.2196/13587
PMID:31507268

Original Paper

Benefits of Blockchain Initiatives for Value-Based Care: Proposed Framework

Rongen Zhang[1]; Amrita George[2], PhD; Jongwoo Kim[3], PhD; Veneetia Johnson[1], DBA; Balasubramaniam Ramesh[1], PhD

[1]Georgia State University, Atlanta, GA, United States
[2]Marquette University, Milwaukee, WI, United States
[3]University of Massachusetts Boston, Boston, MA, United States

Corresponding Author:
Balasubramaniam Ramesh, PhD
Georgia State University
35 Broad Street
Atlanta, GA, 30303
United States
Phone: 1 404 413 7372
Email: bramesh@gsu.edu

Abstract

Background: The potential of blockchain technology to achieve strategic goals, such as value-based care, is increasingly being recognized by both researchers and practitioners. However, current research and practices lack comprehensive approaches for evaluating the benefits of blockchain applications.

Objective: The goal of this study was to develop a framework for holistically assessing the performance of blockchain initiatives in providing value-based care by extending the existing balanced scorecard (BSC) evaluation framework.

Methods: Based on a review of the literature on value-based health care, blockchain technology, and methods for evaluating initiatives in disruptive technologies, we propose an extended BSC method for holistically evaluating blockchain applications in the provision of value-based health care. The proposed method extends the BSC framework, which has been extensively used to measure both financial and nonfinancial performance of organizations. The usefulness of our proposed framework is further demonstrated via a case study.

Results: We describe the extended BSC framework, which includes five perspectives (both financial and nonfinancial) from which to assess the appropriateness and performance of blockchain initiatives in the health care domain.

Conclusions: The proposed framework moves us toward a holistic evaluation of both the financial and nonfinancial benefits of blockchain initiatives in the context of value-based care and its provision.

(J Med Internet Res 2019;21(9):e13595) doi:10.2196/13595

KEYWORDS

blockchain; balanced scorecard; evaluation; value-based care

Introduction

Background

The health care sector has recently been focused on two related challenges: the transition to value-based care and the use of innovative technologies (such as blockchain technology) to facilitate the delivery of health care. The transition to value-based care, which aims to improve the value of care while providing it at a lower cost, places new demands on health care information systems (IS) [1] that current health information technology infrastructure is not designed to support [1]. Adler-Milstein et al [1] identified three major stakeholder groups that must be supported in achieving value-based care: patients, providers, and researchers. Disruptive technologies such as blockchain offer the potential to support these currently inadequately supported stakeholder groups with Health Information Technology (IT) infrastructure.

Blockchain technology, widely celebrated as a technological revolution, is creating unprecedented hype and optimism [2]. Blockchain is a distributed database that maintains a continuously growing list of data records that are secured from tampering and revision [3,4]. A global survey documents the widespread application of blockchain in domains such as health care, manufacturing, legal, government, not for profit, retail, real estate, tourism, and media [5]. The potential of this technology to aid organizations in achieving strategic goals like value-based care is increasingly being recognized by health care providers and other stakeholders (eg, payers, shareholders, accreditation agencies) [6]. However, Iansiti and Lakhani [7] note that practitioners are uncertain about the impact that disruptive technologies such as blockchain might have on organizational performance. Current research and practice lack comprehensive approaches to evaluating the benefits of blockchain and developing appropriate use cases of blockchain applications for value-based care [8].

As IT is increasingly becoming a strategic necessity for improving services and reducing medical errors [9], comprehensive approaches to evaluating the appropriateness and value of disruptive technologies such as blockchain are needed. An evaluation approach should facilitate the assessment of both technical and nontechnical (eg, legal, data ownership and privacy, security) implications. To address this need, we assessed two sets of existing evaluation frameworks: technology evaluation methods (the Zachman framework, human-computer interaction [HCI] guidelines, and the technology-centric framework) and comprehensive evaluation methods (total quality management [TQM], the European foundation quality management excellence model (EFQMEM), the performance pyramid, and the performance prism). Based on this assessment, we identified deficiencies in the existing evaluation methods and subsequently developed an approach that extends the balanced scorecard (BSC) framework that addresses these deficiencies.

The BSC, developed by Robert Kaplan and David Norton nearly two decades ago [10], provides organizations with a structured approach to assessing both the financial and nonfinancial dimensions of organizational initiatives and processes in terms of strategic outcomes. Beyond the purely accounting-based measures traditionally used, the BSC is balanced in that it provides a comprehensive view of organizational performance. It translates high-level organizational vision and strategy into a holistic set of performance and action measures [11]. The BSC is a practical method that is applicable within the health care service sector and health care organizations, and it has previously been used to assess clinical outcomes, for example [12]. However, it has not yet been used to evaluate disruptive innovations, such as blockchain, that can improve patient care and reduce costs but that have regulatory, financial, and operational implications.

A myriad of seemingly promising blockchain projects are being implemented in the health care domain, often without careful consideration of the applicability of the technology [13]. Moreover, questions still linger for early adopters of this technology: “How does an organization holistically assess the performance of blockchain technology in the health care domain?” and “Does the introduction of blockchain technology align with the strategic priorities of a health care organization?”. Answering these questions is critical for health care organizations to achieve the health care IT mission identified by the US federal government, namely: "Improve the health and well-being of individuals and communities through the use of technology and health information that is accessible when and where it matters most" [14].

We sought to answer the above questions through our assessment of existing evaluation frameworks and the development of a new framework that can guide the comprehensive evaluation of the value of blockchain initiatives that seek to enable the delivery of value-based care.

In the sections below, we first discuss the relevant literature on value-based health care and blockchain technology. We then assess existing evaluation frameworks and present our framework, which extends the BSC by addressing some of its limitations. Further, we customize the framework to the context of blockchain applications in health care settings. We then present an illustrative case study on the application of the framework in a pharmaceutical supply chain organization. Finally, we discuss the implications of our framework for both researchers and practitioners.

Information Technology Support for Transitioning to Value-Based Health Care

Health care value, defined as health outcomes (including quality of care achieved per dollar spent), has become a cornerstone of the strategy to restructure the US health care system [15-17]. One of the proposed frameworks for improving health care value is the value-based care model [18]. Value-based care attempts to advance the triple aim of providing better care for individuals, improving population health management strategies, and reducing health care costs. Value-based care models center on patient outcomes and how well health care providers can improve quality of care using measures such as reduced hospital readmissions, improved timeliness and safety of care, more equitable care, shared decision-making, and improved preventative care [17]. This model ties payments for care delivery to the quality of care provided, and rewards providers for both efficiency and effectiveness [19].

Unlike more traditional approaches, value-based care is driven by data because providers must report to payers on specific metrics and demonstrate improvement. Providers are required to use IT systems to track and report metrics such as hospital readmissions, adverse events, population health, and patient engagement. Further, providers are incentivized to use evidence-based medicine, engage patients, upgrade health care IT, use data analytics, and receive payments electronically. When patients receive more coordinated, appropriate, and effective care, providers are rewarded. To achieve these goals, health care organizations need a digital infrastructure that facilitates the provision of comprehensive, affordable, accessible, effective, and error-free care.

While significant progress has been made in digitizing the US health care system, today’s health IT infrastructure largely remains a collection of systems that were not designed to support the transition to value-based care [1]. In fact, prior literature has identified a health IT chasm, which refers to the gaps between the current health IT ecosystem (see Multimedia Appendix 1) and the system that is needed for value-based care [1,20-42].

In fact, a recent study identified several gaps from the perspectives of three stakeholder groups. From the patient perspective, patients are unable to access electronic medical records from most providers, and most care providers do not provide functionalities for patients to submit patient-generated data. Only a small percentage of patients receive clinical trial information from their primary physician, and an even smaller percentage participate in biobanks [1]. From the provider perspective, due to the lack of standardized application interfaces providers have difficulty accessing external data, which hinders the advanced analytics on which personalized assistance is based [43]. In addition, manual credentialing (typically takes more than 120 days) and administration of contracts is complicated and inefficient. Further, pharmaceutical providers find it challenging to ensure the authenticity of pharmacy products due to a lack of transparency in current supply chain systems. Finally, from the researchers' perspective, it is difficult for them to track investigational products to ensure data authenticity, and payments to investigators are delayed due to manual processing [33]. The health IT environment is immature, provides few safeguards for safety and effectiveness, and provides very limited integration of applications used in clinical care or research.

Prior literature has also identified specific goals (eg, improving patients' access to clinical data, improving patient's ability to submit and access data via mobile health technology, more readily engaging patients in clinical research) for addressing the needs of each of these stakeholder groups [1]. Blockchain technology may help achieve these goals.

Blockchain for Enabling Value-Based Care

Blockchain consists of blocks that hold batches of individual transactions. Each block contains a timestamp and a link to a previous block [3,4]. The most salient benefit of blockchain is decentralization and the elimination of a trusted centralized third party in distributed applications. Thus, multiple parties can conduct transactions in a distributed environment without the need for a centralized authority, thereby avoiding a single point of both trust and failure. The absence of a centralized processing entity may reduce time and costs. A consensus mechanism is used to reconcile any discrepancies that may arise between participants in a blockchain network.

Iansiti and Lakhani [7] summarized five basic principles underlying blockchain technology: a distributed database, peer-to-peer transmission, transparency with pseudonymity, irreversibility of records, and computational logic. These unique characteristics of blockchain technology enable the development of solutions that reduce uncertainty and ambiguity and enhance security of stored transactional information by providing full transparency and a single truth for all network participants [44]. Although blockchain technology enjoys the benefits of decentralization, it often comes at the cost of scalability. Blockchains are typically incapable of processing large numbers of transactions in a timely manner [1]. The trustless peer-to-peer network infrastructure, which requires information to be propagated to and validated at each node, is the root of this problem. Several solutions (eg, off-chain transactions, sharding, and a provably neutral cloud) have been proposed to address this issue. For example, Leung et al proposed a design that minimizes storage, bootstrapping costs, and bandwidth costs of joining a network by 90% [45]. Such advances are essential for blockchain to realize its disruptive potential [46]. However, effective management of personal health records using blockchain technology still requires improvements such as reduced data size, strengthened personal information protection, and reduced operational costs [47].

Despite its technological infancy, experimental adoption and customization of blockchain technology appears to be fully underway in the health care domain [8]. One of the most impactful health care applications is expected to be the management of electronic health records (EHRs). The decentralization, immutability, traceability, security, and privacy of blockchain make it well suited for the storing, managing, and sharing of patient-centric data among stakeholders [48-50]. Aligning with the requirements of the European General Data Protection Regulation (GDPR), blockchain can be used to build health care platforms that empower patients to control how their data are used and ensure that sensitive personal data are not revealed without the patients' consent [2,22,51]. Guardtime [2,51], MedRec [23], the Gem Health Network [44], Patientory, and IBM's Watson [21] are some of the key projects in this ecosystem.

Another salient application domain of blockchain is supply chain management in the pharmaceutical industry. Because of the immutability and traceability of blockchain, any modification of a prescription by any party in the supply chain can be detected, which, in turn, can help address the severe problem of counterfeit medications [2,44,49]. In addition, in biomedical research and education, blockchain could facilitate the elimination of falsification of data or the exclusion of undesirable results from clinical trials [31]. Benchoufi [38] and Nugent et al [37] illustrated the ability to trace patients' consent and provide data transparency in clinical trials. Moreover, insurance claim processing is a promising area for blockchain applications because of its transparency, decentralization, immutability, and auditability; a few prototype implementations, such as MIStore [52] and Politdok's initiative partnered with Intel [53], have been reported [44]. Other promising areas include remote patient monitoring [24] and precision medicine [54].

Blockchain technology has the potential to address some of the gaps in the current health IT ecosystem, thereby supporting the three important stakeholder groups involved in value-based care [1]. Multimedia Appendix 1 identifies these gaps and highlights what blockchain can do to address these gaps. Based on a careful study of the needs of the three stakeholder groups, we further outline in the appendix how specific characteristics of blockchain technology may help meet these needs. We also list some proof-of-concept systems that provide some of the desired functionalities.

Methods

Overview

While blockchain offers the potential to address issues (eg, interoperability, difficulty in providing optimal personalized care due to lack of comprehensive medical records, and maintaining integrity of records) that are critical for effective value-based care [55], there is limited research comprehensively evaluating the financial and nonfinancial benefits of blockchain solutions in health care [56]. A review of the literature on value-based health care strongly suggests the need for a framework to holistically evaluate the impacts of technologies such as blockchain. Existing evaluation mechanisms (such as the Level of Information System interoperability reference model [56]) have focused on the operational aspects of blockchain. Motivated by the need for a framework to guide the strategic evaluation of blockchain applications within a health care organization, we extend the BSC approach, which is an already well-established performance evaluation system. Specifically, our approach integrates financial and nonfinancial perspectives (ie, internal processes, learning and growth, external perspectives, and customer perspectives), which are parts of the original BSC, with an external perspective that incorporates the viewpoints of external stakeholders and regulators, especially because of the significant role these parties play in health care delivery. In the following section, we illustrate the use of our framework with a blockchain application for managing a pharmaceutical supply chain.

Performance Evaluation of Health Care Blockchain Implementations and the Balanced Scorecard

Traditional performance measurement systems have either focused purely on financial factors, ignoring the value of nonfinancial factors [12], or have focused solely on the effectiveness of the technical system without considering the external or financial implications. Health care organizations have been using economic evaluations for health care decision-making for several decades. During this period, increased pressure on health care budgets has necessitated the consideration of cost-effectiveness in addition to clinical effectiveness. Economic evaluation approaches have also been applied to other health care–related decision-making in terms of funding, reimbursement, and new technologies [57,58]. Even comprehensive evaluation approaches that include cost-consequences analysis, cost-minimization analysis, cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis [59] are focused on financial factors and give limited consideration to nonfinancial aspects of evaluation targets. For example, Zachman's framework [60] evaluates business-IT alignment in detail but lacks a holistic governance framework. Similarly, the human-computer interaction [61] and technology-centric frameworks [62] provide insights into developing intuitive and interactive IS, but they do not focus on assessing the impact of these systems from external and financial perspectives. Additionally, the interrelationships between the various functional areas in an organization are overlooked in these frameworks. For example, a blockchain implementation in one functional area, such as improving patients' access to their own medical records, may have major impacts in other areas, such as customer service management, internal processes for quality assurance, security checks, or external partnerships (with, say, insurance companies or pharmacies). Finally, the knowledge that results from the long-term growth of organizations or the ability to deal with future threats also needs to be factored into the performance evaluation [12]. The BSC has dual functions as a performance framework and a management methodology, and thus can tackle the shortcomings of traditional performance measurement systems. These shortcomings include the lack of consideration of nonfinancial factors and the lack of strategic focus. Our evaluation suggests that the BSC addresses both shortcomings and is well suited for the evaluation of disruptive technologies, especially in the dynamic environment in which health care organizations operate.

Our comparison of the various performance measurement systems, as presented in Multimedia Appendix 2, suggests that BSC is an appropriate approach for evaluating blockchain initiatives in achieving value-based care for the following reasons. We compared BSC with two sets of existing methods, namely, technology evaluation methods (the Zachman framework, HCI, and the technology-centric framework) and comprehensive performance evaluation methods (TQM, the European foundation quality management excellence model, the performance pyramid, and the performance prism) [63-65]. Technology evaluation methods typically do not provide a holistic view (such as the consideration of external or customer perspectives) and therefore are not appropriate in our setting. Among the comprehensive evaluation methods, TQM's narrow focus on internal process is inadequate, and the European foundation quality management excellence model, designed to improve TQM, lacks a strategic focus. Although both the BSC and the performance pyramid use strategic mapping to link strategy to operational metrics, prior research suggests that the performance pyramid is less effective and harder to understand than the BSC [64]. Moreover, although the performance prism considers stakeholders' perspectives, it does not provide adequate guidelines and neglects to show how the proposed measures can be operationalized [65]. Thus, our comparison of the various technical and comprehensive performance evaluation methods suggests that, among them, the BSC is the most suited to evaluate the performance of disruptive technologies (such as blockchain) in value-based care initiatives.

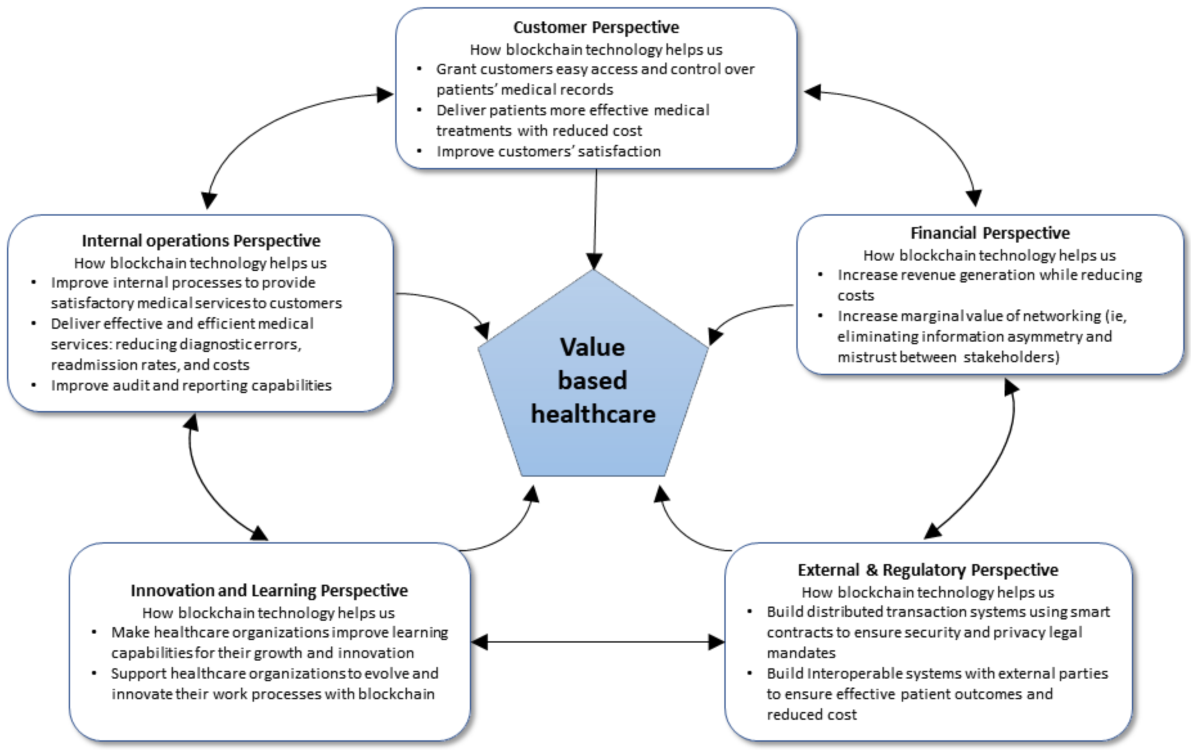
Organizations in multiple domains, including health care, have adopted the BSC [66,67]. In increasingly dynamic business environments, traditional performance evaluation approaches may not work well due to the uncertainty involved in ascertaining both the costs and benefits of new technologies, such as blockchain. However, both theoretical research and practitioner articles support the use of the BSC for evaluating IT initiatives in such contexts. For example, Gartner [68] notes that performance measurement solutions deployed within an organization should include a spectrum of leading measures rather than focusing on lagging financial indicators. To provide a holistic assessment, Gartner [29] recommends using the BSC to measure return on investment (ROI) and the business value of IT services because it enables the consideration of both financial and nonfinancial perspectives and helps develop relevant metrics [68]. Researchers also recognize the BSC framework as a holistic approach that provides managers with a structure to develop metrics that reflect performance from various perspectives [69], hence our selection of the BSC as the basis for the development of our approach to evaluating blockchain applications.

The BSC measures the performance of organizations from the following four linked and balanced perspectives:

1. Financial: How do we increase value for our shareholders (or providers of financial resources)?
2. Customer: How well do we satisfy our customers’ needs?
3. Internal: How well do we perform key internal operational processes? To satisfy our customers, in what processes must we excel?
4. Learning and growth: Are we able to sustain innovation, change, and continual improvement? Do we have the basic infrastructure in place to improve, create, value, and achieve our mission?

Some limitations of the traditional BSC have received attention in the literature [70,71]. One major concern is that the environment external to the organizations, including key groups of stakeholders, is not represented in the framework. For example, Mohobbot [72] points out that the BSC is unable to answer questions concerning the impact of external competitors. Moreover, the BSC does not consider the extended value chain, in which supplier and employee contributions are very significant [73]. This issue is exacerbated in the health care domain due to the complex interactions among the wide variety of organizations and stakeholders that are part of the ecosystem. For example, Norreklit [30] identifies crucial stakeholders like public authorities and suppliers, but other external stakeholders may include insurers, physicians, hospitals, clinics, laboratories, clinical research organizations, supply chain logistics stakeholders (such as pharmaceutical manufacturers, distributors, and retailers), government and regulatory agencies, and charities. To account for the impact of external stakeholders, we extended the BSC with an additional perspective, namely the external and regulatory perspective (see Figure 1). This perspective seeks to answer the following question: "How well does the organization improve value creation through external partnerships while ensuring regulatory compliance?"

Figure 1. Proposed framework for evaluating blockchain initiatives for value-based care.



By integrating financial measures with other crucial performance indicators concerning patients, organizational learning, growth and innovation, internal processes, and external perspectives, this extended BSC framework offers health care organizations a comprehensive view of the performance of blockchain applications.

Results

Summary

In this study we adopted a for-profit health care organization's view, as a majority of current blockchain implementations are in for-profit organizations.

Financial Perspective

From a value-based perspective, one of the key questions health care organizations should ask is: "How do health care organizations use blockchain applications to generate more profits at lower cost?" Typically, the focus of the financial perspective in the BSC has been on traditional financial metrics such as ROI and net income. In the context of value-based health care, patient-centric metrics such as gross revenue, adjusted cost per discharge, in-patient or out-patient revenue mix, contract allowances, discounts as a percentage of operating patient revenue [74], patient-payer mix, Medicare or Medicaid mix, average length of stay, and occupancy rate all deserve consideration.

The auditability and traceability features of blockchain enable more secure and efficient revenue management. As it does not require an intermediary, blockchain can support health care financing tasks, such as automatic claims processing using smart contracts [48,75], preauthorization of payments [36], and alternative payment models [76]. A distributed ledger makes claims processing and payment transactions more efficient and cost-effective. Replacing redundant health care intermediaries (namely, organizations that operate between stakeholders and institutions but that add little value to the health care value chain [54]) with transparent blockchain technology could facilitate processes like real-time claims adjudication [75]. With the data provenance benefits offered by blockchain, providers and patients could have enhanced accessibility to patient data. Blockchain technology can also help eliminate information asymmetry and mistrust between stakeholders in the health care ecosystem. With the innate immutability, transparency, and traceability provided by blockchain technology, medical products can be traced from manufacturer to patient, thereby reducing medication and medical equipment fraud. However, in the short-term, the adoption of blockchain technology will likely involve significant investments in application development, and their integration with legacy systems might initially undermine the financial benefit to shareholders.

Customer (Patient) Perspective

From a value-based care perspective, one of the key questions health care organizations should ask is: "How can we improve our service to customers and satisfy customer needs via blockchain applications?" Improving the performance of health care information systems that support the provision of effective and efficient care to patients is critical for achieving this goal. The patient-centric care paradigm requires the sharing of patients' EHRs, which raises issues such as privacy, confidentiality, integrity, availability, and security [77]. As a valuable personal asset, health care data should be owned and controlled by customers (patients) easily and securely without violating their privacy [41]. With blockchain-supported applications such as FHIRChain and Blockchain-Based Multi-level Privacy-Preserving Location Sharing Scheme (BMPLS), which simplify data authentication and authorization, patients can control access to their medical data easily and quickly. According to a seminal paper on IS success [78], user satisfaction is affected by information quality, system quality, and service quality. Blockchain enables health care stakeholders to access complete, relevant, and secure data on patients, thereby improving information quality. Health care organizations can overcome common challenges, such as data segregation, and achieve better integration of patients' medical data. Blockchain supports data immutability and auditability, thereby improving service quality (eg, reliability, responsiveness, and rapport) of medical IS [79], and as a result health care organizations can enhance their medical service quality and thereby patient satisfaction. Blockchain can help health care organizations easily integrate various elements of clinical data, which can enable medical professionals to make accurate diagnoses at low cost.

Internal Perspective

From a value-based care perspective, one of the key questions that health care organizations should ask is: "What internal processes can blockchain improve to satisfy our customers and the population in general?" Effective internal business processes are critical for providing products and services that satisfy health care organizations' customers' needs in a fiscally responsible manner. These effective processes can be reliable indicators of future financial and operational success [12]. With blockchain applications, health care organizations can build time-stamped, tamper-proof, immutable ledger systems that will improve organizations' auditing and reporting capabilities. These capabilities are crucial for identifying failures in internal processes and remedying those failures. Some benefits that accrue with improvements to internal processes include reduced length of patient stay, accuracy of services provided (both primary and ancillary), optimal surgical capacity utilization, and timeliness of services [12].

A variety of internal processes are candidates for improvements using blockchain technology. Using smart contracts, organizations can encode internal logic (eg, validating identity and tracking the participation of various stakeholders. such as patients and health providers), which will enhance service quality. Service quality can be reflected in measures such as reductions in diagnostic errors, readmission rates, and data security incidents, all of which lower costs. Value for customers can also be improved by instituting newer internal processes, such as Hitech service (eg, digitization of wellness check in Mount Sinai's Lab 100 [80]). Access to longitudinal medical charts using blockchain applications (such as those implemented in FHIRChain) can help health care organizations achieve optimal results with Hitech services, thereby enabling effective long-term care for chronic illnesses (eg, diabetes). Further, such charts can be useful in designing effective population outreach programs. Finally, using peer-to-peer network-enabled blockchain applications (eg, BMPLS), health care organizations can leverage newer mechanisms of health care delivery, such as telecare, to increase their reach, thereby improving health equity while providing care at a reduced cost.

External Perspective

From a value-based care perspective, one of the key questions health care organizations should ask is: "How can we leverage external partnerships to create value while ensuring regulatory compliance, thus satisfying our customers and the general population?" Creating effective partnerships with external stakeholders (eg, payers, accreditation bodies) while remaining compliant with regulations is critical for value creation. These partnerships enable health care organizations to supply products and services that satisfy customer needs in a fiscally and legally responsible manner.

Some multi-level, privacy-preserving. location-sharing blockchain applications (eg, BMPLS) enable interoperability with external systems, thereby enabling access to multi-dimensional medical charts from various stakeholders that can improve long-term medical care at a low cost. Through external partnerships, these health care organizations can seek to create value by taking a proactive role in providing care to their customers (say, by tracking customers' lifestyle and suggesting changes). Naturally, such partnerships can enable future financial and operational success through service innovation, which can help build deeper long-term relationships with customers. Additionally, having access to multi-dimensional population health data will enable health care providers to design outreach services that benefit the community as a whole. Blockchain solutions may also include smart contracts that help meet security and privacy mandates. Further, through the standardization of smart contracts at both the provider's and the external partner's end, interoperability of medical systems for value creation can be achieved.

Learning and Innovation Perspective

From a value-based care perspective, one of the key questions health care organizations should ask is: "How can we use blockchain applications to improve the learning capabilities that lead to growth and innovation?" Blockchain applications can help health care organizations reassess their resources, from employee capabilities to health care delivery processes, and align them to the organization's strategy.

Blockchain enables health care stakeholders to learn and to improve their services, thereby enhancing their competitiveness and sustainability. The systems interoperability enabled by blockchain technology can help health care professionals learn about opportunities to innovate their services. Blockchain technology also supports organizations in reassessing existing processes and resources and identifying opportunities for improvement. For example, auditability and traceability improved by blockchain can help streamline insurance claim processes and make them easier to manage. Blockchain can also significantly reduce administration costs and potentially eliminate some intermediaries that were previously needed for data integration. Aggregated health care data can help health care organizations reconfigure their procedures and innovate medical services for patients. With enhanced traceability and transparency supported by blockchain, organizations can learn how to optimize the health care supply chain.

Interrelationships Among Perspectives

The BSC does not explicitly consider the interactions and trade-offs between perspectives. In dynamic environments, correctly identifying and addressing trade-offs between perspectives can help organizations accurately evaluate the target system and develop effective incentives to improve overall organizational performance. Focusing on the financial perspective alone may motivate organizations to reduce nonfinancial investments that could produce long-term benefits. In particular, if a nonfinancial perspective has no contemporaneously congruent relationship with financial perspective, managers may reduce investments that improve performance in other areas for short-term benefits.

Our approach suggests that in addition to evaluating value-based care with respect to each perspective, health care organizations need to examine the interrelationship among the five perspectives. For example, efforts to improve the efficiency of internal processes (eg, improving quality process within a unit) with blockchain applications can help health care organizations enhance their learning capabilities (eg, creating quality management processes at the organizational level).

While developing relevant key metrics for each perspective (see Multimedia Appendix 3) is crucial for the effective use of the BSC, it is also important to carefully examine the relationships among the perspectives to understand how focus on one affects performance in others in both the short and long term (see Multimedia Appendix 4 for some of the tradeoffs that merit consideration). The relationship is dependent on case characteristics and is therefore not conclusive. For example, as health care organizations learn how to better use blockchain applications, they can use this knowledge to improve their internal processes. Efficient and effective processes can lead to improved service quality, thereby increasing customer satisfaction and revenue in the long term. In turn, organizations can invest more resources in identifying opportunities to learn and develop blockchain applications across the various units. Similarly, an existing health care system may provide a moderate level of data protection that can be achieved with minimal investment, moderate levels of customer satisfaction, and minimal changes to internal processes and learning capabilities. When providing more secure protection of patients' medical data becomes a top priority for compliance with external and regulatory requirements, organizations may consider adopting a blockchain solution. From the financial perspective, adopting blockchain applications may have a negative impact on organizations as it increases costs in the short term. In addition, blockchain adoption may decrease customer satisfaction in the short term until customers become familiar with the new systems and realize value through capabilities such as ease of access and control. These short-term negative impacts from the customer and on financial perspectives may delay the adoption of improvements to internal processes. In the long term, however, improvements to internal processes that are facilitated by the technology may positively affect customer satisfaction. In addition, process improvements can facilitate learning capabilities, which, in turn, positively affect internal processes and organizational finances in the long term.

Case Study: Analysis of the Proposed Extended Balanced Scorecard with a Blockchain Implementation in Health Care

Outline

What follows is a case study applying the BSC framework to the implementation of a blockchain in health care. PharmaChain Inc is a business unit that manages aviation and trucking transportation within the supply chain journey for pharmaceuticals. PharmaChain Inc prides itself on maintaining pharmaceutical supply chain industry certification to handle high-value, temperature-sensitive cargo. The highest impact of blockchain implementation is providing greater visibility and transparency, thereby ensuring the safe transportation of life-saving pharmaceuticals. Business leaders suggest that this blockchain application use case, managing aviation and trucking of pharmaceutical products from manufacturers to health care providers, serves as an example of PharmaChain Inc's commitment to pursuing high impact innovation.

While stakeholders often have varying perspectives and goals, this use case illustrates significant benefits for two important stakeholder groups, namely, customers and providers. The varying stakeholder goals within supply chains results in operational complexity when the process is desynchronized. Blockchain technology helps standardize stakeholder interactions, contributing mutual benefit to the provider and the customer. Standardization of interactions, in turn, reduces human intervention and results in accrual of added business value to all stakeholders. See Figure 2 for a summary of the case study.

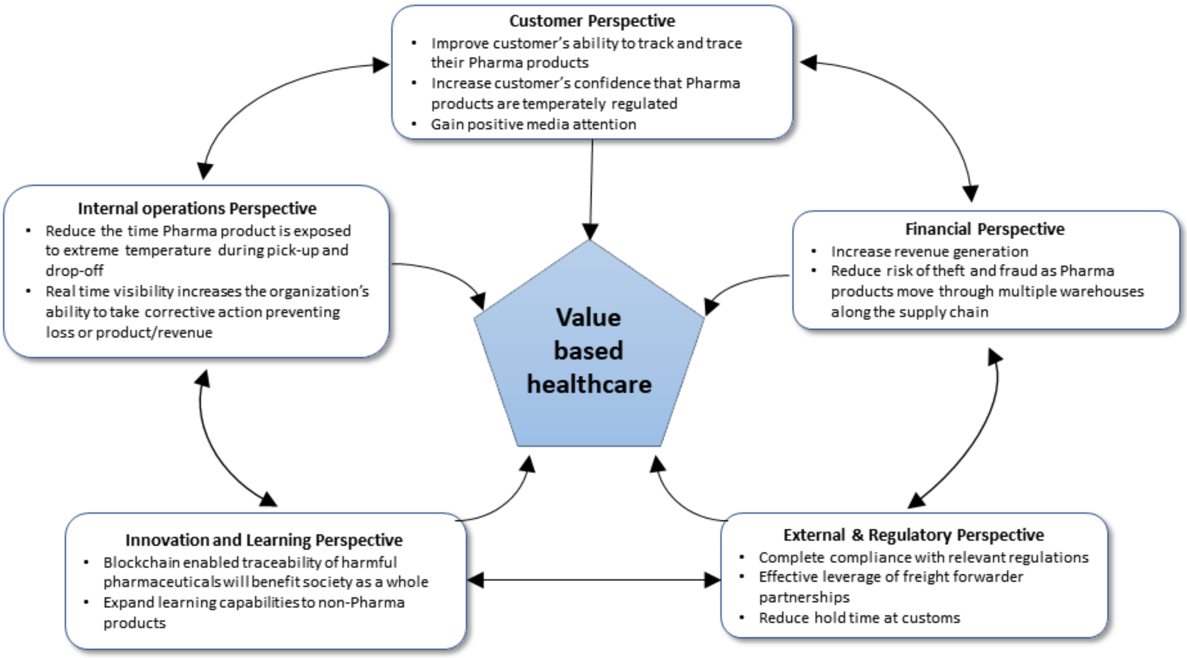


Figure 2. Case study: Application of developed framework in the pharmaceutical supply chain.

Customer Perspective

Customers are at the center of all decisions at PharmaChain Inc. PharmaChain Inc is committed to customer service and innovation, and these two values guide its decision to strengthen its pharmaceutical transportation services. The blockchain solution enables customers to track and trace their temperature-regulated pharmaceutical products, thereby increasing consumer confidence. As an additional benefit, the organization receives positive media attention regarding its commitment to safely transporting pharmaceutical products, which positively strengthens the company's relationship with its customer base. Thus, the customer gains real-time access via mobile device or desktop computer to trustworthy information via the blockchain, and the need to contact customer service, which can be time consuming for the customer and costly for the company, is removed.

Internal Perspective

PharmaChain Inc explores various parts of the internal and external process to solve customer challenges. Blockchain aids in the reduction of lags in the internal processes between temperature measurement and timely corrective actions. Those lags may have otherwise resulted in increased liability, loss of product efficacy, and product destruction. This implementation facilitates the monitoring of the pharmaceutical products' exposure to undesirable conditions (such as temperature extremes and delays in transit).

External Perspective

Conversely, blockchain delivers external process improvements leveraged to resolve legal and compliance issues more rapidly, ultimately allowing lifesaving medicine to reach patients more quickly by eliminating customary hold times in customs. A blockchain initiative was selected to improve visibility and facilitate trust among stakeholders (eg, manufacturers, distributors, transporters, government agencies, and pharmacies). If the freight forwarders do not produce and submit customs approval forms in a timely fashion, the pharmaceutical products cannot be released, with the duration of the hold potentially affecting the quality of the product and negatively affecting the customer experience. A trusted blockchain minimizes the standard 4- to 8-hour hold duration necessary to verify the validity of the customs approval submitted by the freight forwarder, and improved compliance also helps increase trust among external partners.

Learning and Growth (Innovation) Perspective

Blockchain applications support PharmaChain Inc in improving its learning capabilities by enabling it to analyze its business processes and optimize them. The learning capabilities can be extended beyond pharmaceutical products, resulting in organizational efficiencies. The growth opportunity within blockchain applications is enabling traceability along the supply chain journey. Traceability helps reduce fraud in the pharmaceutical supply chain, which is a major societal benefit. Encouraged by the success of the initiative, the organization is deploying blockchain across multiple business products, especially for high value activities and products like pet transportation and food items.

Financial Perspective

For PharmaChain Inc, pharmaceuticals represent one of its highest grossing revenue centers among all its shipping products. With a supply chain industry ripe for innovation, PharmaChain Inc accepts that a financial investment must be made to realize the key benefits of blockchain technology. Blockchain technology reduces the risk of theft and fraud as pharmaceutical products move through multiple warehouses along the supply chain, thus justifying the financial investment. The blockchain solution implemented by PharmaChain Inc positively impacts customer service and internal and external processes, increasing reliability and thereby reducing long-term costs. The risk of theft is minimized due to the automation of security controls, facilitated by the blockchain implementation. In addition, the cost of physical tracking of shipments is also minimized. The organization anticipates a lift of 10% in pharmaceutical sales over an 18-month period due to the initiative.

Tradeoffs Between the Perspectives

Transparency is one of the key characteristics of blockchain that helps to facilitate value within health care. Transparency helps ensure the authenticity of the pharmaceutical products while providing a lone source of the truth for the pharmacy supply chain network. However, transparency comes with tradeoffs between the value-based perspectives. For example, transparency replaces the concept of *need to know* that previously existed between the internal operational perspective and the customer perspective. Prior to the adoption of the blockchain solution, process improvements that were necessary to address internal operational failure were implemented only when the benefits outweighed the costs. With the introduction of blockchain, increased transparency may increase the exposure of failures in internal operations to the entire supply chain network, which, in turn, may reduce confidence in PharmaChain Inc. Therefore, any deficiencies identified in internal processes will be addressed more rapidly. While this increases the cost of the pharmacy product in the short term, it is likely to improve performance in the long term. Since blockchain in pharmaceuticals is transformational in providing trusted information, positive media attention that results from being an innovator in the industry provides additional opportunities for expanding the customer base.

Thus, PharmaChain Inc needs to continuously balance competing demands to improve internal operations and to innovate. Blockchain innovations require financial and human capital investments, which compete with the demands to improve existing internal systems. Thus, at least in the short term, increased quality of services provided to the customer (for example, via the ability to track and trace pharmaceutical products) may negatively affect the financial metrics. However, the benefits are expected to significantly increase financial performance in the long term as the blockchain technology enables PharmaChain Inc to offer superior services in comparison to its competition, thus providing PharmaChain Inc the opportunity to strengthen its competitive position in the industry.

Discussion

Thus, we provide a comprehensive framework that can be used to evaluate blockchain implementation in the value-based health care context, and our study contributes to research streams on blockchain technology, the balanced scorecard framework, and value-based care.

First, our framework can help decision makers in health care organizations evaluate the feasibility and utility of various blockchain proposals that seek to address the health IT chasm reported in prior research [1]. We examined the health IT chasm from three stakeholder perspectives to identify how blockchain-based solutions can resolve these issues based on existing use cases (Multimedia Appendix 1). However, because this disruptive technology is still in its infancy, having a holistic view of the value of blockchain applications is critical to making informed strategic investment decisions [55,81]. Our framework will aid health care organizations in holistically considering the implications of blockchain technology from five critical perspectives. While prior literature has identified three groups of stakeholders central to the delivery of value-based care [1], our study additionally highlights the critical role of external stakeholders and regulations.

In addition, our study extends the BSC framework by emphasizing the importance of the external perspective within the health care domain. The health care domain is a dynamic environment marked by changing regulations as well as competitive forces that are charting the course of the industry more rapidly than ever before. Regulatory compliance and value-based provision of services and products are two salient considerations in the health care industry. While value can be created through external partnerships, interoperability among IT systems and regulatory compliance are two areas of concern that constrain such partnerships. Blockchain's inherent characteristics, such as transparency, immutability, and traceability, facilitate interoperability and enable health care organizations to both cocreate value with their external stakeholders and comply with regulations. Considering the influence of the external environment on a health care organization's existence, our framework enables the examination of the external perspective when evaluating the performance of blockchain-based HIT solutions.

Third, with their emphasis on value-based care, health care organizations need to develop integrated health care IT infrastructure that can improve services and reduce medical errors. Blockchain, with its inherent trust- and security-promoting qualities, has the potential to significantly affect various areas of value provision for patients in health care. While many performance evaluation solutions exist, our study demonstrates the unique aspects of BSC in evaluating IT initiatives for enabling value-based care. The BSC framework enables the consideration of both financial and nonfinancial dimensions of IT initiatives in the short term as well as the long term. When compared with other performance evaluation solutions (such as Zachman's framework, the HCI framework, or the technology-centric framework), our extended BSC framework facilitates consideration of the external perspective. It also defines and assesses performance against operational metrics for each of the five critical perspectives. In addition, our approach highlights the importance of the interrelationships among the perspectives, thus offering another critical extension of the BSC approach. The BSC, however, is limited in its ability to build intuitive and interactive systems like those that HCI and other frameworks provide. Thus, we recommend combining the BSC approach with other appropriate framework(s) to meet an organization's unique needs.

Finally, our case study illustrates how the proposed framework can be utilized to evaluate a health care blockchain application in the for-profit sector. Our approach can also be extended to not-for-profit organizations, which prioritize social goals over financial goals. In such organizations, the financial perspective can be modified to focus on financial sustainability by establishing metrics such as cost reduction, revenue growth, and cost of stakeholder engagement. Similarly, the customer perspective may be widened to include additional stakeholders, such as donors, funding sources, community, volunteers, and employees, that are critical to such organizations [82].

Conflicts of Interest

None declared.

Multimedia Appendix 1
How blockchain can empower value-based care.
[PDF File (Adobe PDF File)230 KB - jmir_v21i9e13595_app1.pdf]

Multimedia Appendix 2
Assessment of performance evaluation frameworks.
[PDF File (Adobe PDF File)188 KB - jmir_v21i9e13595_app2.pdf]

Multimedia Appendix 3
Metrics (KPIs) per perspectives.
[PDF File (Adobe PDF File)182 KB - jmir_v21i9e13595_app3.pdf]

Multimedia Appendix 4
Relationship matrix among perspectives.
[PDF File (Adobe PDF File)104 KB - jmir_v21i9e13595_app4.pdf]

References

1. Adler-Milstein J, Embi PJ, Middleton B, Sarkar IN, Smith J. Crossing the health IT chasm: considerations and policy recommendations to overcome current challenges and enable value-based care. J Am Med Inform Assoc 2017 Sep 01;24(5):1036-1043. [doi: 10.1093/jamia/ocx017] [Medline: 28340128]
2. Mettler M. Blockchain technology in healthcare: the revolution starts here. : The revolution starts here. 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom) IEEE; 2016 Presented at: IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom); 14-16 Sept. 2016; Munich, Germany p. 1-3 URL: https://ieeexplore.ieee.org/document/7749510 [doi: 10.1109/healthcom.2016.7749510]
3. Nakamoto S. Bitcoin. 2008. Bitcoin: a peer-to-peer electronic cash system URL: https://bitcoin.org/bitcoin.pdf [accessed 2018-12-10]

4. Kim HM, Laskowski M. Toward an ontology-driven blockchain design for supply-chain provenance. Intell Sys Acc Fin Mgmt 2018 Mar 28;25(1):18-27. [doi: 10.1002/isaf.1424]
5. Marr B. Forbes. 2018. 35 Amazing real world examples of how blockchain is changing our world URL: https://www.forbes.com/sites/bernardmarr/2018/01/22/35-amazing-real-world-examples-of-how-blockchain-is-changing-our-world/#1f18c5943b56 [accessed 2018-11-10]
6. Lagasse J. Healthcare Finance. 2017. Curisium raises $3.5 million to bring blockchain to value-based contracting URL: https://www.healthcarefinancenews.com/news/curisium-raises-35-million-bring-blockchain-value-based-contracting [accessed 2018-11-20]
7. Iansiti M, Lakhani KR. The truth about blockchain. Harvard Business Review 2017;95(1):118-127 [FREE Full text]
8. Glaser F. Pervasive decentralisation of digital infrastructures: a framework for blockchain enabled system and use case analysis. 2017 Presented at: 50th Hawaii International Conference on System Sciences; Jan 1 2017 - Jan 7 2017; Hawaii URL: https://aisel.aisnet.org/hicss-50/da/open_digital_services/4/ [doi: 10.24251/hicss.2017.186]
9. LeRouge C, Mantzana V, Wilson EV. Healthcare information systems research, revelations and visions. European Journal of Information Systems 2017 Dec 19;16(6):669-671. [doi: 10.1057/palgrave.ejis.3000712]
10. Kaplan RS, Norton DP. The balanced scorecard: measures that drive performance. Harvard Business Review 1992;70(1):71-79 [FREE Full text]
11. Jain R, Benbunan-Fich R, Mohan K. Assessing Green IT Initiatives Using the Balanced Scorecard. IT Prof 2011 Jan;13(1):26-32. [doi: 10.1109/mitp.2010.139]
12. Voelker KE, Rakich JS, French GR. The balanced scorecard in healthcare organizations: a performance measurement and strategic planning methodology. Hosp Top 2001;79(3):13-24. [doi: 10.1080/00185860109597908] [Medline: 11794940]
13. Slabodkin G. A work in progress: blockchain. Health Data Management 2017;25(3):37-39.
14. ONC HHS. The Office of the National Coordinator for Health Information Technology. 2015. Federal health IT strategic plan URL: https://www.healthit.gov/sites/default/files/9-5-federalhealthitstratplanfinal_0.pdf [accessed 2018-11-20]
15. Patel MS, Davis MM, Lypson ML. The VALUE Framework: training residents to provide value-based care for their patients. J Gen Intern Med 2012 Sep;27(9):1210-1214 [FREE Full text] [doi: 10.1007/s11606-012-2076-7] [Medline: 22573146]
16. Porter ME. What Is Value in Health Care? N Engl J Med 2010 Dec 23;363(26):2477-2481. [doi: 10.1056/nejmp1011024]
17. Tseng EK, Hicks LK. Value Based Care and Patient-Centered Care: Divergent or Complementary? Curr Hematol Malig Rep 2016 Aug;11(4):303-310. [doi: 10.1007/s11899-016-0333-2] [Medline: 27262855]
18. Porter ME. A strategy for health care reform--toward a value-based system. N Engl J Med 2009 Jul 09;361(2):109-112. [doi: 10.1056/NEJMp0904131] [Medline: 19494209]
19. LaPointe J. Revcycle Intelligence. 2016 Jun 07. What is value-based care, what it means for providers? URL: https://revcycleintelligence.com/features/what-is-value-based-care-what-it-means-for-providers [accessed 2018-12-10]
20. Office of the National Coordinator for Health Information Technology. 2015. U.S. hospital adoption of patient engagement functionalities URL: https://dashboard.healthit.gov/quickstats/pages/FIG-Hospital-Adoption-of-Patient-Engagement-Functionalities.php [accessed 2018-11-20]
21. Zhang J, Xue N, Huang X. A Secure System For Pervasive Social Network-Based Healthcare. IEEE Access 2016;4:9239-9250. [doi: 10.1109/ACCESS.2016.2645904]
22. Engelhardt MA. Hitching Healthcare to the Chain: An Introduction to Blockchain Technology in the Healthcare Sector. TIM Review 2017 Oct 27;7(10):22-34. [doi: 10.22215/timreview/1111]
23. Ekblaw A, Azaria A, Halamka J, Lippman A. A case study for blockchain in healthcare:. 2016 Presented at: Proceedings of IEEE Open & Big Data Conference; Aug 22, 2016 - Aug 24, 2016; Vienna, Austria.
24. Dey T, Jaiswal S, Sunderkrishnan S, Katre N. HealthSense: a medical use case of Internet of Things and blockchain. 2017 Presented at: International conference on intelligent sustainable systems (ICISS); Dec 7, 2017 - Dec 8, 2017; Palladam, India. [doi: 10.1109/iss1.2017.8389459]
25. Ichikawa D, Kashiyama M, Ueno T. Tamper-Resistant Mobile Health Using Blockchain Technology. JMIR Mhealth Uhealth 2017 Jul 26;5(7):e111 [FREE Full text] [doi: 10.2196/mhealth.7938] [Medline: 28747296]
26. Saravanan M, Shubha R, Marks A, Iyer V. SMEAD: a secured mobile enabled assisting device for diabetics monitoring. 2017 Presented at: IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS); Dec 17, 2017 - Dec 20, 2017; Bhubaneswar, India p. 1-6. [doi: 10.1109/ants.2017.8384099]
27. Carson B, Romanelli G, Walsh P, Zhumaev A. McKinsey & Company. 2018. Blockchain beyond the hype: what is the strategic business value? URL: https://cybersolace.co.uk/CySol/wp-content/uploads/2018/06/McKinsey-paper-about-Blockchain-Myths.pdf [accessed 2018-11-20]
28. Genestier P, Zouarhi S, Limeux P, Excoffier D, Prola A. Blockchain for consent management in the ehealth environment?: a nugget for privacy and security challenges. Journal of the International Society for Telemedicine and eHealth 2017;5:1-4 [FREE Full text]
29. Chandler N. Gartner. 2010 Apr 02. Solutions for scorecards and strategy management URL: https://www.gartner.com/en/documents/1334316/solutions-for-scorecards-and-strategy-management [accessed 2018-12-10]
30. Nørreklit H. The Balanced Scorecard: what is the score? A rhetorical analysis of the Balanced Scorecard. Accounting, Organizations and Society 2003 Aug;28(6):591-619. [doi: 10.1016/s0361-3682(02)00097-1]

31. Glover D, Hermans J. Improving the traceability of the clinical trial supply chain. Applied Clinical Trials 2017;26(12):36-39 [FREE Full text]
32. Funk E, Riddell J, Ankel F, Cabrera D. Blockchain Technology: A Data Framework to Improve Validity, Trust, and Accountability of Information Exchange in Health Professions Education. Acad Med 2018 Dec;93(12):1791-1794. [doi: 10.1097/ACM.0000000000002326] [Medline: 29901658]
33. Moreira C, Farley M, Gebreyes K, McDonnell W. PwC. 2018. A prescription for blockchain and healthcare: reinvent or be reinvented URL: https://www.pwc.com/us/en/health-industries/health-research-institute/assets/pdf/pwc-hri-a-prescription-for-blockchain-in-healthcare_27sept2018.pdf [accessed 2018-12-10]
34. Androulaki E, Barger A, Bortnikov V, Cachin C, Christidis K, De Caro A, et al. Hyperledger fabric: A distributed operating system for permissioned blockchains. In: www.hyperledger.org.: ACM; 2018 Presented at: Thirteenth EuroSys Conference; April 23, 2018 - April 26, 2018; Porto, Portugal. [doi: 10.1145/3190508.3190538]
35. Bocek T, Rodrigues B, Strasser T, Stiller B. Blockchains everywhere - a use-case of blockchains in the pharma supply-chain. 2017 Presented at: IFIP/IEEE Symposium on Integrated Network and Service Management; May 8, 2017 - May 12, 2017; Lisbon, Portugal. [doi: 10.23919/inm.2017.7987376]
36. HealthIT.: IBM; 2016. Blockchain: the chain of trust and its potential to transform healthcare – our point of view URL: https://www.healthit.gov/sites/default/files/8-31-blockchain-ibm_ideation-challenge_aug8.pdf [accessed 2018-12-11]
37. Nugent T, Upton D, Cimpoesu M. Improving data transparency in clinical trials using blockchain smart contracts. F1000Res 2016;5:2541 [FREE Full text] [doi: 10.12688/f1000research.9756.1] [Medline: 28357041]
38. Benchoufi M, Ravaud P. Blockchain technology for improving clinical research quality. Trials 2017 Jul 19;18(1):335 [FREE Full text] [doi: 10.1186/s13063-017-2035-z] [Medline: 28724395]
39. Hollan J, Hutchins E, Kirsh D. Distributed cognition: toward a new foundation for human-computer interaction research. ACM Trans. Comput.-Hum. Interact 2000;7(2):174-196. [doi: 10.1145/353485.353487]
40. Zhang A, Lin X. Towards Secure and Privacy-Preserving Data Sharing in e-Health Systems via Consortium Blockchain. J Med Syst 2018 Jun 28;42(8):140. [doi: 10.1007/s10916-018-0995-5] [Medline: 29956061]
41. Yue X, Wang H, Jin D, Li M, Jiang W. Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control. J Med Syst 2016 Oct;40(10):218. [doi: 10.1007/s10916-016-0574-6] [Medline: 27565509]
42. Cichosz SL, Stausholm MN, Kronborg T, Vestergaard P, Hejlesen O. How to Use Blockchain for Diabetes Health Care Data and Access Management: An Operational Concept. J Diabetes Sci Technol 2019 Mar;13(2):248-253 [FREE Full text] [doi: 10.1177/1932296818790281] [Medline: 30047789]
43. Carson B, Romanelli G, Walsh P, Zhumaev A. McKinsey & Company. 2018. Blockchain beyond the hype: what is the strategic business value? URL: https://cybersolace.co.uk/CySol/wp-content/uploads/2018/06/McKinsey-paper-about-Blockchain-Myths.pdf [accessed 2018-12-10]
44. Agbo C, Mahmoud Q, Eklund J. Blockchain Technology in Healthcare: A Systematic Review. Healthcare (Basel) 2019 Apr 04;7(2):56 [FREE Full text] [doi: 10.3390/healthcare7020056] [Medline: 30987333]
45. Leung D, Suhl A, Gilad Y, Zeldovich N. Network and Distributed Systems Security (NDSS) Symposium. 2019. Vault: fast bootstrapping for the algorand cryptocurrency URL: https://people.csail.mit.edu/nickolai/papers/leung-vault.pdf [accessed 2019-05-20]
46. Kuzmanovic A. Net neutrality. Commun. ACM 2019 Apr 24;62(5):50-55. [doi: 10.1145/3312525]
47. Park YR, Lee E, Na W, Park S, Lee Y, Lee J. Is Blockchain Technology Suitable for Managing Personal Health Records? Mixed-Methods Study to Test Feasibility. J Med Internet Res 2019 Feb 08;21(2):e12533 [FREE Full text] [doi: 10.2196/12533] [Medline: 30735142]
48. Kuo T, Kim H, Ohno-Machado L. Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 2017 Nov 01;24(6):1211-1220 [FREE Full text] [doi: 10.1093/jamia/ocx068] [Medline: 29016974]
49. Liu W, Zhu S, Mundie T, Krieger U. Advanced block-chain architecture for e-health systems. : IEEE; 2017 Presented at: IEEE 19th International Conference on e-Health Networking, Applications and Services, Healthcom; Oct 12, 2017 - Oct 15, 2017; Dalian, China p. 37-42. [doi: 10.1109/healthcom.2017.8210847]
50. Roman-Belmonte JM, De la Corte-Rodriguez H, Rodriguez-Merchan EC. How blockchain technology can change medicine. Postgrad Med 2018 May;130(4):420-427. [doi: 10.1080/00325481.2018.1472996] [Medline: 29727247]
51. Angraal S, Krumholz HM, Schulz WL. Blockchain Technology: Applications in Health Care. Circ Cardiovasc Qual Outcomes 2017 Sep;10(9):1-3. [doi: 10.1161/CIRCOUTCOMES.117.003800] [Medline: 28912202]
52. Zhou L, Wang L, Sun Y. MIStore: a Blockchain-Based Medical Insurance Storage System. J Med Syst 2018 Jul 02;42(8):149 [FREE Full text] [doi: 10.1007/s10916-018-0996-4] [Medline: 29968202]
53. Miller R. TechCrunch. PokitDok teams with Intel on healthcare blockchain solution URL: https://techcrunch.com/2017/05/10/pokitdok-teams-with-intel-on-healthcare-blockchain-solution/ [accessed 2018-12-11]
54. Schumacher A. Blockchain Research Institute.: Brightline; 2018. Reinventing healthcare on the blockchain - toward a new era in precision medicine URL: https://www.brightline.org/resources/reinventing-healthcare-on-the-blockchain/ [accessed 2018-12-11]
55. Vazirani AA, O'Donoghue O, Brindley D, Meinert E. Implementing Blockchains for Efficient Health Care: Systematic Review. J Med Internet Res 2019 Feb 12;21(2):e12439 [FREE Full text] [doi: 10.2196/12439] [Medline: 30747714]

56. Meinert E, Alturkistani A, Brindley D, Knight P, Wells G, Pennington ND. The technological imperative for value-based health care. Br J Hosp Med (Lond) 2018 Jun 02;79(6):328-332. [doi: 10.12968/hmed.2018.79.6.328] [Medline: 29894248]
57. Sullivan SD, Watkins J, Sweet B, Ramsey SD. Health technology assessment in health-care decisions in the United States. Value Health 2009 Jun;12 Suppl 2:S39-S44 [FREE Full text] [doi: 10.1111/j.1524-4733.2009.00557.x] [Medline: 19523183]
58. Chernew ME, Rosen AB, Fendrick AM. Value-based insurance design. Health Aff (Millwood) 2007;26(2):w195-w203. [doi: 10.1377/hlthaff.26.2.w195] [Medline: 17264100]
59. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press; 2015.
60. Sowa JF, Zachman JA. Extending and formalizing the framework for information systems architecture. IBM Syst. J 1992;31(3):590-616. [doi: 10.1147/sj.313.0590]
61. O'Brien MA, Rogers WA, Fisk AD. Developing a Framework for Intuitive Human-Computer Interaction. Proc Hum Factors Ergon Soc Annu Meet 2008 Sep;52(20):1645-1649 [FREE Full text] [doi: 10.1177/154193120805202001] [Medline: 25552895]
62. Cheng LC, Gibson ML, Carrillo EE, Fitch G. A technology‐centric framework for investigating business operations. Industr Mngmnt & Data Systems 2011 Apr 26;111(4):509-530. [doi: 10.1108/02635571111133524]
63. Hoque Z. Total quality management and the balanced scorecard approach: a critical analysis of their potential relationships and directions for research. Critical Perspectives on Accounting 2003 Jul;14(5):553-566. [doi: 10.1016/s1045-2354(02)00160-0]
64. Hasan N. University of Birmingham. 2006. Developing a balanced scorecard model for evaluation of project management and performance URL: https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.494604 [accessed 2018-12-10]
65. Metawie M, Gilman D. Problems with the implementation of performance measurement systems in the public sector where performance is linked to pay: A literature review drawn from the UK. 2005 Presented at: 3rd Conference on Performance Measurement and Management Control; Sept 22, 2005 - Sept 23, 2005; Nice, France URL: http://citeseerx.ist.psu.edu/viewdoc/citations?doi=10.1.1.104.9267
66. Enwere EN, Keating EA, Weber RJ. Balanced scorecards as a tool for developing patient-centered pharmacy services. Hosp Pharm 2014 Jun;49(6):579-584 [FREE Full text] [doi: 10.1310/hpj4906-579] [Medline: 24958976]
67. Zelman WN, Pink GH, Matthias CB. Use of the balanced scorecard in health care. J Health Care Finance 2003;29(4):1-16. [Medline: 12908650]
68. Bandopadhyay T. Gartner. 2013 Dec 6. I&O Teams Can Build Balanced Scorecards for Business Cases on IT Services Transitioning to the Cloud.
69. Schelp J, Stutz M. Via Nova Architectura. 2007. A balanced scorecard approach to measure the value of enterprise architecture URL: https://www.alexandria.unisg.ch/213190/1/SchelpStutz.2007-BalancedScorecardApproach.pdf [accessed 2018-12-15]
70. Rillo M. Limitations of balanced scorecard. 2004 Presented at: Proceedings of the 2nd Scientific and Educational Conference, Business Administration: Business in a Globalizing Economy; Jan 30, 2004 - Jan 31, 2004; Parnu p. 155-161.
71. Banker R, Chang H, Janakiraman S, Konstans C. A balanced scorecard analysis of performance metrics. European Journal of Operational Research 2004 Apr;154(2):423-436 [FREE Full text] [doi: 10.1016/S0377-2217(03)00179-6] [Medline: 27695049]
72. Mohobbot A. The balanced scorecard (BSC): a critical analysis. Journal of Humanities and Social Sciences 2004;18:219-232.
73. Akkermans H, Van Oorschot K. Developing a balanced scorecard with system dynamics. Journal of the Operational Research Society 2002;40(56):931-941 [FREE Full text]
74. Grigoroudis E, Orfanoudaki E, Zopounidis C. Strategic performance measurement in a healthcare organisation: A multiple criteria approach based on balanced scorecard. Omega 2012 Jan;40(1):104-119. [doi: 10.1016/j.omega.2011.04.001]
75. Culver K. HealthIT. 2016. Blockchain technologies: a whitepaper discussing how the claims process can be improved URL: https://www.healthit.gov/sites/default/files/3-47-whitepaperblockchainforclaims_v10.pdf [accessed 2019-04-12]
76. Yip K. HealthIT. 2017. Blockchain and alternative payment models URL: https://www.healthit.gov/sites/default/files/15-54-kyip_blockchainapms_080816.pdf [accessed 2019-09-14]
77. Pankomera R, van Greunen D. Privacy and security issues for a patient-centric approach in public healthcare in a resource constrained setting. In: IEEE. 2016 Presented at: IST-Africa Week Conference; May 11, 2016 - May 13, 2016; Durban, South Africa.
78. DeLone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. Journal of Management Information Systems 2003;19(4):9-30. [doi: 10.1080/07421222.2003.11045748]
79. Kettinger, Lee. Zones of Tolerance: Alternative Scales for Measuring Information Systems Service Quality. MIS Quarterly 2005;29(4):607-623. [doi: 10.2307/25148702]
80. Hu C. Business Insider. 2018 Aug 10. Mount Sinai teamed up with the designers who created projects for Nike and Beyonce to build a futuristic, new clinic ? and it?s reimagining how healthcare is delivered URL: https://www.businessinsider.com.au/inside-mount-sinais-lab100s-clinic-of-the-future-2018-8 [accessed 2018-12-10]
81. Zhang P, Walker M, White J, Schmidt D, Lenz G. Metrics for assessing blockchain-based healthcare decentralized apps. 2017 Presented at: IEEE 19th International Conference on E-Health Networking, Applications and Services (Healthcom); Oct 12, 2017 - Oct 15, 2017; Dalian, China. [doi: 10.1109/healthcom.2017.8210842]

82. Somers AB. Shaping the balanced scorecard for use in UK social enterprises. Social Enterprise Journal 2005 Mar;1(1):43-56. [doi: 10.1108/17508610580000706]

Abbreviations

BMPLS: Blockchain-Based Multi-level Privacy-Preserving Location Sharing Scheme
BSC: Balanced Scorecard
EHR: electronic health record
EFQMEM: European Foundation Quality Management Excellence Model
GDPR: European General Data Protection Regulation
HCI: human-computer interaction
HIT: Health Information Technology
IS: Information Systems
IT: Information Technology
ONC: Office of the National Coordinator for Health Information Technology
ROI: return on investment
TQM: total quality management

Edited by K Clauson, P Zhang; submitted 01.02.19; peer-reviewed by H Oh, C Reis, K McLeroy; comments to author 01.04.19; revised version received 27.05.19; accepted 16.07.19; published 27.09.19.

Please cite as:
Zhang R, George A, Kim J, Johnson V, Ramesh B
Benefits of Blockchain Initiatives for Value-Based Care: Proposed Framework
J Med Internet Res 2019;21(9):e13595
URL: https://www.jmir.org/2019/9/e13595
doi:10.2196/13595
PMID:31573899

Original Paper

Connected Health Services: Framework for an Impact Assessment

Ioanna Chouvarda[1], PhD; Christos Maramis[1], PhD; Kristina Livitckaia[1], MSc; Vladimir Trajkovik[2], PhD; Serhat Burmaoglu[3], PhD; Hrvoje Belani[4], MSc; Jan Kool[5], PhD; Roman Lewandowski[6], PhD; The ENJECT Working Group 1 Network

[1]Lab of Computing, Medical Informatics and Biomedical Imaging Technologies, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece

[2]Faculty of Computer Science and Engineering, Saints Cyril and Methodius University, Skopje, North Macedonia

[3]Department of Health Management, Faculty of Economics and Administrative Sciences, Izmir Katip Celebi University, Izmir, Turkey

[4]IT Division, Ministry of Health, Zagreb, Croatia

[5]Rehabilitation Centre Valens, Valens, Switzerland

[6]Faculty of Management, University of Social Sciences, Lodz, Poland

Corresponding Author:
Ioanna Chouvarda, PhD
Lab of Computing, Medical Informatics and Biomedical Imaging Technologies
School of Medicine, Faculty of Health Sciences
Aristotle University of Thessaloniki
Aristotle University Campus, Building 16d
Thessaloniki, 54124
Greece
Phone: 30 2310999247
Fax: 30 2310999263
Email: ioanna@med.auth.gr

Abstract

Background: Connected health (CH), as a new paradigm, manages individual and community health in a holistic manner by leveraging a variety of technologies and has the potential for the incorporation of telehealth and integrated care services, covering the whole spectrum of health-related services addressing healthy subjects and chronic patients. The reorganization of services around the person or citizen has been expected to bring high impact in the health care domain. There are a series of concerns (eg, contextual factors influencing the impact of care models, the cost savings associated with CH solutions, and the sustainability of the CH ecosystem) that should be better addressed for CH technologies to reach stakeholders more successfully. Overall, there is a need to effectively establish an understanding of the concepts of CH impact. As services based on CH technologies go beyond standard clinical interventions and assessments of medical devices or medical treatments, the need for standardization and for new ways of measurements and assessments emerges when studying CH impact.

Objective: This study aimed to introduce the CH impact framework (CHIF) that serves as an approach to assess the impact of CH services.

Methods: This study focused on the subset of CH comprising services that directly address patients and citizens on the management of disease or health and wellness. The CHIF was developed through a multistep procedure and various activities. These included, as initial steps, a literature review and workshop focusing on knowledge elicitation around CH concepts. Then followed the development of the initial version of the framework, refining of the framework with the experts as a result of the second workshop, and, finally, composition and deployment of a questionnaire for preliminary feedback from early-stage researchers in the relevant domains.

Results: The framework contributes to a better understanding of what is CH impact and analyzes the factors toward achieving it. CHIF elaborates on how to assess impact in CH services. These aspects can contribute to an impact-aware design of CH services. It can also contribute to a comparison of CH services and further knowledge of the domain. The CHIF is based on 4 concepts, including CH system and service outline, CH system end users, CH outcomes, and factors toward achieving CH impact. The framework is visualized as an ontological model.

Conclusions: The CHIF is an initial step toward identifying methodologies to objectively measure CH impact while recognizing its multiple dimensions and scales.

(J Med Internet Res 2019;21(9):e14005) doi:10.2196/14005

KEYWORDS

connected health; health impact assessment; framework; outcome assessment; enablers and barriers

Introduction

Background Concepts

Presently, information and communication technologies, including a growing number of consumer and medical devices as well as patient services, have created new opportunities to improve the health and well-being of individuals and populations. Such improvements are expected to be successful through behavior change at a personal level, better health care coordination, and multilevel information sharing, gradually building the connected health (CH) landscape [1].

CH, as a new paradigm, manages individual and community health in a connected and holistic manner by leveraging a variety of technologies [2,3]. CH is a promising vehicle for the incorporation of telehealth and integrated care services, covering the whole spectrum of health-related services from the ones directing the *healthy subject* (as a citizen who seeks health service support or a wellness service consumer) to those addressing the *chronic patient* as an integrated (tele) care service beneficiary. The evolution of the CH ecosystem and related concepts (eg, telemedicine) has been discussed from a bibliometric viewpoint in a study by Burmaoglu [4].

The reorganization of services around the person or citizen, with person-centered care being a promising area [5], is expected to bring an important impact in the health domain. This will require addressing a series of concerns in an effective manner to more successfully reach stakeholders: contextual factors influencing the impact of care models, the cost savings associated with CH solutions, and the sustainability of the CH ecosystem. Overall, there is a need to shed light on the concepts of CH impact.

The Aim

This paper aimed to introduce the CH impact framework (CHIF) that serves in the assessment of CH services' impact. CHIF was born from the European Network for the Joint Evaluation of Connected Health Technologies (ENJECT), a network actively involved in the evaluation of CH technologies, funded by European Cooperation in Science and Technology (COST) Action [6]. Within the complete spectrum of CH, this study focused on the subset of CH comprising *services that directly address patients and citizens* at large on the *management of disease or health and wellness*. These CH patient services are heavily dependent on new technologies. Nevertheless, CH services are not considered detached from established health information technology (either secondary care medical technology or technology primarily oriented to the health care professional [HCP], such as ePrescription) and the respective health care services.

The reason for this study focusing on CH patient services, on the verge of consumer informatics, was because this is a new, highly promising area that is unmapped and in a gray zone concerning health care, meaning that there are no explicit care models incorporating CH services, health policies, or guidelines or standard ways of assessing these services.

Related Work and Rationale

When considering health impact assessment (HIA) of a policy, program, or project, its potential effects on the health of a population and the distribution of those effects are evaluated [7,8] so as to produce (1) recommendations supporting decision makers and other stakeholders in making choices about alternatives and (2) improvements to avoid risks, prevent disease or injury, and actively promote health. The impact of CH technologies and services needs to be well defined [9], providing relevant evidence linking to and extending the HIA procedures.

On a broader scale, a relevant study regarding the assessment of integrated care services and scaling up of integrated care in European regions has been conducted by the European Innovation Partnership on Active and Healthy Ageing B3 [10] group (the Action Group on integrated care). The topics addressed included the following: (1) assessment of the health care system's capacity to adopt integrated approaches to deal with challenges of aging, (2) assessment of the uptake of a particular good practice by a health care system, (3) identification of maturity characteristics necessary for adoption and scale-up of good practice, and (4) understanding the context and conditions in adopting and transferring practices among regions. The Maturity Model of B3 group was developed [11] as a tool to assess maturity along 12 dimensions reflecting the various aspects that need to be managed to deliver integrated care.

In addition, with regard to the assessment of telemedicine applications and services, the Model for Assessment of Telemedicine (MAST) tool was developed [12] to describe the effectiveness of telemedicine applications and their contribution to the quality of care. MAST summarizes and evaluates information about the medical, social, economic, and ethical issues related to the use of telemedicine, considering 7 assessment domains (ie, health problem, safety, clinical effectiveness, patient perspectives, financial aspects, organizational aspects, and sociocultural, ethical, and legal aspects). A framework for the emerging area of behavioral interventions was recently proposed [13], yet not elaborating on impact. Methodological aspects for CH evaluation were introduced by O'Leary [14] and Carroll [15], although not uniquely focusing on impact. Consumer health informatics assessment is discussed in the study by Gibbons et al [16]; this includes users, barriers at the system and individual level, implementation of applications (ie, function and process), and outcomes at different levels and directions (intermediate, health care processes, and clinical, economic, and relationship-centered).

There are not many papers that specifically refer to CH and its impact or to the use of specific frameworks for the impact of

CH. In the CH review of Colorafi [17], the theoretical construct of a study by Ryan and Sawin [18] for self-management is adopted. It applies to both chronic conditions and health promotion and considers work, context, process, and proximal and distal outcomes. More specifically, according to this framework, self-management takes place in the context of (1) risk and protective factors specific to the condition, (2) a particular physical and social environment (eg, health care access, culture, and transportation), and (3) a set of individual and family factors (eg, literacy and family structure and capacity to self-manage). Self-management is a process involving individuals and families that includes (1) knowledge, facts, and beliefs (eg, self-efficacy), (2) self-regulation skills and abilities (eg, goal setting, decision making, and emotional control), and (3) social facilitation, including influence, support, and collaboration, to achieve positive health-related outcomes. Interventions to the person and family consider both process and context. The proximal or short-term outcomes lead to the achievement of distal outcomes. Thus, a temporal causal relation is introduced. Proximal outcomes mainly include individual and family self-management behaviors, such as engagement in activities and recommendations of treatment, symptom management, and adherence to recommended pharmacological therapies. Secondly, engagement in health-related behaviors may positively impact the cost of health care services in the short term. The distal outcomes are threefold: (1) health status as an indicator of the disease trajectory (indicating prevention, attenuation, stabilization, and worsening of the condition), (2) quality of life and perceived well-being, and (3) direct and indirect costs.

In the same vein, as services based on CH technologies go beyond standard clinical interventions and assessments of medical devices or medical treatments, the need for standardization and for new ways of measurements emerge when studying CH impact in depth. As mentioned in the study by Colorafi [17], "we likely need more sophisticated study designs if we are to adequately assess which element of a comprehensive program is affecting the outcome, asking how exactly do the ' *interventions impact the psychosocial aspects of the lives of people with diabetes?* '." Therefore, this area needs further research and disambiguation, especially, with regard to outcomes and impact. Although the abovementioned efforts are relevant to the concept of CH and offer valuable insights, CH services constitute more complex constructs that are not compartmentalized and assessed in the same manner as pharmacological trials.

The emerging CH technologies impose the definition of a CHIF. The CHIF was created in the process of exploring concepts around *what CH impact is* and *how it can be described*, *assessed*, and *achieved*. Specifically, CHIF is based on the inputs from the 2 workshops conducted in the scope of ENJECT within the last 2 years. In the following sections, the steps taken for deriving CHIF are presented, and the framework itself is described in detail along with a preliminary assessment tool based on CHIF. The paper also discusses challenges and future steps.

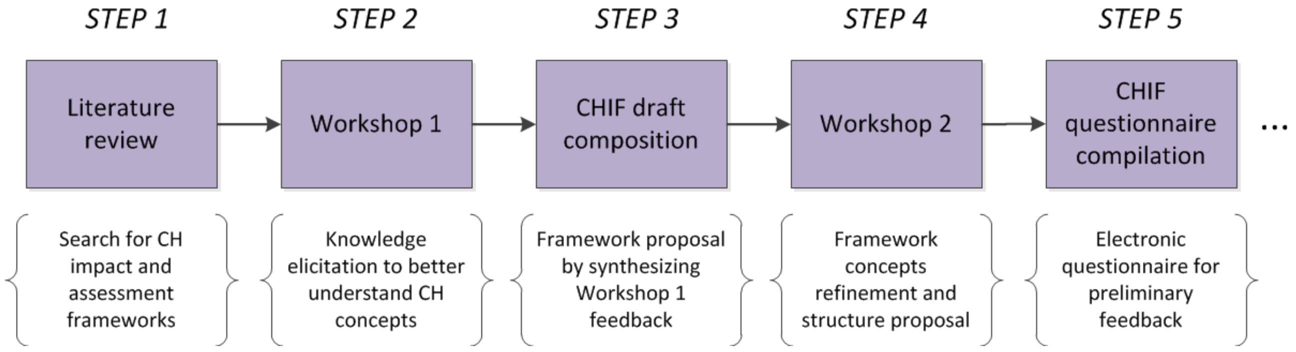
Methods

The formulation of the CHIF framework took place in a multistep process, as delineated below, and it is visually outlined in Figure 1:

- Step 1: A literature review was conducted on the topics of CH impact and assessment frameworks. This step helped identify the main concepts and issues discussed in the domain and helped us further shape our research.
- Step 2: A workshop for further knowledge and insight elicitation toward better understanding the concepts around CH impact was conducted (workshop 1). The methodology employed was based on the structured group feedback approach [19].
- Step 3: Following workshop 1, knowledge elicitation took place, which resulted in a proposal for a CHIF based on a synthesis of inputs.
- Step 4: In workshop 2, the first CHIF proposal was presented and discussed among the participants, and this was the basis for the further effort to organize and propose the framework for reporting impact, including discussion and refinement of the previously established concepts and framework structure. The result was the consolidation of the CHIF structure. In addition, different visual representations of CHIF were suggested, for example, the ontological model.
- Step 5: CHIF was implemented in an electronic questionnaire for preliminary feedback.

More details about the workshops can be found in the Multimedia Appendix 1.

Figure 1. Methodological steps for the derivation of connected health (CH) impact framework.



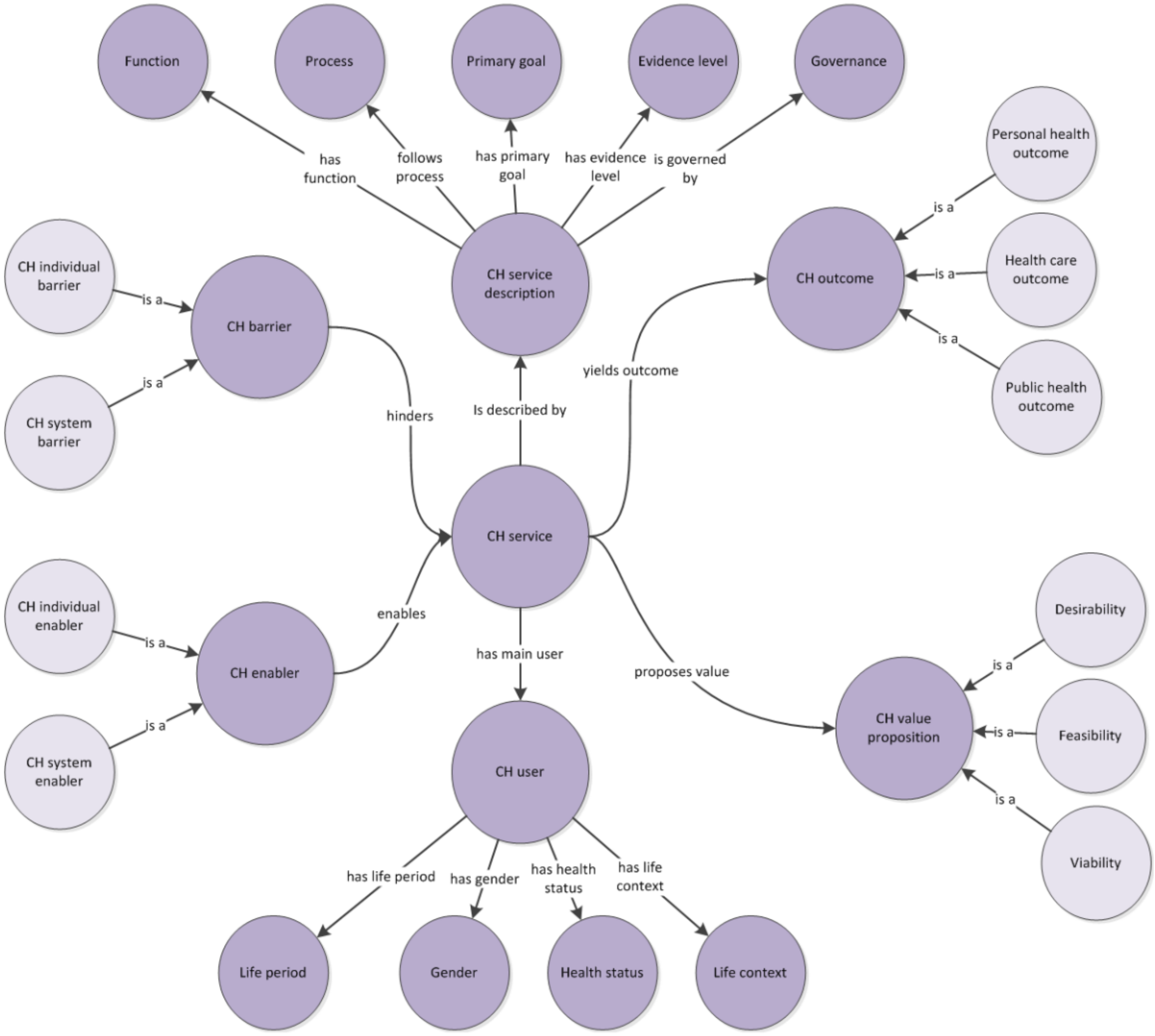
Results

Overview of the Connected Health Impact Framework

The CHIF serves in the assessment of CH services impact. It aims to contribute to (1) a better understanding of what is CH impact, (2) exploring how to achieve impact and thus support in better designing of CH services, and (3) methods to measure and assess impact, which can also help compare the impact of CH services and gather further knowledge. To meet these aims, 4 axes are considered: (1) CH system or service outline—of note, both the concepts of *system* and *service* are mentioned, as the focus is sometimes on the developed application and sometimes on the provided service, which adds a broader scope, (2) CH system end users and their profile, including the profile of primary users that the system targets and secondary users, (3) CH outcomes and measures of impact at different levels, and (4) factors toward achieving CH impact, including barriers and enablers, and a clear value proposition.

The framework is visualized as an ontological model (see Figure 2). CHIF is organized as a tree with the concept of CH service at its root. Nodes beyond the third tree level are not depicted in the figure to reduce the complexity of visualization; however, all the nodes are presented in the following sections.

Figure 2. The ontological model of connected health impact framework; the framework is organized as a tree whose root is the concept of connected health (CH) system and service. Nodes beyond the third tree level are not depicted here, to reduce the complexity of visualization. Note that the arrows tagged as is a denote a subsumption relationship (ie, lighter colored nodes are subclasses of darker colored nodes).



Connected Health Impart Framework Axes

The following subsections refer to the description of the CHIF.

Connected Health System and Service Description

As a prerequisite, CH services need a basic level of functionality description.

We propose 5 elements, helping to describe a CH service through its *function*, *process*, *primary goal*, *evidence level*, and *control*. The *function* element reveals the functionality behind the service, such as assessment or monitoring, knowledge building, disease or condition management, or lifestyle management. The *process* element is responsible for describing how the function is implemented (eg, receiving a measurement from the user and returning automated feedback to the user). This element can include specific components supporting user personalization, such as social interactions and other. The *primary goal* reflects the health-related intent for optimization (eg, daily activity through the number of steps a day and night sleep duration). The *evidence level* describes the validation and evaluation of the service, including technical validation, clinical testing, and user experience. The last proposed element, *control*, refers to the governance of the service on a higher level. The control may belong to the patient or consumer, health care representative, social services, or payers, depending also on the CH services' funding (private or public insurance).

These CH functionality elements can be directly or indirectly linked to impact and further support a better understanding of the service as well as compare services.

Connected Health Users

When addressing the personal CH outcomes coming from a specific technology or service, one has to explicitly specify the offered functionality and aim, as well as the users it addresses or applies to. Particularly, with regard to services and interventions, the targeted users should be well described. It is necessary to note that CH has many *contextual factors* influencing its adoption that should be reflected. These include geographical, social, demographics, human factors, educational, regulation, interoperability, and big data contexts of the specific CH systems within particular deployments. The contextual factors influence the clinical trials concerning the CH. Similarly, CH systems might vary significantly in different geographical regions, because of, for example, different environmental influences (eg, a training or coaching application should provide a different sports suggestion for the desert region). The same argument is valid for different demographics and socioeconomic contexts. The acceptance of services is often determined by human factors, such as the engagement, education level, and—especially—digital literacy of the end users. These contextual factors are essential to understanding that, for example, *the CH system X is efficient for female elder users with dementia in rural areas*. This fact is crucial when designing the system and reporting outcomes [20], especially if aiming at personalized interventions and avoiding *one-size-fits-all* ones.

We propose an initial approach where the *primary user* (ie, patient or citizen) is described with 4 elements: Phase of *life (young, working, retired*, or *dependent)*, *gender*, *health status (healthy, chronic patient, comorbid, acute disease*, or *disabled)*, and *life context*. The concept *young* includes childhood, adolescence, and transition to young adulthood. The last element comprises a list of factors that help describe the life circumstances of the user, including the location of living (eg, rural area), social activity, financial status, and others.

When not targeting consumer apps but health services, other involved users have to be identified, in addition to the direct beneficiary (ie, patient or citizen). The *secondary users* could be HCPs, the state, and policymakers, as well as businesses. CH may have an impact on all the user groups.

Connected Health Outcomes

The health-related outcome of a CH service can be viewed from 3 different perspectives: (1) the personal perspective, (2) the health care process–related perspective, or (3) the wider socioeconomic or public health perspective (see Figure 3). These outcomes belonging to the 3 different perspectives are potentially intertwined.

Personal Outcomes

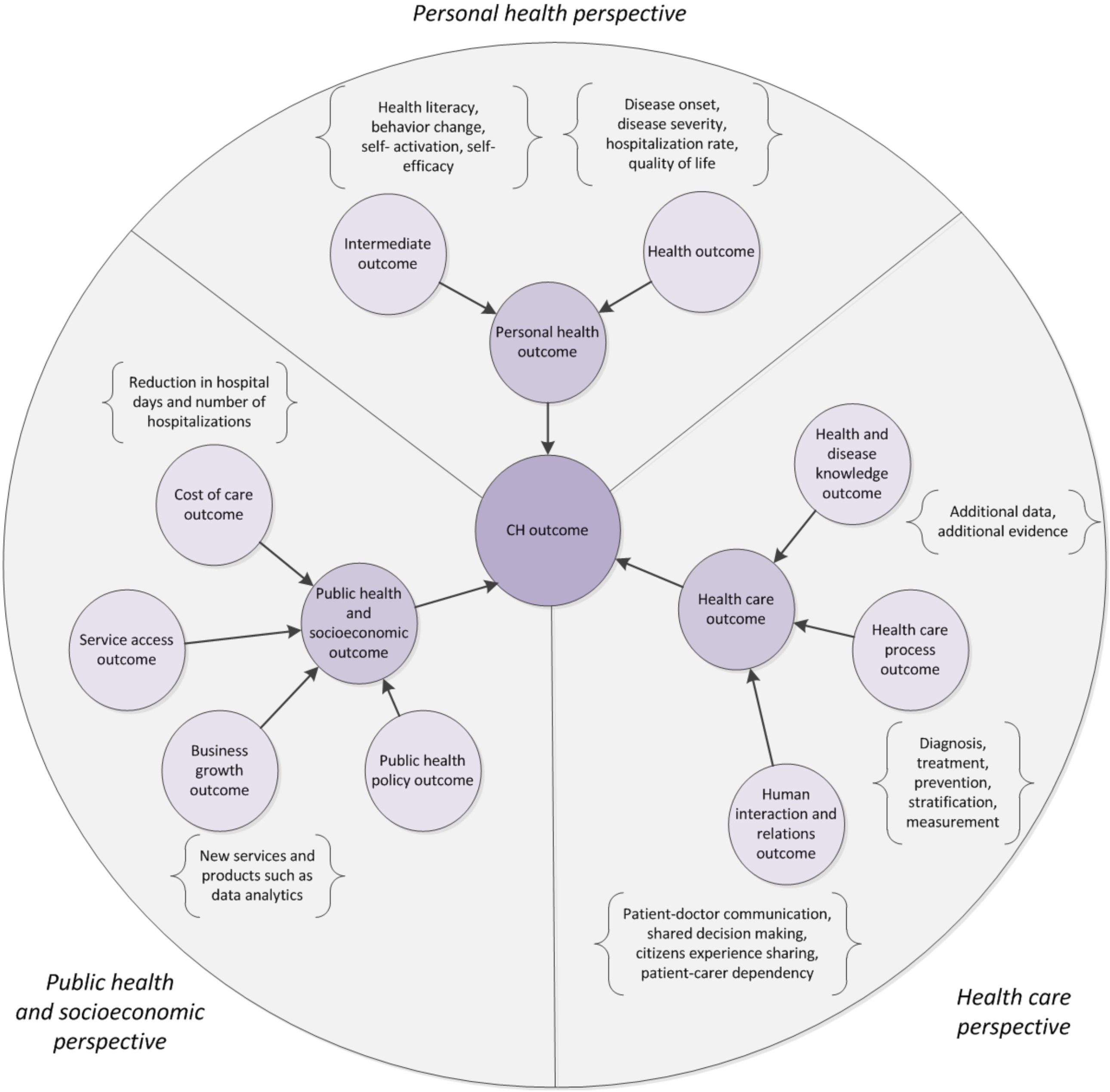
First, CH can affect the patient's or citizen's empowerment and engagement, as well as compliance with treatment [21,22] or other health behavior. These outcomes are expected to lead to the promotion of a healthy lifestyle, with further positive care and social consequences, improved health, and a better quality of life. In this regard, personal health outcomes are divided into the following categories: (1) intermediate outcomes (health literacy, behavior change, self-activation, and self-efficacy) [23] and (2) health outcomes (disease onset, disease deterioration, hospitalization rate, and quality of life), where the former is considered as potential mediators of the latter.

The introduction of CH tools may bring improved self-efficacy, understood here as a person's ability to implement situation-specific behaviors toward attaining established goals, expectations, or designated types of outcomes [24]. Individuals knowing more about their health status may better cope with their health-related problems by themselves. Improved knowledge and understanding about health indicators, achieved through CH, while a person suffers from health problems can also reduce uncertainty in illness.

The timescales are expected to differ depending on the 2 types of the outcome, and many pilot studies decide to report on either but not both. However, reporting on both effects would help better understand the mechanisms of outcome formation and its further impact on the personal level.

A significant challenge is how to best measure these outcomes in a consistent manner, including both subjective or qualitative parts that are mostly measured with questionnaires and objective parts that are quantitatively measured through the use of various devices (eg, number of steps on a pedometer or heart rate on a smartwatch). Another challenge is when to measure the outcomes and, more importantly, how to express their temporal nature. Importantly, personal health outcomes are also linked to the care process outcomes (eg, improved access and accessibility to health care services), especially when combined with health literacy [25].

Figure 3. Perspectives on health-related outcomes that are associated with the impact of connected health (CH) services. All arrows denote is a relationships.



Health Care Process Outcomes

The utilization of CH relates to better patient safety, decreased duration of diagnostic processes (eg, early diagnoses), and better disease management (eg, identification of the risk of deterioration and primary and secondary prevention of disease) [26,27]. In the scope of CH, it is also expected to offer better access to the data, which can be used to improve understanding of the disease (especially in the case of chronic diseases) and provide evidence for health policy makers and other involved stakeholders. CH technologies offer great opportunities for a unified collection of patient-reported outcomes, which can affect the health care process [28].

The introduction of CH services impacts the models of care by enabling novel pathways for health monitoring, which include new interaction models supporting the involvement and empowerment of all stakeholders. These perspectives outline the need for novel clinical health care and social care guidelines, which can influence long-term health strategy design by promoting the economic efficiency of these services.

Overall, the directions identified as regards the care-related impacts of CH can be organized in 3 axes: (1) health care process (diagnosis, treatment, prevention, stratification, and measurement of outcome) [29], (2) human interaction and relations (patient-doctor communication and shared decisions, information and experience sharing, and patient-carer dependency), and (3) new health and disease knowledge (more data and evidence)

Public Health and Socioeconomic Outcomes

There are also horizontal aspects in CH outcomes, which affect multiple stakeholders and levels of health, and thus can be considered both as drivers and outcomes of the CH. These mainly include the facilitation of communication and information flow between health stakeholders and the improvement of health data analytics and management.

A characteristic example is the MyData Nordic Model [30]. This is an infrastructure for human-centered personal data management and processing, aiming to provide individuals with the practical means to access, obtain, and use datasets containing their personal information (ie, medical records, financial information, and data derived from various Web-based services). This approach introduces interesting dynamics at the societal and business level. Another example that relates CH data to public health policies is the BigO program, a European research project that analyzes daily living behavioral patterns of the youth to propose optimal physical activity–, diet-, and nutrition-related policies [31].

Cost reduction as an outcome can be expected at different levels, from the personal to the level of public health. There are studies on the cost-effectiveness of various telemedicine services. The main CH horizontal socioeconomic effects include (1) reduction of cost of care (eg, reduction in hospital days and number of hospitalizations), (2) improved and cost-efficient access to services, (3) improved public health policy, and (4) industrial activity and business growth, related to new services and products (eg, analytics services).

Factors Toward Achieving Connected Health Impact

Connected Health Value Proposition

To achieve scalability and impact, CH value proposition must be clearly articulated. In this respect, it is important to elaborate on what critical information regarding the CH applications is required for understanding the value proposition pertinent to each of the different stakeholders (eg, consumers and patients, their families, clinicians, developers, and payers). This is a clear statement of how the proposed solution relates to some improvement for the user, what specific benefits it brings, and how it differentiates from others. Although the value proposition is a consumer informatics concept, rather than a health care one, this concept may help crystallize the virtues of the CH application and its adoption.

As suggested in a McKinsey report [32], the 3 main properties that generally describe the value proposition of a CH solution are as follows: (1) *desirability* for all involved users (custom-centered and easy to use), (2) *feasibility* both technical and organizational, and (3) *viability* and sustainability (eg, via a supporting ecosystem, involving smart elements, and involving integration and collaboration of stakeholders).

To support these properties, the new solutions should be designed following a user-centered approach to (1) *respect the activities* a potential user has to perform, (2) *meet the expectations* (eg, comfort of use and easy to learn how to use), and (3) *minimize the fears* associated with the solution (eg, fears related to the new technology and fear of high costs). A CH service or product should be proposed based on the abovementioned elements.

Barriers and Enablers of Connected Health Impact

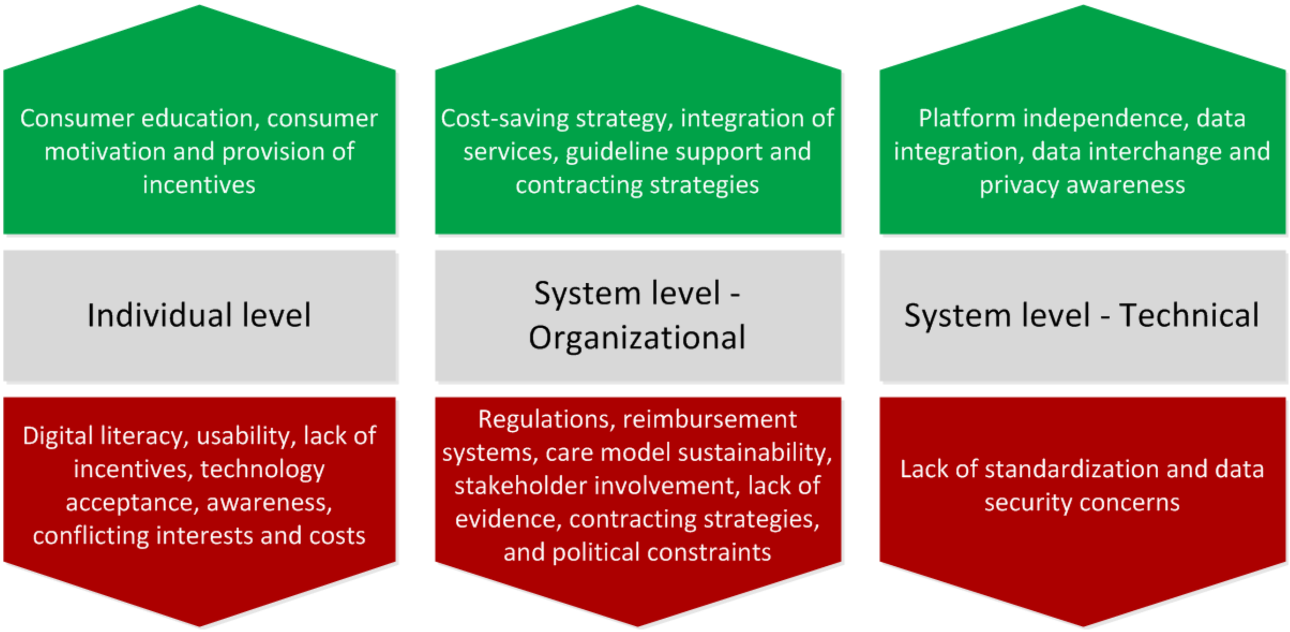
Figure 4 provides a visual outline in the CH impact enablers and barriers that have been identified by CHIF.

When designing and later evaluating a CH service and system or an application, it is necessary to recognize and report barriers that users (clinicians, developers, consumers, their families, caregivers, and policy makers) encounter and that can potentially limit the implementation or utilization of the CH solution. Considering *the user-CH system dipole*, we propose to classify the barriers, as *individual level* ones related to internal user features and abilities (eg, literacy gaps) and barriers external to the user and attributed at the *system level*. The latter might be *technical* (eg, design problems that limit usability) or *organizational* (eg, regulations). The individual level barriers include *digital literacy*, *usability*, *lack of incentives*, *technology acceptance*, *awareness*, *conflicting interests*, and *costs*. A series of system-level barriers related to an organization have been identified, including obstacles in *regulations*, *reimbursement systems*, *caremodel sustainability*, *stakeholder involvement*, *lack of evidence*, *contracting strategies*, and *political constraints*. Significant technical barriers at the system level include, among others, lack of standardization and data security concerns.

The enablers that a CH system employs toward achieving impact can be reported at the same levels as the barriers, that is, *individual* and *system* (categorized as *organizational* or *technical*) level, and span beyond merely overcoming technical challenges to address the problems identified at these levels and contribute to their solution. At the individual level, it is essential to *educate and motivate* consumers, potentially through incentives, for example, via offering a clear user-perceived health-related benefit.

At the system level, both organizational and technical issues need to be addressed. From the organizational side, issues within the health care organization (eg, *cost-saving strategy*, *integration of services*, and *guideline support*) need to be identified, and facilitating factors should be clearly defined. A vital organizational enabler is *contracting strategies*. In the developed countries, health care is delivered by the state, and most of the medical services are bought by a paying organization (payer) on behalf of patients or consumers as a third party of the transaction. This method shows that end users are not a party in the contract and are not directly interested in cost savings. Therefore, contracting strategies, as a way in which medical services are reimbursed, have to incentivize providers to implement innovative CH solutions. Payers already use many types of contracts that promote a different kind of provider behavior, for example, capitation, pay for performance, and case-mix contracts (diagnosis-related group, health regulation division, and shared saving models). On the technical side, best practices such as platform independence of applications, data integration and interchange, privacy awareness [33] are straightforward CH enablers.

Figure 4. Classification of enablers (upward arrows) and barriers (downward arrows) of connected health impact as identified by the connected health impact framework; Individual (left) and system-level (right) enablers and barriers have been identified.



Although proactively working to leverage enablers and overcome barriers, based on previous experience, the barriers and enablers are not in a static relation to the CH system. Besides, their relationship is not readily observable and quantified, that is, the power of the association, the extent to which they affect each other, and the outcome at a user level and beyond.

Connected Health Impact Framework Preliminary Evaluation

To examine how understandable and usable the CH concepts and terms of the proposed framework and CHIF in general are, we developed an *electronic questionnaire assessment*, which followed a semistructured form [34]. The questionnaire, which also encouraged insights entered in free text, was introduced to early-stage researchers that work in the domain of CH in multiple disciplines (information technology, business, and health), in the scope of the ENJECT summer school (London, September 2017). A screen capture of the electronic questionnaire is provided in Multimedia Appendix 2. The questions included in the questionnaire are provided in Multimedia Appendix 3. The actual questionnaire was eventually completed anonymously by 5 volunteers from this group. The summarized answers are presented in Multimedia Appendix 4.

Although we received a limited number of responses, it helped in observing how the CHIF concepts were perceived. Analyzing the answers of the assessment, we noticed a pattern—most of the participants tended to fill in specific parts of the questionnaire, whereas other parts were consistently left without an answer. We assume that the answered questions represent the concepts of the CHIF that are understood and accepted, and the questions that remained unanswered suggest the concepts and terms of the framework that were not clear enough to understand. The answered questions were associated with the following concepts: CH User, CH outcomes toward impact, and partially the means to achieve impact. The parts presenting gaps in the answers included the primary goal of the framework, information on time scales of the different effects, industrial and business growth dimensions, and barriers and enablers beyond the patient or consumer user (system and secondary users). The value proposition section was left unanswered. Open questions were mostly unanswered.

Although preliminary, this limited assessment indicates the directions for improving the framework (especially the fuzzy areas) and shows the need for a better understanding and contemplating around the concept of CH services and their impact, beyond a solely technological perspective.

Discussion

Principal Findings

This paper provides an overview of CH impact concepts and proposes the CHIF for the consistent description and assessment of CH services' impact in its different dimensions. Although this framework may benefit from further refinement, it is an attempt for setting the basis for complete and consistent reporting of this rather vague area, and it is expected to contribute in better evidence building and better designing of CH services.

A series of steps is foreseen that will lead to the CHIF deployment and use. The development of CHIF ontology is necessary for knowledge standardization and interoperability that lies in the very heart of CH. Following the example of mobile health (mHealth)–reporting guidelines [35], CHIF has the potential to evolve as a tool for reporting CH impact, either as a checklist or as an eQuestionnaire. This will also be useful for comparing interventions. In this regard, it can be part of a broader CH framework and form the basis for a digital registry of CH interventions to be further studied and compared.

An important step before that is to proceed with a thorough evaluation of CHIF-based tools in terms of clarity, completeness, and redundancy. The existing questionnaire will be the basis for that. A future evaluation will include compare between free text and structured entry of CH impact information and a post reporting questionnaire for user experience.

The following subsections discuss different aspects regarding CH impact that present challenges and could benefit from further investigation.

The Multiple Dimensions and Scales of Connected Health Impact

We see CH impact as a multilevel concept, where potentially some CH outcomes at one level can influence those at other levels. The above interaction indicates that outcomes can also include causative relations. In essence, this requires approaching the concept of impact in a different manner: moving beyond the static clinical aggregate key performance indicators, toward linking outcomes in a more dynamic way (ie, personal, health care, and socioeconomic as well as horizontal aspects of the outcomes). Such approaches can be envisioned within a big data framework that can potentially reshape health policies. Among others, it could help to better elaborate on how each type of health outcome is linked with potential care benefit and cost reduction.

The concepts and directions stated above can set the basis for a CH taxonomy. The taxonomy in its turn can support consistent reporting, evidence building, and systematic reviewing purposes. Similar approaches have been used in consumer health [36], integrated care, and behavioral informatics applications [13].

One of the most crucial and challenging issues is the multiscale character of the CH impact. The impact can be at different structural (eg, micro-, meso-, or macroscale) and temporal scales. For example, a CH system can provide a more efficient cost-benefit ratio in the midterm to long term because of the initial increase of the cost of service and even more increased benefits and savings in the long term, owing to the abundance of information and knowledge produced by CH services.

Connected Health Beyond Consumer Health Applications

CH extends beyond primary and secondary health care settings to the whole daily life, and therefore, inevitably uses technology that is beyond medical devices as established in clinical care. The *person-centered care* approach is well-suited to utilize consumer health technologies [37]. The role of consumer health electronics and systems in daily life has been recently recognized and appreciated. As described in the study by Gibbons et al [16], *consumer health informatics applications or tools* were defined as any electronic tool, technology, or system, which is in accordance with the following: (1) primarily designed to interact with health information users or consumers (ie, anyone who seeks or uses health care information for nonprofessional work), (2) interacts directly with the consumer who provides personal health information to the system and receives personalized health information from the tool, application, or system, and (3) the data, information, recommendations, or other benefits provided to the consumer may be used in coordination with an HCP but is not dependent on an HCP. In this regard, patients (individuals who have entered in the health care process) are distinguished from citizens and consumers.

By repurposing or extending their initial aims, such tools can be used and have already been used (eg, smartwatches and activity trackers) for (1) disease management to facilitate knowing, tracking, or understanding clinical parameters, (2) monitoring and understanding daily living observations (quantified-self perspective), (3) lifestyle management assistance (eg, calendar, reminder), (4) prevention and health promotion, (5) self-care, and (6) assisted care and caregiving. When considering the quantified self, socialization, or patient–health professional relationship domain, few published studies have investigated the determinants of the efficacy of these smart connected devices and their impact on individual behaviors and professional health practices [38].

A valid point for disambiguation is whether *CH basically is driven by consumer health electronics and applications*, a point extensively discussed in ENJECT workshops. The answers and views seem contradictory. From one side, it is believed that CH impact is mainly driven by consumer electronics in daily life, in other words, CH impact heavily relies on consumer health informatics. The reasoning behind this is that information and communication technologies and consumer electronics indeed influence and facilitate different aspects of everyday life and societal needs, including health. The culture of permanent self-monitoring (quantified self) is a typical case of this transformational power.

There is also an opinion about consumer electronics as a partial or moderate contributor to the broader impact of CH. The main arguments for the partial contribution are the lack of de-facto integration of consumer electronics data with medical data and also the lack of actionability at the medical level (professional interpretation of the data).

Connected Health Impact Beyond Electronic and Mobile Health

Electronic health (eHealth) has been the generic platform for organizing and delivering digital health content and electronic and remote care services, whereas mHealth contributed to the wealth of mobile services focusing on the patient, the elderly, and the continuity of chronic care. The added value that CH can bring to the previous efforts in the eHealth and mHealth domain seems to span across 3 axes: (1) data and service integration and interaction, (2) validation of health-related services, and (3) overall health.

Data and Service Integration and Interaction

Traditionally, the technological framework for standardization and interoperability has been built within eHealth (eg, the data exchange standard health level 7 (HL7) [39]. However, from a functional perspective, the integration and interaction between personal and clinical information in a continuum, instead of overlapping eHealth, mHealth, and telemedicine, is a central point in CH. To a certain extent, this interaction can be regarded as a transfer of evidence from self-management data to the clinical treatment of patients and vice versa. Although this can now be technically leveraged by HL7 FHIR (Fast Health Care Interoperability Resources) and similar technology and standards

[40], neither the organizational capability that is required on the health care side nor the scientific evidence on the use of such resources is entirely evident.

Validation of Health-Related Services

This can be regarded as a secondary outcome of data and service integration and interaction, supporting CH evidence formation.

Health

The CH services have the potential to contribute to the improvement in the diagnostic process (eg, shorter time to diagnosis), wellness, and evidence of self-management. Telemedicine services for the elderly and patients with chronic diseases and those targeting accessibility to health care services (eg, people with disabilities or rural area residents) have been recognized and adopted to some extent. Other aspects, including patient and consumer empowerment, treatment adherence, prevention of behaviors contributing to health-related risks, and health literacy, are candidates future research and development targets toward achieving impact.

Overall, CH impact beyond eHealth and mHealth should focus on integration and access to a wealth of information and services. Therefore, there is a need for the explicit descriptions of services and data that will be linked and integrated, from both, technical and organizational perspective.

Which Future Research Activities Can Facilitate Connected Health Impact?

CH is a new promising direction for improved health and well-being services [2]. Therefore, further research and investigations should concentrate on how CH can be interwoven into other important initiatives leading to cost containment and improvement of care.

Person-centered care and health promotion are both vital fields where CH tools are potentially able to prove their usefulness. CH in person-centered health care systems can support patients or consumers to cope with the health and well-being problems using their own resources, and as needed, help make informed decisions on when to invite others, including professionals, to act on their behalf. In this approach, well-designed CH tools may be able to prolong the period when patients and consumers would be capable to successfully manage their health and care according to their lifestyle, preferences, and goals. Patient-centered design and patients’ and consumers’ data analytics are the essential methods under the theoretical foundations of health behavior informatics.

This direction of the CH development needs studies to investigate what kind of contracting strategies and incentives could facilitate implementations of CH tools that enable cost containment by keeping people longer out of health care facilities or providers. The integration of CH services with new promising cost containment and quality improvement policies should be a research priority, and new business models should be designed. Field studies should be promoted to collect evidence and understand needs. Health economics and finance should be revised based on new political guidance.

CH technologies can be employed for adapting public health policies, addressing a broader health-related impact, which also involves transitions to new models of care. The availability of CH data combined with *big data analytics* can be of added value toward supporting the learning health system cycle [41].

Besides the cost and business perspective, it is essential to recognize the role and the rising needs for CH education, entailing for interprofessional aspects. CH education is related to preparing the stakeholders and addressing barriers and concerns as well as contextual factors. Elaboration of new curricula for HCPs and health researchers, while addressing CH literacy for citizens in an organized and inspiring manner, could have a transformative power toward CH impact.

Limitations

A limitation of this study is the lack of extensive evaluation of the proposed CHIF framework. In addition, the lack of standard terminology may pose challenges toward extended use of the framework for comparison and new knowledge elicitation. The addition of formal descriptions and semantics and the link to standard terminologies is considered a necessary next step. The adoption of standards and semantics is expected to alleviate some of the possible difficulties and ambiguities related to the current implementation and lead to broader use and evaluation of the framework.

Conclusions

CH technologies offer new vehicles for implementing *anytime and anywhere* health and care services. Being an emerging and diverse field, CH will benefit from the disambiguation of concepts. In addition, scaling up of these services is closely related to a means for understanding and measuring their impact. In this regard, this study introduces CHIF, a framework for CH impact assessment that contributes to the formalization of the CH domain, also paving the way toward the introduction of methods for measuring and comparison in multiple scales and dimensions related to CH outcomes. CHIF can evolve toward the creation of a CH impact tool and contribute to the generation of a service registry for further comparison and investigation.

Acknowledgments

This article/publication is based upon work from COST Action ENJECT TD 1405, supported by COST (European Cooperation in Science and Technology; www.cost.eu). The participants of the 2 ENJECT working group 1 workshops, in Munich and Zagreb, are especially acknowledged. The contribution of RL was supported by funding from National Science Centre, Poland (Grant Number: 2015/17/B/HS4/02747).

The ENJECT Working Group 1 Network: collaborators included, besides the auhors of this work, the following scientists: Anne Moen, Nikolaos Maglaveras, Helen Schonenberg, Violeta Gaveikaite, Cecilia Lee, Brian Caulfield, Nicola Mountford, Liepa

Bikulciene, Igor Todorovic, Estefania Guisado Fernandez, George Dimitrakopoulos, Marcin Kautsch, Tatjana Loncar Turukalo, Maritta Perala-Heape, Luis Fernandez Luque.

Authors' Contributions

Special thanks to native English speaker, KL, collaborator at Aristotle University of Thessaliniki, for proofreading the paper.
IC organized the research and was the main editor of the paper. CM and KL contributed to the visualizations and ontological formalization and framework consistency. KL, VT, SB, HB, JK, and RL contributed to the refinement of the framework. All coauthors contributed to editing and revision of the paper. The ENJECT Network collaborators contributed as workshop participants and supporters of the study, overall.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Workshops description.
[DOCX File, 15KB - jmir_v21i9e14005_app1.docx]

Multimedia Appendix 2
Screen captures of the electronic questionnaire based on the connected health impact framework (CHIF).
[DOCX File, 392KB - jmir_v21i9e14005_app2.docx]

Multimedia Appendix 3
The electronic questionnaire.
[PDF File (Adobe PDF File), 684KB - jmir_v21i9e14005_app3.pdf]

Multimedia Appendix 4
The electronic questionnaire summarized answers.
[PDF File (Adobe PDF File), 1MB - jmir_v21i9e14005_app4.pdf]

References

1. Mountford N, Kessie T, Quinlan M, Maher R, Smolders R, van Royen P, et al. Connected Health in Europe: Where are We Today?. Dublin: University College Dublin; 2016.
2. Caulfield BM, Donnelly SC. What is connected health and why will it change your practice? QJM 2013 Aug;106(8):703-707. [doi: 10.1093/qjmed/hct114] [Medline: 23676416]
3. Chouvarda IG, Goulis DG, Lambrinoudaki I, Maglaveras N. Connected health and integrated care: toward new models for chronic disease management. Maturitas 2015 Sep;82(1):22-27. [doi: 10.1016/j.maturitas.2015.03.015] [Medline: 25891502]
4. Burmaoglu S, Saritas O, Kıdak LB, Berber İ. Evolution of connected health: a network perspective. Scientometrics 2017 Jun 15;112(3):1419-1438. [doi: 10.1007/s11192-017-2431-x]
5. Ekman I, Swedberg K, Taft C, Lindseth A, Norberg A, Brink E, et al. Person-centered care--ready for prime time. Eur J Cardiovasc Nurs 2011 Dec;10(4):248-251. [doi: 10.1016/j.ejcnurse.2011.06.008] [Medline: 21764386]
6. European Network for the Joint Evaluation of Connected Health Technologies (ENJECT). 2015. URL: http://enject.eu/ [accessed 2019-05-18] [WebCite Cache ID 78SUIoxSw]
7. Health Education Partners. 1999. Health Impact Assessment: Main Concepts and Suggested Approach (Gothenburg Consensus Paper) URL: http://www.healthedpartners.org/ceu/hia/hia01/01_02_gothenburg_paper_on_hia_1999.pdf [accessed 2019-05-18] [WebCite Cache ID 78SfNom71]
8. World Health Organization. 2015. Health Impact Assessment URL: http://www.who.int/topics/health_impact_assessment/en/ [accessed 2019-05-18] [WebCite Cache ID 78SfOBaPp]
9. Lluch M, Abadie F. Exploring the role of ICT in the provision of integrated care--evidence from eight countries. Health Policy 2013 Jun;111(1):1-13. [doi: 10.1016/j.healthpol.2013.03.005] [Medline: 23587547]
10. European Commission. 2016. A Maturity Model for Adoption of Integrated Care within Regional Healthcare Systems (B3 Action Group) URL: https://ec.europa.eu/eip/ageing/repository/maturity-model-adoption-integrated-care-within-regional-healthcare-systems-b3-action_en [accessed 2019-05-18] [WebCite Cache ID 78SfPRHGh]
11. SCIROCCO Project. 2016. URL: http://www.scirocco-project.eu [accessed 2019-05-18] [WebCite Cache ID 78SUctDmr]
12. Kidholm K, Ekeland AG, Jensen LK, Rasmussen J, Pedersen CD, Bowes A, et al. A model for assessment of telemedicine applications: MAST. Int J Technol Assess Health Care 2012 Jan;28(1):44-51. [doi: 10.1017/S0266462311000638] [Medline: 22617736]

13. Maramis C, Chouvarda I, Maglaveras N, Isomursu M. Introducing a Framework for Reporting Behavioral Informatics Interventions. In: Proceedings of the 14th International Conference on Telecommunications. 2017 Presented at: ConTEL'17; June 28-30, 2017; Zagreb, Croatia p. 63-68. [doi: 10.23919/ConTEL.2017.8000040]
14. O'Leary P, Carroll N, Clarke P, Richardson I. Untangling the Complexity of Connected Health Evaluations. In: Proceedings of the International Conference on Healthcare Informatics. 2015 Presented at: CHI'15; October 21-23, 2015; Dallas, Texas, US p. 272-281. [doi: 10.1109/ICHI.2015.39]
15. Carroll N, Travers M, Richardson I. Evaluating Multiple Perspectives of a Connected Health Ecosystem. In: Proceedings of the 9th International Conference on Health Informatics. 2016 Presented at: HEALTHINF'16; February 21-23, 2016; Rome, Italy p. 17-27. [doi: 10.5220/0005623300170027]
16. Gibbons MC, Wilson RF, Samal L, Lehmann CU, Dickersin K, Lehmann HP, et al. Consumer health informatics: results of a systematic evidence review and evidence based recommendations. Transl Behav Med 2011 Mar;1(1):72-82 [FREE Full text] [doi: 10.1007/s13142-011-0016-4] [Medline: 24073033]
17. Colorafi K. Connected health: a review of the literature. Mhealth 2016;2:13 [FREE Full text] [doi: 10.21037/mhealth.2016.03.09] [Medline: 28293591]
18. Ryan P, Sawin KJ. The individual and family self-management theory: background and perspectives on context, process, and outcomes. Nurs Outlook 2009;57(4):217-25.e6 [FREE Full text] [doi: 10.1016/j.outlook.2008.10.004] [Medline: 19631064]
19. O'Donoghue G, Doody C, Cusack T. Using student centred evaluation for curriculum enhancement: an examination of undergraduate physiotherapy education in relation to physical activity and exercise prescription. Stud Educ Eval 2011 Jun;37(2-3):170-176. [doi: 10.1016/j.stueduc.2011.04.004]
20. Livitckaia K, Koutkias V, Maglaveras N, Kouidi E, van Gils M, Chouvarda I. Overview of Health Behavior Change Interventions to Promote Physical-Activity-Related Adherence in Patients with Heart Disease. In: Proceedings of the European Medical and Biological Engineering Conference. 2017 Presented at: EMBEC & NBC'17; June 11-15, 2017; Tampere, Finland p. 286-289. [doi: 10.1007/978-981-10-5122-7_72]
21. Anglada-Martinez H, Riu-Viladoms G, Martin-Conde M, Rovira-Illamola M, Sotoca-Momblona JM, Codina-Jane C. Does mhealth increase adherence to medication? Results of a systematic review. Int J Clin Pract 2015 Jan;69(1):9-32. [doi: 10.1111/ijcp.12582] [Medline: 25472682]
22. 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]
23. Argent R, Daly A, Caulfield B. Patient involvement with home-based exercise programs: can connected health interventions influence adherence? JMIR Mhealth Uhealth 2018 Mar 1;6(3):e47 [FREE Full text] [doi: 10.2196/mhealth.8518] [Medline: 29496655]
24. Hoffman DA. Enhancing self-efficacy for optimized patient outcomes through the theory of symptom self-management. Cancer Nurs 2013;36(1):E16-E26 [FREE Full text] [doi: 10.1097/NCC.0b013e31824a730a] [Medline: 22495550]
25. Weinstein RS, Lopez AM. Health literacy and connected health. Health Aff (Millwood) 2014 Jun;33(6):1103-1104. [doi: 10.1377/hlthaff.2014.0396] [Medline: 24889963]
26. Ford AR, Gibbons CM, Torres J, Kornmehl HA, Singh S, Young PM, et al. Access to dermatological care with an innovative online model for psoriasis management: results from a randomized controlled trial. Telemed J E Health 2019 Jul;25(7):619-627 [FREE Full text] [doi: 10.1089/tmj.2018.0160] [Medline: 30222518]
27. Rothman BS, Gupta RK, McEvoy MD. Mobile technology in the perioperative arena: rapid evolution and future disruption. Anesth Analg 2017 Mar;124(3):807-818. [doi: 10.1213/ANE.0000000000001858] [Medline: 28151816]
28. Weenink JW, Braspenning J, Wensing M. Patient reported outcome measures (PROMs) in primary care: an observational pilot study of seven generic instruments. BMC Fam Pract 2014 May 6;15:88 [FREE Full text] [doi: 10.1186/1471-2296-15-88] [Medline: 24884544]
29. Carroll N, Richardson I. Mapping a careflow network to assess the connectedness of connected health. Health Informatics J 2019 Mar;25(1):106-125. [doi: 10.1177/1460458217702943] [Medline: 28438102]
30. Poikola A, Kuikkaniemi K, Honko H. Valto Etusivu - Valtioneuvosto. 2014. MyData – A Nordic Model for Human-Centered Personal Data Management and Processing URL: http://julkaisut.valtioneuvosto.fi/handle/10024/78439 [accessed 2019-05-18] [WebCite Cache ID 78SfSYoRy]
31. BigO. 2017. Big Data Against Childhood Obesity URL: https://bigoprogram.eu/ [accessed 2019-05-08] [WebCite Cache ID 78D4kZTkI]
32. Rosenberg R, Vanlare J, Reinholt B, Rao S, Dertouzos J. McKinsey & Company. 2015. Capturing Value From Connected Health URL: https://www.mckinsey.com/~/media/McKinsey/Industries/Pharmaceuticals%20and%20Medical%20Products/Our%20Insights/Capturing%20value%20from%20connected%20health/Capturing%20value%20from%20connected%20health.ashx [accessed 2019-05-18] [WebCite Cache ID 78SfUqpxK]
33. Allaert FA, Mazen NJ, Legrand L, Quantin C. The tidal waves of connected health devices with healthcare applications: consequences on privacy and care management in European healthcare systems. BMC Med Inform Decis Mak 2017 Jan 17;17(1):10 [FREE Full text] [doi: 10.1186/s12911-017-0408-6] [Medline: 28095843]

34. Chouvarda I. Google Docs. 2017. Connected Health Impact Reporting URL: https://goo.gl/forms/jq99GeloLpZluFB62 [accessed 2019-05-18] [WebCite Cache ID 78SfVv4Dl]
35. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mhealth) evidence reporting and assessment (mERA) checklist. Br Med J 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
36. Agency for Healthcare Research and Quality. 2009. Impact of Consumer Health Informatics Applications URL: https://www.ahrq.gov/downloads/pub/evidence/pdf/chiapp/impactchia.pdfhttps://www.ahrq.gov/downloads/pub/evidence/pdf/chiapp/impactchia.pdf [accessed 2019-05-18] [WebCite Cache ID 78SfWxXSS]
37. Ekman I, Busse R, van Ginneken E, van Hoof C, van Ittersum L, Klink A, et al. Health-care improvements in a financially constrained environment. Lancet 2016 Feb 13;387(10019):646-647. [doi: 10.1016/S0140-6736(16)00285-3] [Medline: 26876711]
38. Petit A, Cambon L. Exploratory study of the implications of research on the use of smart connected devices for prevention: a scoping review. BMC Public Health 2016 Jul 11;16(1):552. [doi: 10.1186/s12889-016-3225-4] [Medline: 27401769]
39. Health Level 7 International. 2007. URL: http://www.hl7.org/ [accessed 2019-05-18] [WebCite Cache ID 78SfYI8sJ]
40. Beredimas N, Kilintzis V, Chouvarda I, Maglaveras N. A Reusable Ontology for Primitive and Complex HL7 FHIR Data Types. In: Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2015 Presented at: EMBC'15; August 25-29, 2015; Milan, Italy p. 2547-2550. [doi: 10.1109/EMBC.2015.7318911]
41. The Learning Healthcare Project. 2015. Background: Learning Healthcare System URL: http://www.learninghealthcareproject.org/section/background/learning-healthcare-system [accessed 2019-05-18] [WebCite Cache ID 78Sen1w0k]

Abbreviations

CH: connected health
CHIF: connected health impact framework
COST: Cooperation in Science and Technology
eHealth: electronic health
ENJECT: European Network for the Joint Evaluation of Connected Health Technologies
HCP: health care professional
HIA: health impact assessment
HL7: health level 7
MAST: Model for Assessment of Telemedicine
mHealth: mobile health

Edited by B Caulfield; submitted 14.03.19; peer-reviewed by D Kriksciuniene, R Argent; comments to author 27.04.19; revised version received 22.05.19; accepted 24.05.19; published 03.09.19.

Please cite as:
Chouvarda I, Maramis C, Livitckaia K, Trajkovik V, Burmaoglu S, Belani H, Kool J, Lewandowski R, The ENJECT Working Group 1 Network
Connected Health Services: Framework for an Impact Assessment
J Med Internet Res 2019;21(9):e14005
URL: https://www.jmir.org/2019/9/e14005/
doi:10.2196/14005
PMID:31482857

Review

The Extent and Coverage of Current Knowledge of Connected Health: Systematic Mapping Study

Maria Karampela[1], MSc; Minna Isomursu[2], PhD; Talya Porat[3], PhD; Christos Maramis[4], PhD; Nicola Mountford[5], PhD; Guido Giunti[2], MD, PhD; Ioanna Chouvarda[4], PhD; Fedor Lehocki[6], PhD

[1]IT University of Copenhagen, Copenhagen S, Denmark
[2]University of Oulu, Oulu, Finland
[3]Imperial College London, London, United Kingdom
[4]Aristotle University of Thessaloniki, Thessaloniki, Greece
[5]Maynooth University, Maynooth, Ireland
[6]Slovak University of Technology, Bratislava, Slovakia

Corresponding Author:
Maria Karampela, MSc
IT University of Copenhagen
Rued Langgaards Vej 7
Copenhagen S, 2300
Denmark
Phone: 45 25540984
Email: makar@itu.dk

Abstract

Background: This study examines the development of the connected health (CH) research landscape with a view to providing an overview of the existing CH research. The research field of CH has experienced rapid growth coinciding with increasing pressure on health care systems to become more proactive and patient centered.

Objective: This study aimed to assess the extent and coverage of the current body of knowledge in CH. In doing so, we sought to identify specific topics that have drawn the attention of CH researchers and to identify research gaps, in particular those offering opportunities for further interdisciplinary research.

Methods: A systematic mapping study that combined scientific contributions from research in the disciplines of medicine, business, computer science, and engineering was used. Overall, seven classification criteria were used to analyze the papers, including publication source, publication year, research type, empirical type, contribution type, research topic, and the medical condition studied.

Results: The search resulted in 208 papers that were analyzed by a multidisciplinary group of researchers. The results indicated a slow start for CH research but showed a more recent steady upswing since 2013. The majority of papers proposed health care solutions (77/208, 37.0%) or evaluated CH approaches (49/208, 23.5%). Case studies (59/208, 28.3%) and experiments (55/208, 26.4%) were the most popular forms of scientific validation used. Diabetes, cancer, multiple sclerosis, and heart conditions were among the most prevalent medical conditions studied.

Conclusions: We conclude that CH research has become an established field of research that has grown over the last five years. The results of this study indicate a focus on technology-driven research with a strong contribution from medicine, whereas the business aspects of CH have received less research attention.

(J Med Internet Res 2019;21(9):e14394) doi:10.2196/14394

KEYWORDS

connected health; health services research; interdisciplinary research; empirical research; telemedicine; information technology; wireless technology; health informatics; information systems

Introduction

Background

A variety of terms and concepts exist that describe the use of technology in health including *Health IT*, *eHealth*, *Telemedicine*, and *Health Informatics*. In the early 2000s [1], the term *Connected Health* (CH) began to appear. It did so in the context of research that investigated how information and communication technologies could advance health by connecting people, knowledge, technical artifacts, and organizations. Some years later, Poon and Zhang introduced a new CH information system based on a “four-layered architecture: personal, home, community, and hospital” [2]. One of the most cited definitions for CH was introduced in 2013 [3]:

> *Connected Health encompasses terms such as wireless, digital, electronic, mobile, and tele-health and refers to a conceptual model for health management where devices, services or interventions are designed around the patient’s needs, and health related data is shared, in such a way that the patient can receive care in the most proactive and efficient manner possible. All stakeholders in the process are connected by means of timely sharing and presentation of accurate and pertinent information regarding patient status through smarter use of data, devices, communication platforms, and people.*

The European Network for the Joint Evaluation of Connected Health Technologies defines the CH vision as “a paradigm shift looking after the individual and community health in a process that speaks to the health journey of the person, through the entire lifespan, leveraging a variety of technologies to do so” [4]. Achieving this vision will require attention to be paid to policy and regulation, technology and interoperability, training and education, business and revenue models, as well as citizen and clinician engagement [5]. Successful implementation of the CH vision needs time and effort from all health care stakeholders.

Objectives

Given the recent growth in CH research, we set out to map its evolution up to the present day. We use a systematic mapping study to chart the research landscape combining scientific contributions from the research disciplines of medicine, business, computer science, and engineering. Although previous CH reviews do exist, they are confined to descriptions of how CH solutions are being used in specific conditions such as cancer [6,7] or in specific measurement technologies such as measurement of vital signs [8] or weight [9]. Past efforts also concentrated on presenting CH-related literature in systematic ways [6,10,11]. Our goal is to provide a comprehensive, interdisciplinary overview of existing CH research. This will help researchers understand how the field has developed since its earliest studies in the late 1990s. It will also identify those topics that have drawn the attention of the research community. This knowledge will add value by identifying gaps or interdisciplinary opportunities in the study of CH. Perhaps most importantly, it may also underpin future work to develop an integrated and interdisciplinary research agenda for CH that will answer efficacy, design, policy, and sustainability questions for patients, clinicians, technology developers, and businesses.

Methods

Overview

This paper followed a systematic mapping study method [12]. Systematic mapping studies aim primarily to present an overview of a research area to report the quantity and type of literature and results that are published within it. The systematic mapping process comprises 3 steps: (1) the identification of relevant literature, (2) the composition of a classification scheme, and (3) the mapping of literature [12]. Figure 1 presents the mapping process including the search for relevant literature, the definition of a scheme, and the mapping of relevant publications.

The method was used to examine the body of existing research conducted by researchers in medicine, business, computer science, and engineering to understand the nature of research conducted in the area of CH. A systematic mapping study was found to be suitable for this task as it provided a high-level framework for combining interdisciplinary research efforts as well as an analytical framework that spanned disciplinary boundaries.

Mapping Questions

The aim of this study was to present an overview of the available publications pertinent to CH. Following the systematic mapping study method, the study was guided by a set of mapping questions. Table 1 presents the 6 mapping questions (MQs) and the rationale for conducting this study. More detail as to the logic supporting the selection of suitable MQs is included in the Data Extraction Strategy section below. The study search strategy as well as the inclusion and exclusion criteria (EC) were based on these 6 MQs.

Figure 1. Systematic mapping process.

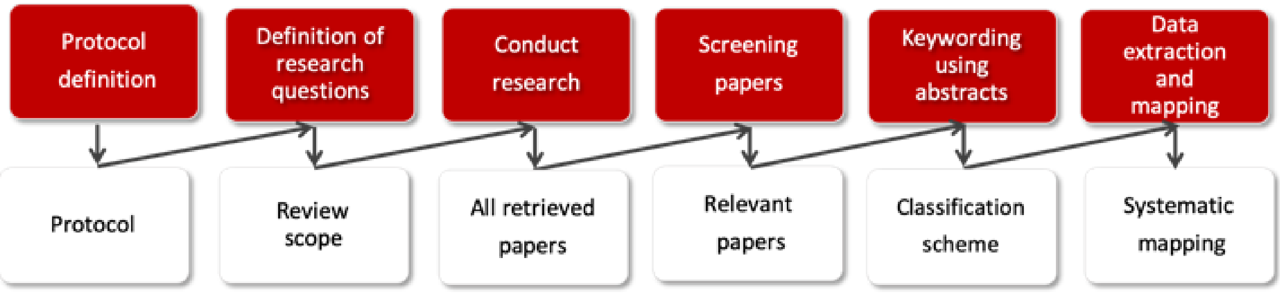


Table 1. Mapping questions.

ID	Questions	Rationale
MQ[a]1	Which publication channels are the main targets for CH[b] research?	To identify where CH research can be found and to identify targets for publication of future studies
MQ2	How has the frequency of studies related to CH changed over time?	To identify the publication trends over time of CH literature
MQ3	What are the research types of CH studies?	To explore the different types of research reported in the literature concerning CH
MQ4	Are CH studies empirically validated?	To discover whether research on CH has been validated through empirical studies
MQ5	What are the approaches that were reported in CH research?	To discover the CH approaches reported in the existing CH literature
MQ6	What are the main topics and conditions in CH literature?	To identify the research areas and health conditions discussed in papers

[a] MQ: mapping question.

[b] CH: connected health.

Search Strategy

As CH is inherently interdisciplinary, our goal was to use the systematic mapping method to study research contributions on the topic across disciplines. We, therefore, searched papers from the most recognized scientific literature databases in each of the chosen disciplines.

The papers of the study were retrieved from 7 databases of scientific literature, namely, (1) Institute of Electrical and Electronics Engineers Xplore Digital Library, (2) Association for Computing Machinery Digital Library, (3) ScienceDirect, (4) SpringerLink, (5) MEDLINE and PubMed, (6) Business Source Complete (EBSCO), and (7) ABI and INFORM Collection (ProQuest), with the help of the corresponding search engines. The search was performed in October 2018. Different search strings were proposed and discussed over the course of joint meetings to arrive at a set of primary keywords. Two of the authors tested different strings of potential keywords such as “Connect Health,” “Connecting Health,” “Connect-Health,” or “Connecting-Health.” After evaluating the search results, the authors agreed to proceed using the following search strings: (“Connected” AND “Health”) OR (“Connected” AND “-” AND “Health”).

The search was applied to the title, abstract, and keywords to include relevant papers. On the basis of our methodology, we included a wide selection of papers on the first iteration and thereafter relied upon the inclusion criteria and EC to identify the relevant literature [13].

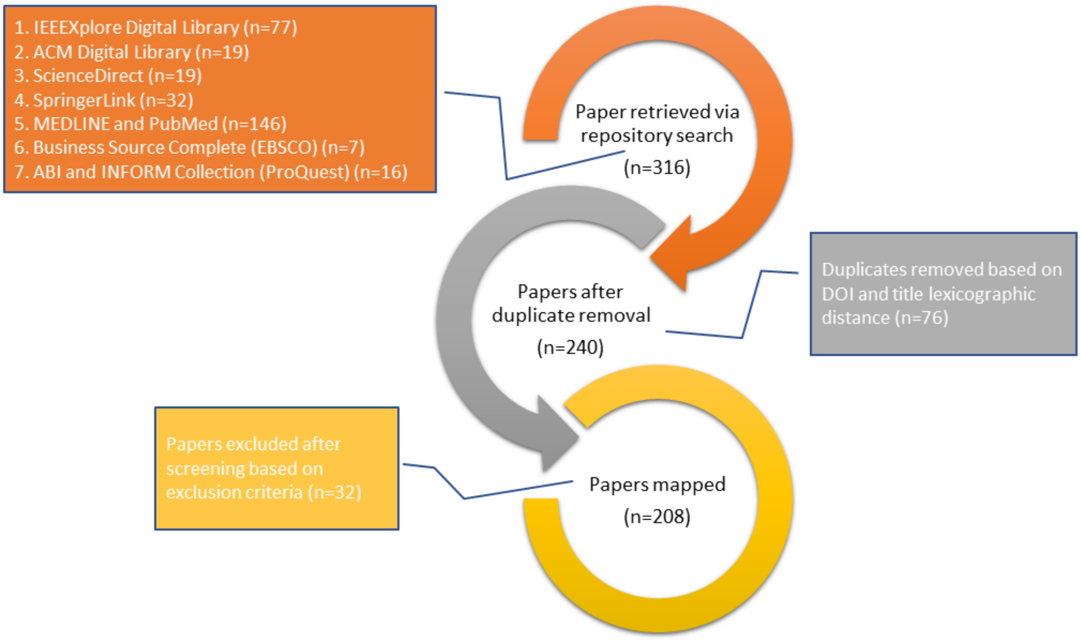
Paper Selection Criteria

The search results that were retrieved from the 7 chosen search engines were merged in a single list and saved in a spreadsheet document. Duplicate entries were removed based on the digital object identifier (DOI) and the intertitle lexical distance (the Levenshtein metric was used to measure the lexical distance) with the help of a custom Python script [14]. A total of 7 papers had to be manually reviewed as potential duplicates.

The final list of papers was distributed to all authors for analysis. On the basis of the area of expertise, pairs of authors were assigned to analyze papers retrieved from each database. A total of 3 areas of expertise were identified as follows: technical (computer science and engineering), medical, and business. Each pair of authors reviewed the title, abstract, and keywords and made a recommendation as to whether that paper should be included or excluded. Discrepancies across coding teams were resolved through further scrutiny of the paper.

The inclusion criteria were limited to studies that discussed CH. In total, 240 papers were identified after the removal of duplicates, whereas 32 papers were excluded after meeting at least one of the following exclusion criteria: EC1—papers that focus on CH comorbidities, that is, the paper discusses the connection between health conditions, not CH as a concept; EC2—papers that focus on law; EC3—papers that focus on medical procedures only, without connection to CH; and EC4—papers that focus on evaluating climate but not in relation to CH. Figure 2 presents the process of the study selection.

Figure 2. Study selection process. ACM: Association for Computing Machinery; DB: database; DOI: digital object identifier; IEEE: Institute of Electrical and Electronics Engineers.



Data Extraction Strategy

The selected studies were analyzed to collect information that would give answers to the MQs according to the data extraction strategy outlined in Table 2 below.

Table 2. Data extraction strategy.

Mapping questions	Description of classification categories
MQ[a]1	Publication source and publication channel
MQ2	Publication year
MQ3	Research types [15]: evaluation research—real-world CH[b] approaches are implemented and undergo evaluation; solution proposal—a new CH solution or a significant extension of an existing solution is proposed, and the evaluation of the solution is based either on empirical data or theoretical argumentation; opinion paper—a CH study that is based on the personal opinion of the author(s); review—studies that present a review of existing CH literature; other—the remainder of research types associated with CH studies. This is assigned to studies where research type is either unknown or does not fall into one of the aforementioned main categories (eg, experience papers, which express the personal experience of author(s) without providing any scientific evidence to support it)
MQ4	Empirical types [16,11]: case study—an empirical inquiry that investigates a CH approach within its real-life context; survey—an empirical inquiry method for collecting quantitative information concerning a CH approach, for example, a questionnaire; experiment—an empirical method applied under controlled conditions to evaluate a CH approach; history-based evaluation—nonempirical studies evaluating CH approaches in previously completed projects; theory—nonempirical research approaches or theoretical evaluation of a CH approach; other—the remainder of CH studies that do not fit within the previous types
MQ5	Contribution types: method—a manner of procedure and steps taken to acquire knowledge in CH; tool-based technique—a technique based on a software tool to accomplish CH tasks; model—a system representation that allows CH to be investigated through a hierarchical structure; framework—a real or conceptual structure intended to serve as a support or guide for CH; other—the remainder of CH approaches. This includes CH studies of approach not fitting other classes, along with very rare approaches that have been grouped in this category to facilitate abstraction and visualization. The approaches grouped herein are feasibility study, field research, process, guideline, and network analysis
MQ6	Main topics and medical conditions

[a]MQ: mapping question.

[b]CH: connected health.

MQ6 comprised 2 parameters: the topic of the paper and the medical condition examined within it. With regard to the topic of the paper, we did not have a predefined list of topics but relied on an open coding process where the researchers conducting the analysis selected a descriptive word for the topic of the paper. To identify the main topics of the included papers, the authors relied on the title, abstract, and keywords. After compiling all topic words, we curated the topic list to come up with a consistent list (eg, using the same word for *body-worn sensors* and *wearables*). Similarly, with regard to the medical condition examined, no predefined classification was used that allowed coders to assign the medical condition, if any, of each CH paper. Again, the resulting condition list was curated to develop a more consistent list (eg, using the same word for *ageing* and *elderly*).

Synthesis Method

The synthesis method used was based on the following steps: (1) enumerating the number of papers per publication channel and the number of papers per bibliographic source per year; (2) enumerating the primary studies that are classified in each MQ's response; (3) presenting visualizations for the classification results, which have been used in the analysis; and (4) presenting a narrative summary to discuss the principal findings.

Results

Overview

This section describes the results related to the systematic MQs presented in Table 1. Multimedia Appendix 1 gives an overview of the classification results for all the included papers [6-11,17-217]. Custom Python scripts have been developed to process the classification data and generate the tables (Multimedia Appendix 1) and figures (Figures 3 and 4) of this section. The Pandas and the Matplotlib Python [218,219] libraries were used to manipulate the tabular input data and plot the results, respectively.

Figure 3. Publication trend per year−total number of connected health papers published per year; the number of papers reported for the year 2018 only includes papers published until October 2018, with a projection of estimated papers based on linear extrapolation (presented as superimposed gray line).

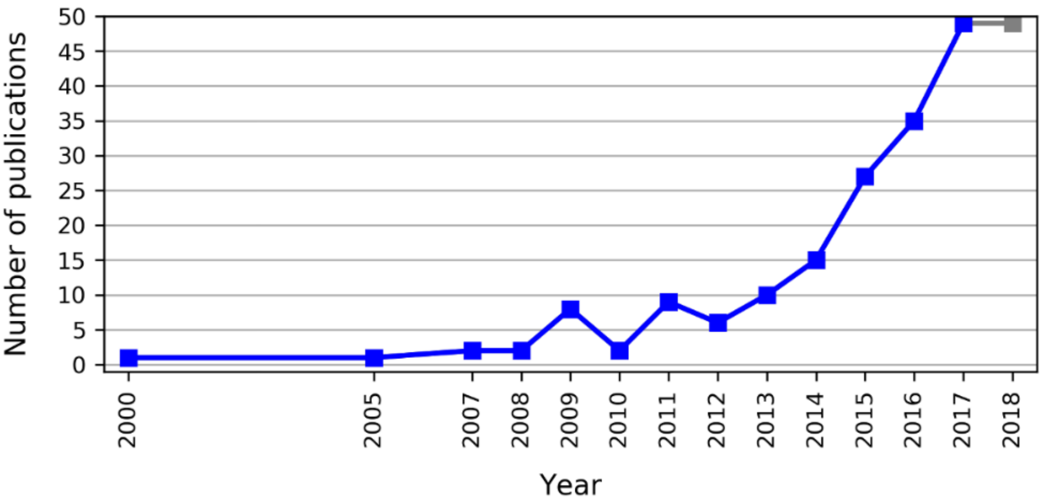
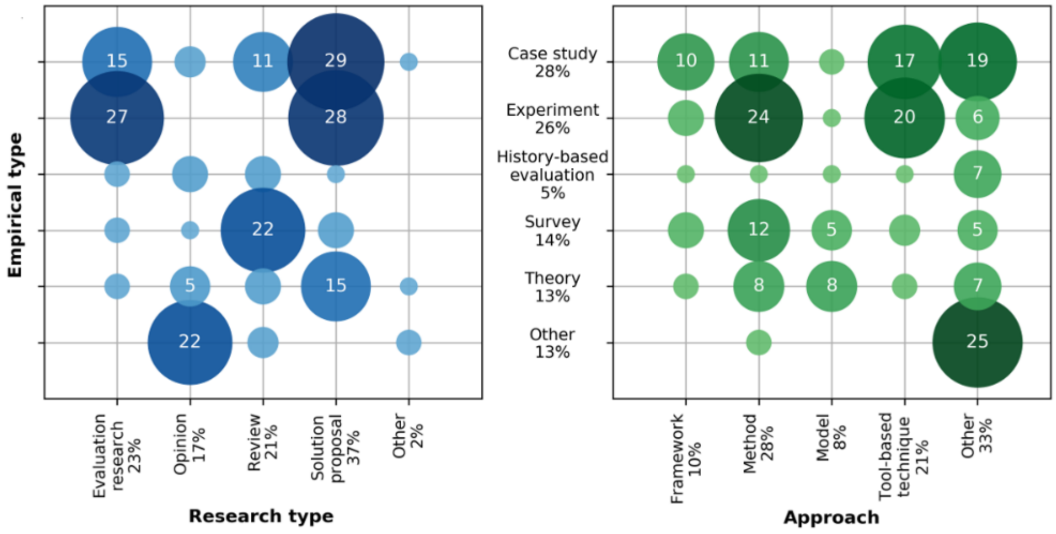


Figure 4. Bubble graphs associating the empirical types with the research types (left) and the approaches (right) of the included connected health studies. The vertical axis (empirical type) is shared between the two graphs. The size and shade of each bubble represents the absolute frequency of connected health papers belonging to a given pair of empirical type and research type (left) or approach (right); absolute frequencies less than 5 are not typed inside the bubble because of space limitations. The horizontal and vertical axes labels are accompanied by the relative frequency (ie, percentage) of the class.



Mapping Question 1: Which Publication Channels are the Main Targets for Connected Health Research?

The majority of the CH papers were published in scientific journals (139/208, 66.8%), whereas 32.9% (68/208) were published in scientific conferences. Table 3 lists the publication forums that have published at least two CH papers.

Mapping Question 2: How Has the Frequency of Studies Related to Connected Health Changed Over Time?

Figure 3 presents the publication trend per year from 2000 to 2018. It must be noted that the analysis does not include the full body of CH work published in 2018 as our search was performed before the end of that year. We have added a projection for year 2018, estimating the total number of papers. Our projection assumes that publication frequency would remain consistent till the year end. As Figure 3 shows, there is a gap in publications from 2001 to 2004. Less than 5 papers per year were published until 2008, whereas from 2013 to 2017, the number of publications rises steadily.

Mapping Question 3: What are the Research Types of Connected Health Studies?

Figure 4 shows the research types of the included studies. The majority of the papers are solution proposal studies (77/208, 37.0%), 23.5% (49/208) of the selected studies evaluated CH approaches, 20.6% (43/208) reviewed literature, 16.8% (35/208) reported the opinions of authors about CH, and 1.9% (4/208) of the papers analyzed were classified as other.

Mapping Question 4: Are Connected Health Studies Empirically Validated?

Figure 4 presents our findings as to whether or not the included studies were empirically validated and, if so, the empirical validation approaches used. The majority of the studies were, in fact, empirically validated. More specifically, 28.3% (59/208) were validated in case studies, 26.4% (55/208) in controlled experiments, and 13.4% (28/208) with surveys. Nonempirical papers included, for example, those focused on theory (25/208, 12.0%) and those using history-based evaluation (11/208, 5.2%).

Mapping Question 5: What Are the Approaches That Were Reported in Connected Health Research?

Figure 4 shows the wide range of approaches taken in the included CH studies. The approaches most frequently reported belonged to other categories (66/208, 31.7%), followed by methods (56/208, 26.9%) and tool-based techniques (43/208, 20.6%). Only 10.0% (21/208) of the papers were classified as presenting frameworks, with 8.1% (17/208) suggesting models.

Mapping Question 6: What Are the Main Topics and Diseases in Connected Health Literature?

Tables 4 and 5 present the results for the main topic and condition of the included CH studies. Topics with fewer than two occurrences have been omitted from Table 4. The most common research topics of the included CH papers were health care (×16) and disease management (×13), followed by telemedicine/telehealth (×11) and electronic health (eHealth)/mobile health (×9). Table 4 lists all research topics and their frequencies, whereas Table 5 lists the conditions and their frequencies. As for the main conditions reported in our study, diabetes (×12), aging (×10), cardiovascular diseases (×10), cancer (×7), chronic diseases (x7), and dementia (x4) were among the prevalent conditions found in the selected studies.

Table 3. Publication sources that have published 2 or more connected health papers.

Publication source	References	Number of published papers
Studies in Health Technology and Informatics	Topaz and Pruinelli (2017) [37], Wu et al (2017) [44], O'Neill et al (2012) [73], Feied et al (2009) [81], Goossen (2017) [82], Singh and Kvedar (2009) [84], Maglaveras et al (2016) [85], Saranto et al (2017) [96], Skiba et al (2016) [101], Kreuzthaler et al (2017) [106], Goossen-Baremans et al (2017) [107], Sermeus et al (2016) [110], Silvello (2018) [206], Tonheim and Babic (2018) [207]	14
Journal of Medical Systems	Wen (2013) [26], Das and Goswami (2013) [27], Santos et al (2016) [28], Chang et al (2013) [29], Vlahu-Gjorgievska et al (2016) [57] Xu and Wu (2015) [59], Lin et al (2016) [65], Xie et al (2014) [80], Kim and Lee (2014) [97]	9
Telemedicine and eHealth[a]	Ford et al (2018) [31], McConnochie et al (2016) [41], Kvedar et al (2009) [49], Aberger et al (2014) [72], Trout et al (2017) [90], Ternullo et al (2013) [99], Kvedar et al (2009) [122]	7
JMIR mHealth and uHealth	Wang et al (2018) [21], El Amrani et al (2017) [51], Dur et al (2018) [62], Wang et al (2018) [64], Harte et al (2017) [77], Argent et al (2018) [100], Sathyanarayana et al (2016) [108]	7
Health Affairs (Millwood)	Frist (2014) [48], Kvedar et al (2014) [52], Iglehart (2014) [55], Weinstein and Lopez (2014) [75]	4
JMIR[b]	Sperrin et al (2016) [9], Gay and Leijdekkers (2015) [39], Loiselle and Ahmed (2017) [89], Agboola et al (2015) [115]	4
QJM[c]: An International Journal of Medicine	Caulfield (2013) [88], Ansary et al (2013) [114], Agboola et al (2013) [116], Caulfield and Donnelly (2013) [3]	4
Journal of Diabetes Science and Technology	Watson et al (2008) [38], Watson et al (2009) [63], Pelletier et al (2011) [78], Helal et al (2009) [109]	4
Journal of the American Academy of Audiology	Saunders and Jacobs (2015) [36], Krupinski (2015) [83], Gladden et al (2015) [112]	3
JMIR Research Protocols	Mountford et al (2018) [32], Wang et al (2018) [58]	2
International Journal of Medical Informatics	Giunti et al (2018) [6], Karampela et al (2018) [11]	2
Maturitas	Chouvarda et al (2015) [17], Stara et al (2018) [68]	2
JMIR Human Factors	Harte et al (2017) [22], Harte et al (2018) [71]	2
Journal of Personalized Medicine	Agboola and Kvedar (2012) [7], Jethwani et al (2010) [35]	2
Revue de l'infirmiere	Warnet (2017) [18], Raymond and Léo (2017) [19]	2
In the 31st International Symposium on computer-based medical systems	Barbosa et al (2018) [138], Barbosa et al (2018) [189]	2
In the 14th International Conference on Telecommunications	Starič et al (2017) [135], Maramis et al (2017) [204]	2
In Proceedings of the International Workshop on Software Engineering in Healthcare Systems	Carroll and Richardson (2016) [172], Abdullah et al (2018) [183]	2
BMC[d] Medical Informatics and Decision Making	Allaert et al (2017) [118], Tharmalingam et al (2016) [119]	2
In the First International Conference on Connected Health: Applications, Systems, and Engineering Technologies	Sinharay et al (2016) [132], Gawanmeh (2016) [194]	2
Journal of Evaluation in Clinical Practice	Barr et al (2012) [50], Barr et al (2014) [103]	2
American Journal of Hospice and Palliative Medicine	Aktas et al (2015) [54], Thomas et al (2017) [111]	2

[a] eHealth: electronic health.

[b] JMIR: Journal of Medical Internet Research.

[c] QJM: Quarterly Journal of Medicine.

[d] BMC: BioMed Central.

Table 4. Frequencies of the main topics associated with the reviewed articles; main topics with a single occurrence have not been included because of space limitations (N=208).

Topics	Frequency (%)
Health care	16 (7.6)
Disease management	13 (6.2)
Telemedicine/telehealth	11 (5.2)
Electronic health/mobile health	9 (4.3)
Monitoring	8 (3.8)
Security	8 (3.8)
Consumer health informatics	7 (3.3)
Sensors	6 (2.8)
Information and communication technologies challenges	5 (2.4)
Personal health devices	4 (1.9)
Education	3 (1.4)
Usability	3 (1.4)
Medical education	3 (1.4)
Innovation	3 (1.4)
User-centered design	3 (1.4)
Interoperability	3 (1.4)
The internet of things	3 (1.4)
Software engineering	2 (0.9)
Privacy	2 (0.9)
Medication adherence	2 (0.9)
Lifestyle coaching	2 (0.9)
Ageing	2 (0.9)
Personalization	2 (0.9)
Service delivery	2 (0.9)
Elderly	2 (0.9)

Table 5. Frequencies of the target conditions associated with the reviewed articles; of note, only 77 of the reviewed articles have been mapped to a target condition (N=208).

Condition	Frequency (%)
No (condition)	131 (62.9)
Diabetes	12 (5.7)
Ageing	10 (4.8)
Cardiovascular diseases	10 (4.8)
Cancer	7 (3.3)
Chronic diseases	7 (3.3)
Dementia	4 (1.9)
Multiple sclerosis	3 (1.4)
Stroke	2 (0.9)
Mental health	2 (0.9)
General health	2 (0.9)
Psychosis	1 (0.4)
Chronic skin disease (psoriasis, dermatology)	1 (0.4)
Vital signs	1 (0.4)
Stress	1 (0.4)
Renal conditions	1 (0.4)
Malaria	1 (0.4)
Hemodialysis	1 (0.4)
Dental issues	1 (0.4)
Arrhythmia	1 (0.4)
Obesity	1 (0.4)
Palliative care	1 (0.4)
Urinary incontinence	1 (0.4)
Epilepsy	1 (0.4)
Rheumatoid arthritis	1 (0.4)
Glaucoma	1 (0.4)
Blood transfusion service	1 (0.4)
Hearing issues	1 (0.4)
Environmental exposure	1 (0.4)

Discussion

This section discusses the results and main findings of this study. First, each mapping question is discussed in its specific subsection. Finally, the limitations of this study are discussed.

Principal Findings

Publication Channels

Publication channels provide information about how research activities in CH have been established in the scientific community. The results of this study show that the majority of CH publications appear in peer-reviewed scientific journals (139/208, 66.8%). Although both journals and conferences aim to disseminate research and contribute to the development of a field, journals are typically considered more prestigious because of manuscript review criteria and acceptance practices. In addition, journals usually present more extended pieces of research work and contribute toward the establishment of a knowledge base for a field. Although the process of publication is longer compared with conferences, journals potentially have a larger impact in terms of visibility and audience reach [220-222]. CH studies are often published in peer-reviewed journals with high impact factors, such as the *Journal of Medical Systems*, *Studies in Health Technology and Informatics*, *Telemedicine and eHealth*, and *Journal of Medical Internet Research*. This indicates that CH research efforts to date have focused on the establishment of a body of knowledge. On the other hand, conferences offer a dynamic environment that enables researchers to communicate with colleagues, exchange research interests, and receive specialized feedback. Conference publications could, therefore, be considered innovation-laden venues that denote the evolution of research in a field, as they often facilitate the presentation of novel ideas based on preliminary results. Conferences appear to be channels for the introduction and establishment of innovative research to experts of a particular area of research [220-222]. The first published conference papers on the topic of CH appeared in 2016. In the following two years, CH research was accepted and published by two international conferences. This finding could indicate efforts to further develop a research community to highlight and nurture CH research. The establishment of the European Network for the Joint Evaluation of Connected Health Technologies research coordination program, the funding of two Innovative Training Network CH projects (cancer: activating technology for connected health [CATCH] and connected health early stage researcher support system [CHESS]), and the expansion of the European Connected Health Alliance organization across Europe have all served to connect stakeholders under the umbrella of CH [223-225]. This coordination of effort has potentially led to the increase in both journal and conference publications as researchers come together around a shared interest in CH.

Publication Trend

The papers matching our inclusion criteria, while spanning a publishing period from 2000 to 2018, show that CH research has attracted increasing attention since 2013. A publication gap between 2001 and 2004 coincides with the infancy of the CH field, likely reflecting the fact that the evolution of CH research was a slow and gradual process underpinned by the fermentation of experts from different academic disciplines. Overall, Figure 3 draws a very typical picture of the evolution of a research field, with sporadic publications in the beginning followed by an exponential increase in the yearly production of scientific literature. One could argue that the discourse around personal health records (PHRs) and patient services started around early 2000s, whereas the focus until that point had tended toward classic telemedicine. The PHR concept, and the shift from *telemedicine* to patient-centered services may have then driven the publication of CH-related papers. The growth of publications after 2009 can be attributed to various factors. The emergence of CH in the past decade coincides with a demographic shift where the older population is ultimately projected to outnumber the young people [226]. This increasing aging population with its chronic and degenerative diseases has been projected to exert severe financial pressure on future health care systems [226,227]. At the same time, the development of new technologies has facilitated the promotion of CH solutions. The proliferation of devices and apps enabled by *internet of things* in health care over the last decade [228], along with the adoption of smartphones and wearables by everyday users, has transformed health care delivery, enabling remote health monitoring and personalization of health care services [17,229,230]. Likewise, providers such as Amazon, Google, Salesforce, International Business Machines corporation, and Microsoft began to establish new data centers for hosting cloud computing apps in 2009 [231]. Taking that into consideration, we could argue that the emergence and evolution of the CH approach from 2009 onward reflects a demand for the provision of CH services to exploit technological advances and bring together patients and stakeholders to “offer the correct information to the correct person at the correct time” and make better decisions for health and care [17]. In terms of eHealth policies and regulations, the trends are consistent with the development of the CH paradigm as, according to the World Health Organization, the number of countries with eHealth and telehealth policies or strategies has started to increase significantly since 2009 [232].

Research Types

Our results show that CH researchers focused primarily on suggesting novel solutions or extending existing research to explain, identify, and provide details of the CH approach (77/208, 37.0%). The 23.5% (49/208) of papers that centered on evaluation also represent attempts to comprehend and develop previous research through evaluating a solution with a valid approach [13,15]. This leads us to conclude that much of CH research is at a development stage, where new concepts are proposed, developed, and evaluated to demonstrate their potential value. However, the existence of a significant number of literature reviews (43/208, 20.6%) indicates that there is maturity in the discipline that allows reviews of the existing research. Through these reviews, researchers aim to identify the research gaps to drive the growth of future research endeavors. If we view this finding alongside the publication trend, which shows that the body of CH research grew over the last few years, then we could argue that there is rapid growth in this discipline. This argument is given weight by the commission of European Union–funded projects such as CHESS in September 2015 and CATCH in 2016 [223,224]. Another factor to consider is that CH as a vision builds upon the best possible utilization of health data. Therefore, the growth of CH research can be considered to also be related to the availability of relevant health data. From 2016 onward, the amount of available networked data is more than 10.000 billion GB, almost double the amount of available data in 2014 [233].

Empirical Studies

Our findings support the idea that over half of the CH studies are based on empirically informed approaches. More specifically, the majority of the solution proposals were empirically validated with case studies (59/208, 28.3%), followed by experiments (55/208, 26.4%), whereas 13.4% (28/208) used surveys. Case studies have been shown to be particularly suited to *how* and *why* questions, real-life contexts, and the building of theory [234]. They offer an opportunity to use real-time methodologies where the collection of data and empirical material “takes place at the same time as such data are unfolding and where events depend on each other in a sequential order” [235]. In particular, case studies can offer an “opportunity to observe and analyze a phenomenon previously inaccessible to scientific investigation” [236]. This predominance of case study approaches to the study of CH reflects the relative immaturity of this field of research and a need to build understanding and theoretical contributions in the area through better understanding of individual cases and their contextual parameters. An experiment, on the other hand, is predicated on the analysis of covariance and assumes that participants can be assigned at random, that there are equal numbers of cases in each cell of the factorial design, and " the correlations between or among the independent variables of a factorial design are zero" [237]. The high number of experiment-based studies can be likely explained by the high number of papers exploring specific technical solutions that can be validated in controlled conditions. This raises the possibility that CH researchers are channeled into empirical approaches that have long been considered *gold standard* in engineering and health research.

Approaches

The results for this MQ show that the majority of the included papers belong to other types of studies (66/208, 31.7%), which means that they used approaches that were not defined by our classification scheme. This finding could arise from a variety of factors. More than half of the included papers were from the medical discipline, a fact that has given rise to classification challenges, as reviewers were not able to fit them into one of the defined categories. Although mapping studies are common in the medical discipline [12], classifying research approaches that spanned studies from different disciplines was difficult. A more detailed classification scheme that incorporated classification systems from all disciplines may have delivered more precise results but would have made aggregation difficult and may have obscured any similarities across disciplines. Methods (56/208, 26.9%) were reported to be among the most frequent approaches, followed by tool-based techniques (43/208, 20.6%). This potentially reflects both the relative immaturity and the inherently interdisciplinary nature of the field of CH at this stage. Before researchers can develop models or frameworks that might be applied to CH, they must come to an agreement as to those methods and techniques that are both feasible and acceptable across the new field. The interdisciplinary nature of CH research makes it even more important to focus on methods. As Klein puts it: interdisciplinarity is "a means of solving problems and answering questions that cannot be satisfactorily addressed using single methods or approaches" [238]. Researchers prefer, however, those methods that are traditional to their discipline [239] and so time must be spent discussing and agreeing to those methods and techniques that can span the interdisciplinary boundaries of CH.

Research Topics and Conditions

The topic analysis shows that technical and medical disciplines dominate the research topics of the papers, with some references to related disciplines such as education or innovation research. The topic descriptions show a large body of papers discussing measurement-based monitoring with sensors and wearables. This reflects the fast development of body-worn sensors and wireless communication methods that allow the transfer and storage of large amounts of data for further analysis. Another body of papers focused on patient perspectives in CH solutions through a consumer viewpoint or user-centered design, which might reflect an increasing interest in patient empowerment and self-management solutions. The growth of CH as a research field might also reflect regulatory moves toward a data economy where rules for using personal data are clearer (eg, General Data Protection Regulation). Related topics included security, privacy, and interoperability issues. It has been suggested that leveraging interoperable CH technologies for chronic disease management can have multiple positive effects not only on patients but also on clinical outcomes, thus contributing, for example, to the promotion of outpatient care [87]. Although the implementation of CH interoperable scenarios in real-life contexts, such as the Whole System Demonstrator Program in the United Kingdom, had overall positive outcomes [87], issues related to security and lack of data standardization are among the challenges yet to be overcome [106,203].

Diabetes, cancer, and chronic heart conditions dominate the medical conditions covered. This is unsurprising given that these are leading health problems on the global stage [240]. This may also indicate that lifestyle-related conditions are especially suited to CH, as lifestyle changes require patient empowerment and may benefit from technologies used for unobtrusive measurement and personal health devices. Papers discussing issues related to aging, including dementia and falls were often present; among less common conditions, multiple sclerosis seemed to draw more attention.

Limitations

Owing to the interdisciplinary nature of our topic, we used an interdisciplinary team of researchers for analysis. Having researchers with different backgrounds could decrease inter-rater reliability, especially where we did not have a predefined list of values, as with the *topic* and *condition* parameters. Nevertheless, for the analysis of results, the authors relied upon the interpretation of descriptive statistics and visualizations, thus decreasing the threat to validity. In the same vein, to alleviate the authors' influence on the classification process, the development of the classification scheme relied on widely accepted guidelines [13]. The differences between content and style of abstracts in different research disciplines and traditions may have resulted in slight differences in information retrieval process. In some cases, the abstract did not include all information needed to classify the paper, and the researchers had to read parts of the full text to obtain all relevant information. However, the vast majority of the 208 publications identified have been classified purely based on the title, abstract, and keywords. In our view, subtle differences in the publication screening process have had only minor impacts on the main conclusions drawn from the 208 publications identified in our study.

The differences in publication practices between the disciplines probably had an influence on how the results of MQ1 were interpreted, as the role of conference and journal publications differs between disciplines [241]. For example, many highly regarded conferences within the business discipline do not publish conference proceedings (eg, European Group for Organizational Studies). Indeed, even those conferences that do publish proceedings may only do so for a subsection of the best papers, and authors will still be offered the option of removing their papers from those proceedings to protect future publishing opportunities (eg, Academy of Management). These factors combine to mean that verbal discussions may well have commenced within the business academic community in conferences that are not reflected within our review as it deals only with published material.

As the methodology of the systematic mapping study that we used in this research was originally developed in the context of software engineering, it is likely that some of the analysis parameters were less optimal in other disciplines. This is reflected, for example, in the large number of studies classified into the category *other* for the parameter *approach*. In the joint analysis meetings, the coding authors shared experiences of their difficulties in classifying the papers within the agreed analysis parameters. Alternative solutions were discussed, but it was difficult to reach a consensus that would have been satisfactory across disciplines. Ultimately, we decided to use the classification parameters proposed by our methodology. The validity of our conclusions is only applicable within the CH context.

To limit the threat related to the identification of primary studies and to include as many relevant papers as possible, 2 of the authors ran several iterations to test different strings of keywords. The adoption of the final set of keywords was used as it returned the largest number of studies. However, the list of studies might be incomplete, as additional or different terms might have an impact on the final selection of papers [242]. Nevertheless, in light of the interdisciplinary endeavor and the scope of the study, we believe that we have included the majority of the relevant literature. For the bibliographic search, no timeframe was defined; hence, the representativeness of the included studies was not affected by this factor. The results of this study should be considered under the prism of the specific search string and classification scheme and offer a baseline for future endeavors.

Our search strategy and inclusion criteria have omitted studies that are referenced in grey literature. However, the literature search was conducted in the world’s most leading and comprehensive databases for scientific knowledge. Furthermore, to alleviate the threat of publications’ nonavailability because of subscription paywalls, we performed the initial screening using a combination of university libraries to improve our access to papers. To address validity threats because of duplication, duplicate entries were removed based on the DOI and the intertitle lexical distance with the help of a custom Python script, which was ultimately manually reviewed to ensure duplicate removal.

Conclusions

On the basis of our results, we can conclude that CH research is an established field of research. The interdisciplinary nature of the field can be seen especially in papers at the intersection of the medical and technical disciplines. The number of papers in the business research publication forums is still smaller. However, business-related themes are visible in topics of papers, such as consumer orientation and innovation research, although at a much smaller scale than the topics of more technical and medical nature. For CH to succeed, money needs to move differently around the health care system. Most developed health care systems continue to reimburse care in a *cure* rather than a *prevention* mode. Cutting-edge technologies and redesigned care pathways may fail if they run contrary to the flow of health care finance. Our findings emphasize the need to increase business research in the area of CH or to find the vocabularies and keywords necessary to link existing business research with CH endeavors.

There is a growing need to involve and engage patients in their own care and, by extension, in the design of digital solutions to improve their efficacy. Tailored CH interventions may more effectively reach the intended audience in a meaningful way, but this requires in-depth understanding of the condition’s needs, barriers, and facilitators. The important role that health care professionals play in the health care system is in contrast with their lack of involvement in the design of CH. In the same vein, recent research suggests that health care professionals’ education in Europe is lacking in the area of health care information technologies [243]. Emerging trends such as user-centered design and the inclusion of patient representatives attempt to address these problems by creating CH solutions that are tailored to the characteristics and tasks of the intended users. Adoption of these and similar approaches could further the field of CH research and implementation.

Given that CH research has become far more widespread in the years since 2013, perhaps it is time to devote more research resources to the scalability of CH as reflected in empirical approaches that facilitate the use of larger populations.

We see our findings as the foundations of a research roadmap for CH researchers that challenge current thinking in health care. Such a research agenda would go beyond investigations into the feasibility of individual technical solutions to examine and develop ecosystems of stakeholders, technologies, and infrastructures that together form new kinds of systemic solutions. Such an agenda would require more focus on research that addresses interdisciplinary methodological questions alongside the creation of vocabularies and frameworks for researchers working in different disciplines to effectively collaborate and examine interdisciplinary research questions through joint methodological approaches.

Acknowledgments

This paper is based upon work from European Cooperation in Science and Technology (COST) Action ENJECT TD 1405, supported by COST.

This research has been supported by a grant from the Tatra banka Foundation (2018vs108).

Authors' Contributions

MK and MI conceived, designed, led overall study conduct. MK and CM carried out the data collection and curation of visualizations. MK and MI and drafted the first version of the manuscript. All the authors except CM led analysis and interpretation of the data, NM proofread the final manuscript and suggested modifications. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Classification results for all the included papers.
[XLSX File (Microsoft Exel File), 17KB - jmir_v21i9e14394_app1.xlsx]

References

1. Mountford N, Chouvarda I, Isomursu M, Caulfield B. Vectors and drivers of connected health in Europe: a foundation for integrated care. Int J Integr Care 2017 Oct 17;17(5):305. [doi: 10.5334/ijic.3622]
2. Poon CC, Zhang YT. Perspectives on high technologies for low-cost healthcare. IEEE Eng Med Biol Mag 2008;27(5):42-47. [doi: 10.1109/MEMB.2008.923955] [Medline: 18799389]
3. Caulfield BM, Donnelly SC. What is connected health and why will it change your practice? QJM 2013 Aug;106(8):703-707. [doi: 10.1093/qjmed/hct114] [Medline: 23676416]
4. Mountford N, Kessie T, Quinlan M, Maher R, Smolders R, van Royen P, et al. Connected Health in Europe: Where are We Today?. Dublin, Ireland: ENJECT Consortium; 2017.
5. Wait S, Han D, Muthu V, Oliver K, Chrostowski S, Florindi F, et al. Towards sustainable cancer care: reducing inefficiencies, improving outcomes—a policy report from the All.Can initiative. J Cancer Policy 2017 Sep;13:47-64. [doi: 10.1016/j.jcpo.2017.05.004]
6. Giunti G, Giunta DH, Guisado-Fernandez E, Bender JL, Fernandez-Luque L. A biopsy of breast cancer mobile applications: state of the practice review. Int J Med Inform 2018 Feb;110:1-9 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.022] [Medline: 29331247]
7. Agboola SO, Kvedar JC. Institutional profile: center for connected health. Per Med 2012 May;9(3):337-340. [doi: 10.2217/pme.12.24] [Medline: 29758782]
8. Khan Y, Ostfeld AE, Lochner CM, Pierre A, Arias AC. Monitoring of vital signs with flexible and wearable medical devices. Adv Mater 2016 Jun;28(22):4373-4395. [doi: 10.1002/adma.201504366] [Medline: 26867696]
9. Sperrin M, Rushton H, Dixon WG, Normand A, Villard J, Chieh A, et al. Who self-weighs and what do they gain from it? A retrospective comparison between smart scale users and the general population in England. J Med Internet Res 2016 Jan 21;18(1):e17 [FREE Full text] [doi: 10.2196/jmir.4767] [Medline: 26794900]
10. Ostherr K, Killoran P, Shegog R, Bruera E. Death in the digital age: a systematic review of information and communication technologies in end-of-life care. J Palliat Med 2016 Apr;19(4):408-420 [FREE Full text] [doi: 10.1089/jpm.2015.0341] [Medline: 26713368]
11. Karampela M, Ouhbi S, Isomursu M. Personal health data: a systematic mapping study. Int J Med Inform 2018 Oct;118:86-98. [doi: 10.1016/j.ijmedinf.2018.08.006] [Medline: 30153927]
12. Petersen K, Feldt R, Mujtaba S, Mattsson M. Systematic Mapping Studies in Software Engineering. In: Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering. 2008 Presented at: EASE'08; June 26-27, 2008; Swindon, UK p. 68-77 URL: http://dl.acm.org/citation.cfm?id=2227115.2227123
13. Kitchenham BA, Budgen D, Brereton OP. Using mapping studies as the basis for further research – a participant-observer case study. Inf Softw Technol 2011 Jun;53(6):638-651. [doi: 10.1016/j.infsof.2010.12.011]
14. Levenshtein VI. Binary codes capable of correcting deletions, insertions, and reversals. Sov Phys Dokl 1996;10(8):707-710 [FREE Full text]
15. Brereton P, Kitchenham BA, Budgen D, Turner M, Khalil M. Lessons from applying the systematic literature review process within the software engineering domain. J Syst Software 2007 Apr;80(4):571-583. [doi: 10.1016/j.jss.2006.07.009]
16. Jorgensen M, Shepperd M. A systematic review of software development cost estimation studies. IIEEE Trans Software Eng 2007;33(1):33-53. [doi: 10.1109/TSE.2007.256943]
17. Chouvarda IG, Goulis DG, Lambrinoudaki I, Maglaveras N. Connected health and integrated care: toward new models for chronic disease management. Maturitas 2015 Sep;82(1):22-27. [doi: 10.1016/j.maturitas.2015.03.015] [Medline: 25891502]
18. Warnet S. ['Health professionals will become experts in the world of connected health']. Rev Infirm 2017 Nov;66(235):21-23. [doi: 10.1016/j.revinf.2017.09.003] [Medline: 29127973]
19. Raymond G, Léo M. [E-health applications and their users]. Rev Infirm 2017 Nov;66(235):24-25. [doi: 10.1016/j.revinf.2017.09.004] [Medline: 29127974]
20. Gorwood P. [The contribution of digital technology to mental healthcare: much hope, a few concerns]. Soins Psychiatr 2017;38(311):12-15. [doi: 10.1016/j.spsy.2017.04.006] [Medline: 28683879]

21. Wang J, Cai C, Padhye N, Orlander P, Zare M. A behavioral lifestyle intervention enhanced with multiple-behavior self-monitoring using mobile and connected tools for underserved individuals with type 2 diabetes and comorbid overweight or obesity: pilot comparative effectiveness trial. JMIR Mhealth Uhealth 2018 Apr 10;6(4):e92 [FREE Full text] [doi: 10.2196/mhealth.4478] [Medline: 29636320]
22. Harte R, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
23. Harte R, Quinlan LR, Glynn L, Rodriguez-Molinero A, Scharf T, Carenas C, et al. A multi-stage human factors and comfort assessment of instrumented insoles designed for use in a connected health infrastructure. J Pers Med 2015 Dec 16;5(4):487-508 [FREE Full text] [doi: 10.3390/jpm5040487] [Medline: 26694468]
24. Bloss CS, Wineinger NE, Peters M, Boeldt DL, Ariniello L, Kim JY, et al. A prospective randomized trial examining health care utilization in individuals using multiple smartphone-enabled biosensors. PeerJ 2016;4:e1554 [FREE Full text] [doi: 10.7717/peerj.1554] [Medline: 26788432]
25. Patel MS, Volpp KG, Rosin R, Bellamy SL, Small DS, Fletcher MA, et al. A randomized trial of social comparison feedback and financial incentives to increase physical activity. Am J Health Promot 2016 Jul;30(6):416-424 [FREE Full text] [doi: 10.1177/0890117116658195] [Medline: 27422252]
26. Wen F. A robust uniqueness-and-anonymity-preserving remote user authentication scheme for connected health care. J Med Syst 2013 Dec;37(6):9980. [doi: 10.1007/s10916-013-9980-1] [Medline: 24146334]
27. Das AK, Goswami A. A secure and efficient uniqueness-and-anonymity-preserving remote user authentication scheme for connected health care. J Med Syst 2013 Jun;37(3):9948. [doi: 10.1007/s10916-013-9948-1] [Medline: 23660745]
28. Santos DF, Gorgônio KC, Perkusich A, Almeida HO. A standard-based and context-aware architecture for personal healthcare smart gateways. J Med Syst 2016 Oct;40(10):224. [doi: 10.1007/s10916-016-0580-8] [Medline: 27624493]
29. Chang YF, Yu SH, Shiao DR. A uniqueness-and-anonymity-preserving remote user authentication scheme for connected health care. J Med Syst 2013 Apr;37(2):9902. [doi: 10.1007/s10916-012-9902-7] [Medline: 23321970]
30. DePasse JW, Chen CE, Sawyer A, Jethwani K, Sim I. Academic medical centers as digital health catalysts. Healthc (Amst) 2014 Sep;2(3):173-176. [doi: 10.1016/j.hjdsi.2014.05.006] [Medline: 26250503]
31. Ford AR, Gibbons CM, Torres J, Kornmehl HA, Singh S, Young PM, et al. Access to dermatological care with an innovative online model for psoriasis management: results from a randomized controlled trial. Telemed J E Health 2019 Jul;25(7):619-627 [FREE Full text] [doi: 10.1089/tmj.2018.0160] [Medline: 30222518]
32. Mountford N, Zubiete ED, Kessie T, Garcia-Zapirain B, Nuño-Solinís R, Coyle D, et al. Activating technology for connected health in cancer: protocol for a research and training program. JMIR Res Protoc 2018 Jan 24;7(1):e14 [FREE Full text] [doi: 10.2196/resprot.8900] [Medline: 29367184]
33. Fernandez-Luque L, Staccini P. All that glitters is not gold: consumer health informatics and education in the era of social media and health apps. Findings from the yearbook 2016 section on consumer health informatics. Yearb Med Inform 2016 Nov 10(1):188-193 [FREE Full text] [doi: 10.15265/IY-2016-045] [Medline: 27830250]
34. Kuziemsky CE, Gogia SB, Househ M, Petersen C, Basu A. Balancing health information exchange and privacy governance from a patient-centred connected health and telehealth perspective. Yearb Med Inform 2018 Aug;27(1):48-54 [FREE Full text] [doi: 10.1055/s-0038-1641195] [Medline: 29681043]
35. Jethwani K, Kvedar J, Kvedar J. Behavioral phenotyping: a tool for personalized medicine. Per Med 2010 Nov;7(6):689-693. [doi: 10.2217/pme.10.62] [Medline: 29788562]
36. Saunders G, Jacobs P. Beyond the audiology clinic: innovations and possibilities of connected health. J Am Acad Audiol 2015 Oct;26(9):748-749. [doi: 10.3766/jaaa.26.9.1] [Medline: 26415967]
37. Topaz M, Pruinelli L. Big data and nursing: implications for the future. Stud Health Technol Inform 2017;232:165-171. [doi: 10.3390/ijerph13101015] [Medline: 28106594]
38. Watson AJ, Grant RW, Bello H, Hoch DB. Brave new worlds: how virtual environments can augment traditional care in the management of diabetes. J Diabetes Sci Technol 2008 Jul;2(4):697-702 [FREE Full text] [doi: 10.1177/193229680800200422] [Medline: 19885247]
39. Gay V, Leijdekkers P. Bringing health and fitness data together for connected health care: mobile apps as enablers of interoperability. J Med Internet Res 2015 Nov 18;17(11):e260 [FREE Full text] [doi: 10.2196/jmir.5094] [Medline: 26581920]
40. Spruijt-Metz D, Hekler E, Saranummi N, Intille S, Korhonen I, Nilsen W, et al. Building new computational models to support health behavior change and maintenance: new opportunities in behavioral research. Transl Behav Med 2015 Sep;5(3):335-346 [FREE Full text] [doi: 10.1007/s13142-015-0324-1] [Medline: 26327939]
41. McConnochie KM, Wood NE, Alarie C, Ronis S. Care offered by an information-rich pediatric acute illness connected care model. Telemed J E Health 2016 Jun;22(6):465-472. [doi: 10.1089/tmj.2015.0161] [Medline: 26701609]
42. Hao T, Chang H, Ball M, Lin K, Zhu X. cHRV uncovering daily stress dynamics using bio-signal from consumer wearables. In: Gundlapalli AV, Jaulent MC, Zhao D, editors. MEDINFO 2017: Precision Healthcare through Informatics. Amsterdam, Netherlands: IOS Press; 2017:98-102.

43. Weinstein RS, Krupinski EA, Doarn CR. Clinical examination component of telemedicine, telehealth, mhealth, and connected health medical practices. Med Clin North Am 2018 May;102(3):533-544. [doi: 10.1016/j.mcna.2018.01.002] [Medline: 29650074]
44. Wu Y, Wang Y, Ji M. Competencies related to informatics and information management for practicing nurses in select countries in Asia. Stud Health Technol Inform 2017;232:86-96. [doi: 10.1891/9780826121844.0010] [Medline: 28106586]
45. Adeshina AM, Hashim R. Computational approach for securing radiology-diagnostic data in connected health network using high-performance GPU-accelerated AES. Interdiscip Sci 2017 Mar;9(1):140-152. [doi: 10.1007/s12539-015-0140-9] [Medline: 26754740]
46. Cohen M. Connected health and multiple sclerosis. Rev Neurol (Paris) 2018 Jun;174(6):480-485. [doi: 10.1016/j.neurol.2018.03.008] [Medline: 29680178]
47. McElnay JC. Connected health and the pharmacist. Int J Pharm Pract 2011 Feb;19(1):3-4. [doi: 10.1111/j.2042-7174.2010.00088.x] [Medline: 21235653]
48. Frist WH. Connected health and the rise of the patient-consumer. Health Aff (Millwood) 2014 Feb;33(2):191-193. [doi: 10.1377/hlthaff.2013.1464] [Medline: 24493759]
49. Kvedar JC, Herzlinger R, Holt M, Sanders JH. Connected health as a lever for healthcare reform: dialogue with featured speakers from the 5th annual connected health symposium. Telemed J E Health 2009 May;15(4):312-319. [doi: 10.1089/tmj.2009.9972] [Medline: 19441948]
50. Barr PJ, McElnay JC, Hughes CM. Connected health care: the future of health care and the role of the pharmacist. J Eval Clin Pract 2012 Feb;18(1):56-62. [doi: 10.1111/j.1365-2753.2010.01522.x] [Medline: 20698917]
51. El Amrani L, Engberink AO, Ninot G, Hayot M, Carbonnel F. Connected health devices for health care in french general medicine practice: cross-sectional study. JMIR Mhealth Uhealth 2017 Dec 21;5(12):e193 [FREE Full text] [doi: 10.2196/mhealth.7427] [Medline: 29269336]
52. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014 Feb;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]
53. Colorafi K. Connected health: a review of the literature. Mhealth 2016;2:13 [FREE Full text] [doi: 10.21037/mhealth.2016.03.09] [Medline: 28293591]
54. Aktas A, Hullihen B, Shrotriya S, Thomas S, Walsh D, Estfan B. Connected health: cancer symptom and quality-of-life assessment using a tablet computer: a pilot study. Am J Hosp Palliat Care 2015 Mar;32(2):189-197. [doi: 10.1177/1049909113510963] [Medline: 24212101]
55. Iglehart JK. Connected health: emerging disruptive technologies. Health Aff (Millwood) 2014 Feb;33(2):190. [doi: 10.1377/hlthaff.2014.0042] [Medline: 24493758]
56. Séblain D, Capitaine A, Khonsari RH. Connected toothbrush, augmented reality and oral hygiene in children with cleft lip and palate. J Stomatol Oral Maxillofac Surg 2019 Feb;120(1):83-85. [doi: 10.1016/j.jormas.2018.09.001] [Medline: 30282054]
57. Vlahu-Gjorgievska E, Koceski S, Kulev I, Trajkovik V. Connected-health algorithm: development and evaluation. J Med Syst 2016 Apr;40(4):109. [doi: 10.1007/s10916-016-0466-9] [Medline: 26922593]
58. Wang J, Coleman DC, Kanter J, Ummer B, Siminerio L. Connecting smartphone and wearable fitness tracker data with a nationally used electronic health record system for diabetes education to facilitate behavioral goal monitoring in diabetes care: protocol for a pragmatic multi-site randomized trial. JMIR Res Protoc 2018 Apr 2;7(4):e10009 [FREE Full text] [doi: 10.2196/10009] [Medline: 29610111]
59. Xu L, Wu F. Cryptanalysis and improvement of a user authentication scheme preserving uniqueness and anonymity for connected health care. J Med Syst 2015 Feb;39(2):10. [doi: 10.1007/s10916-014-0179-x] [Medline: 25631840]
60. Bond R, Finlay DD, Guldenring D, Breen C. Data driven computer simulation to analyse an ECG limb lead system used in connected health environments. Methods Inf Med 2016 May 17;55(3):258-265. [doi: 10.3414/ME15-01-0120] [Medline: 27096217]
61. Anderson D, Liu R, Subramony JA, Cammack J. Design control considerations for biologic-device combination products. Adv Drug Deliv Rev 2017 Mar;112:101-105. [doi: 10.1016/j.addr.2017.01.003] [Medline: 28088344]
62. Dur O, Rhoades C, Ng MS, Elsayed R, van Mourik R, Majmudar MD. Design rationale and performance evaluation of the wavelet health wristband: benchtop validation of a wrist-worn physiological signal recorder. JMIR Mhealth Uhealth 2018 Oct 16;6(10):e11040 [FREE Full text] [doi: 10.2196/11040] [Medline: 30327288]
63. Watson AJ, Kvedar JC, Rahman B, Pelletier AC, Salber G, Grant RW. Diabetes connected health: a pilot study of a patient- and provider-shared glucose monitoring web application. J Diabetes Sci Technol 2009 Mar 1;3(2):345-352 [FREE Full text] [doi: 10.1177/193229680900300216] [Medline: 20144366]
64. Wang J, Chu CF, Li C, Hayes L, Siminerio L. Diabetes educators' insights regarding connecting mobile phone- and wearable tracker-collected self-monitoring information to a nationally-used electronic health record system for diabetes education: descriptive qualitative study. JMIR Mhealth Uhealth 2018 Jul 26;6(7):e10206 [FREE Full text] [doi: 10.2196/10206] [Medline: 30049667]

65. Lin C, Song Z, Song H, Zhou Y, Wang Y, Wu G. Differential privacy preserving in big data analytics for connected health. J Med Syst 2016 Apr;40(4):97. [doi: 10.1007/s10916-016-0446-0] [Medline: 26872779]
66. Sapci AH, Sapci HA. Digital continuous healthcare and disruptive medical technologies: m-health and telemedicine skills training for data-driven healthcare. J Telemed Telecare 2018 Aug 22:- (epub ahead of print). [doi: 10.1177/1357633X18793293] [Medline: 30134779]
67. Novillo-Ortiz D, Dumit EM, D'Agostino M, Becerra-Posada F, Kelley ET, Torrent-Sellens J, et al. Digital health in the Americas: advances and challenges in connected health. BMJ Innov 2018 Jul;4(3):123-127 [FREE Full text] [doi: 10.1136/bmjinnov-2017-000258] [Medline: 30101033]
68. Stara V, Harte R, di Rosa M, Glynn L, Casey M, Hayes P, et al. Does culture affect usability? A trans-European usability and user experience assessment of a falls-risk connected health system following a user-centred design methodology carried out in a single European country. Maturitas 2018 Aug;114:22-26. [doi: 10.1016/j.maturitas.2018.05.002] [Medline: 29907242]
69. Kvedar JC, Nesbitt T, Kvedar JG, Darkins A. E-patient connectivity and the near term future. J Gen Intern Med 2011 Nov;26(Suppl 2):636-638 [FREE Full text] [doi: 10.1007/s11606-011-1763-0] [Medline: 21989615]
70. Turnin MC, Schirr-Bonnans S, Martini J, Buisson JC, Taoui S, Chauchard MC, et al. Educ@dom: comparative study of the telemonitoring of patients with type 2 diabetes versus standard monitoring-study protocol for a randomized controlled study. Diabetol Metab Syndr 2017;9:52 [FREE Full text] [doi: 10.1186/s13098-017-0252-y] [Medline: 28702091]
71. Harte R, Hall T, Glynn L, Rodríguez-Molinero A, Scharf T, Quinlan LR, et al. Enhancing home health mobile phone app usability through general smartphone training: usability and learnability case study. JMIR Hum Factors 2018 Apr 26;5(2):e18 [FREE Full text] [doi: 10.2196/humanfactors.7718] [Medline: 29699969]
72. Aberger EW, Migliozzi D, Follick MJ, Malick T, Ahern DK. Enhancing patient engagement and blood pressure management for renal transplant recipients via home electronic monitoring and web-enabled collaborative care. Telemed J E Health 2014 Sep;20(9):850-854 [FREE Full text] [doi: 10.1089/tmj.2013.0317] [Medline: 25046403]
73. O'Neill SA, Nugent CD, Donnelly MP, McCullagh P, McLaughlin J. Evaluation of connected health technology. Technol Health Care 2012;20(3):151-167. [doi: 10.3233/THC-2012-0662] [Medline: 22735731]
74. Bushko R. Future of ehealth: can consumers cure themselves? Stud Health Technol Inform 2009;149:178-184. [Medline: 19745480]
75. Weinstein RS, Lopez AM. Health literacy and connected health. Health Aff (Millwood) 2014 Jun;33(6):1103-1104. [doi: 10.1377/hlthaff.2014.0396] [Medline: 24889963]
76. Harte RP, Glynn LG, Broderick BJ, Rodriguez-Molinero A, Baker PM, McGuiness B, et al. Human centred design considerations for connected health devices for the older adult. J Pers Med 2014 Jun 4;4(2):245-281 [FREE Full text] [doi: 10.3390/jpm4020245] [Medline: 25563225]
77. Harte R, Quinlan LR, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, et al. Human-centered design study: enhancing the usability of a mobile phone app in an integrated falls risk detection system for use by older adult users. JMIR Mhealth Uhealth 2017 May 30;5(5):e71 [FREE Full text] [doi: 10.2196/mhealth.7046] [Medline: 28559227]
78. Pelletier A, Jethwani K, Bello H, Kvedar J, Grant R. Implementing a web-based home monitoring system within an academic health care network: barriers and facilitators to innovation diffusion. J Diabetes Sci Technol 2011 Jan 1;5(1):32-38 [FREE Full text] [doi: 10.1177/193229681100500105] [Medline: 21303622]
79. Casey M, Hayes PS, Heaney D, Dowie L, Ólaighin G, Matero M, et al. Implementing transnational telemedicine solutions: a connected health project in rural and remote areas of six northern periphery countries series on European collaborative projects. Eur J Gen Pract 2013 Mar;19(1):52-58. [doi: 10.3109/13814788.2012.761440] [Medline: 23432039]
80. Xie Q, Liu W, Wang S, Han L, Hu B, Wu T. Improvement of a uniqueness-and-anonymity-preserving user authentication scheme for connected health care. J Med Syst 2014 Sep;38(9):91. [doi: 10.1007/s10916-014-0091-4] [Medline: 24994512]
81. Feied CF, Handler JA, Gillam M, Smith MS. Indistinguishable from magic: health and wellness in a future of sufficiently advanced technology. Stud Health Technol Inform 2009;149:29-48. [Medline: 19745470]
82. Goossen W. Informatics competencies in connected health: annotated bibliography. Stud Health Technol Inform 2017;232:241-251. [Medline: 28106604]
83. Krupinski EA. Innovations and possibilities in connected health. J Am Acad Audiol 2015 Oct;26(9):761-767. [doi: 10.3766/jaaa.14047] [Medline: 26415969]
84. Singh K, Kvedar JC. Innovations in connected health. Stud Health Technol Inform 2009;149:111-129. [doi: 10.3766/jaaa.14047] [Medline: 19745476]
85. Maglaveras N, Kilintzis V, Koutkias V, Chouvarda I. Integrated care and connected health approaches leveraging personalised health through big data analytics. Stud Health Technol Inform 2016;224:117-122. [doi: 10.3233/978-1-61499-653-8-117] [Medline: 27225565]
86. Sweeney KT, Leamy DJ, Ward TE, McLoone S. Intelligent Artifact Classification for Ambulatory Physiological Signals. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology. 2010 Presented at: EMBS'10; August 31-September 4, 2010; Buenos Aires, Argentina p. 6349-6352. [doi: 10.1109/iembs.2010.5627285]
87. Hay I, Lim K, Wartena F. Interoperability for device manufacturers: building a personal, connected health ecosystem. Biomed Instrum Technol 2012;46(Suppl 2):71-77. [doi: 10.2345/0899-8205-46.s2.71] [Medline: 23039780]

88. Caulfield BM. Introduction: review series--connected health. QJM 2013 Aug;106(8):701. [doi: 10.1093/qjmed/hct138] [Medline: 23863952]
89. Loiselle CG, Ahmed S. Is connected health contributing to a healthier population? J Med Internet Res 2017 Nov 10;19(11):e386 [FREE Full text] [doi: 10.2196/jmir.8309] [Medline: 29127077]
90. Trout KE, Rampa S, Wilson FA, Stimpson JP. Legal mapping analysis of state telehealth reimbursement policies. Telemed J E Health 2017 Oct;23(10):805-814. [doi: 10.1089/tmj.2017.0016] [Medline: 28430029]
91. Carroll N, Richardson I. Mapping a careflow network to assess the connectedness of connected health. Health Informatics J 2019 Mar;25(1):106-125. [doi: 10.1177/1460458217702943] [Medline: 28438102]
92. Campo D, Khettab H, Yu R, Genain N, Edouard P, Buard N, et al. Measurement of aortic pulse wave velocity with a connected bathroom scale. Am J Hypertens 2017 Sep 1;30(9):876-883 [FREE Full text] [doi: 10.1093/ajh/hpx059] [Medline: 28520843]
93. Wang J. Mobile and connected health technologies for older adults aging in place. J Gerontol Nurs 2018 Jun 1;44(6):3-5. [doi: 10.3928/00989134-20180509-01] [Medline: 29846735]
94. Rothman BS, Gupta RK, McEvoy MD. Mobile technology in the perioperative arena: rapid evolution and future disruption. Anesth Analg 2017 Mar;124(3):807-818. [doi: 10.1213/ANE.0000000000001858] [Medline: 28151816]
95. Merchant RM, Groeneveld PW. Neighborhood-level disparities in resuscitation and the potential of connected health. JAMA Cardiol 2017 Oct 1;2(10):1118-1119. [doi: 10.1001/jamacardio.2017.2763] [Medline: 28854314]
96. Saranto K, Ronquillo C, Velez O. Nursing competencies for multiple modalities of connected health technologies. Stud Health Technol Inform 2017;232:172-182. [doi: 10.3233/978-1-61499-738-2-172] [Medline: 28106595]
97. Kim KW, Lee JD. On the security of two remote user authentication schemes for telecare medical information systems. J Med Syst 2014 May;38(5):17. [doi: 10.1007/s10916-014-0017-1] [Medline: 24777588]
98. Lee D, Helal S, Anton S, de Deugd S, Smith A. Participatory and persuasive telehealth. Gerontology 2012;58(3):269-281 [FREE Full text] [doi: 10.1159/000329892] [Medline: 21893945]
99. Ternullo J, Jethwani K, Lane S, Myint UK, Havasy R, Carter M, et al. Partners healthcare center for connected health. Telemed J E Health 2013 May;19(5):363-367. [doi: 10.1089/tmj.2012.0294] [Medline: 23330595]
100. Argent R, Daly A, Caulfield B. Patient involvement with home-based exercise programs: can connected health interventions influence adherence? JMIR Mhealth Uhealth 2018 Mar 1;6(3):e47 [FREE Full text] [doi: 10.2196/mhealth.8518] [Medline: 29496655]
101. Skiba DJ, Barton AJ, Estes K, Gilliam E, Knapfel S, Lee C, et al. Preparing the next generation of advanced practice nurses for connected care. Stud Health Technol Inform 2016;225:307-313. [Medline: 27332212]
102. Haluza D, Naszay M, Stockinger A, Jungwirth D. Prevailing opinions on connected health in Austria: results from an online survey. Int J Environ Res Public Health 2016 Aug 11;13(8):813 [FREE Full text] [doi: 10.3390/ijerph13080813] [Medline: 27529261]
103. Barr PJ, Brady SC, Hughes CM, McElnay JC. Public knowledge and perceptions of connected health. J Eval Clin Pract 2014 Jun;20(3):246-254. [doi: 10.1111/jep.12118] [Medline: 24661432]
104. Weinhandl ED, Collins AJ. Relative risk of home hemodialysis attrition in patients using a telehealth platform. Hemodial Int 2018 Jul;22(3):318-327. [doi: 10.1111/hdi.12621] [Medline: 29210164]
105. Bollyky JB, Bravata D, Yang J, Williamson M, Schneider J. Remote lifestyle coaching plus a connected glucose meter with certified diabetes educator support improves glucose and weight loss for people with type 2 diabetes. J Diabetes Res 2018;2018:3961730 [FREE Full text] [doi: 10.1155/2018/3961730] [Medline: 29888288]
106. Kreuzthaler M, Martínez-Costa C, Kaiser P, Schulz S. Semantic technologies for re-use of clinical routine data. Stud Health Technol Inform 2017;236:24-31. [Medline: 28508775]
107. Goossen-Baremans A, Collins S, Park HA. Semanticification in connected health. Stud Health Technol Inform 2017;232:133-151. [doi: 10.4258/hir.2010.16.4.201] [Medline: 28106592]
108. Sathyanarayana A, Joty S, Fernandez-Luque L, Ofli F, Srivastava J, Elmagarmid A, et al. Sleep quality prediction from wearable data using deep learning. JMIR Mhealth Uhealth 2016 Nov 4;4(4):e125 [FREE Full text] [doi: 10.2196/mhealth.6562] [Medline: 27815231]
109. Helal A, Cook DJ, Schmalz M. Smart home-based health platform for behavioral monitoring and alteration of diabetes patients. J Diabetes Sci Technol 2009 Jan;3(1):141-148 [FREE Full text] [doi: 10.1177/193229680900300115] [Medline: 20046657]
110. Fernandez-Luque L, Vilmarlund V, Borycki E, Schulz S, Kuziemsky C, Marschollek M, et al. Social media as catalyzer for connected health: hype or hope? Perspectives from IMIA working groups. In: Sermeus W, Procter PM, Weber P, editors. Nursing Informatics 2016: eHealth for All: Every Level Collaboration - From Project to Realization. Amsterdam, Netherlands: IOS Press; 2016:602-604.
111. Thomas S, Walsh D, Shrotriya S, Aktas A, Hullihen B, Estfan B, et al. Symptoms, quality of life, and daily activities in people with newly diagnosed solid tumors presenting to a medical oncologist. Am J Hosp Palliat Care 2017 Aug;34(7):611-621. [doi: 10.1177/1049909116649948] [Medline: 27217423]
112. Gladden C, Beck L, Chandler D. Tele-audiology: expanding access to hearing care and enhancing patient connectivity. J Am Acad Audiol 2015 Oct;26(9):792-799. [doi: 10.3766/jaaa.14107] [Medline: 26415973]

113. Landers SH. The case for 'connected health' at home. Cleve Clin J Med 2013 Jan;80(Electronic Suppl 1):eS27-eS29. [doi: 10.3949/ccjm.80.e-s1.06] [Medline: 23420799]
114. Ansary A, Azuma A, Komatireddy R, Barrett PM. The connected health of cardiovascular medicine: current status and future directions. QJM 2013 Nov;106(11):979-982 [FREE Full text] [doi: 10.1093/qjmed/hct159] [Medline: 23904521]
115. Agboola SO, Ju W, Elfiky A, Kvedar JC, Jethwani K. The effect of technology-based interventions on pain, depression, and quality of life in patients with cancer: a systematic review of randomized controlled trials. J Med Internet Res 2015 Mar 13;17(3):e65 [FREE Full text] [doi: 10.2196/jmir.4009] [Medline: 25793945]
116. Agboola SO, Ball M, Kvedar JC, Jethwani K. The future of connected health in preventive medicine. QJM 2013 Sep;106(9):791-794. [doi: 10.1093/qjmed/hct088] [Medline: 23598385]
117. Ni M, Brown LG, Lawler D, Ellis TD, Deangelis T, Latham NK, et al. The rehabilitation enhancing aging through connected health (REACH) study: study protocol for a quasi-experimental clinical trial. BMC Geriatr 2017 Sep 20;17(1):221 [FREE Full text] [doi: 10.1186/s12877-017-0618-x] [Medline: 28931377]
118. Allaert FA, Mazen NJ, Legrand L, Quantin C. The tidal waves of connected health devices with healthcare applications: consequences on privacy and care management in European healthcare systems. BMC Med Inform Decis Mak 2017 Jan 17;17(1):10 [FREE Full text] [doi: 10.1186/s12911-017-0408-6] [Medline: 28095843]
119. Tharmalingam S, Hagens S, Zelmer J. The value of connected health information: perceptions of electronic health record users in Canada. BMC Med Inform Decis Mak 2016 Jul 16;16:93 [FREE Full text] [doi: 10.1186/s12911-016-0330-3] [Medline: 27422571]
120. Deng Y, Liu NY, Tsow F, Xian X, Krajmalnik-Brown R, Tao N, et al. Tracking personal health-environment interaction with novel mobile sensing devices. Sensors (Basel) 2018 Aug 14;18(8):E2670 [FREE Full text] [doi: 10.3390/s18082670] [Medline: 30110932]
121. Weaver SJ, Che XX, Petersen LA, Hysong SJ. Unpacking care coordination through a multiteam system lens: a conceptual framework and systematic review. Med Care 2018 Mar;56(3):247-259. [doi: 10.1097/MLR.0000000000000874] [Medline: 29356720]
122. Kvedar JC, Hwang J, Moorhead T, Orlov LM, Ubel PA. Up from crisis: overhauling healthcare information, payment, and delivery in extraordinary times. Dialogue with featured speakers from the 6th annual connected health symposium. Telemed J E Health 2009 Sep;15(7):634-641. [doi: 10.1089/tmj.2009.9948] [Medline: 19747071]
123. Aldeer M, Martin RP. Medication Adherence Monitoring Using Modern Technology. In: Proceedings of the 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference. 2017 Presented at: UEMCON'17; October 19-21, 2017; New York, USA p. 491-497. [doi: 10.1109/uemcon.2017.8249101]
124. Härmä A, Helaoui R. Probabilistic Scoring of Validated Insights for Personal Health Services. In: Proceedings of the Symposium Series on Computational Intelligence (SSCI). 2016 Presented at: SSCI'16; December 6-9, 2016; Athens, Greece p. 1-6. [doi: 10.1109/ssci.2016.7849892]
125. Majma N, Babamir SM, Monadjemi A. Utilizing Fuzzy Colored Petri-Nets to Monitor Cardiac Pacemaker Behavior. In: Proceedings of the 10th International Conference on Application of Information and Communication Technologies. 2016 Presented at: AICT'16; October 12-14, 2016; Baku, Azerbaijan p. 1. [doi: 10.1109/icaict.2016.7991824]
126. Ahouandjinou AS, Assogba K, Motamed C. Smart and Pervasive ICU Based-IoT for Improving Intensive Health Care. In: Proceedings of the International Conference on Bio-engineering for Smart Technologies. 2016 Presented at: BioSMART'16; December 4-7, 2016; Dubai, United Arab Emirates p. 1-4. [doi: 10.1109/biosmart.2016.7835599]
127. Djelouat H, Baali H, Amira A, Bensaali F. Joint Sparsity Recovery for Compressive Sensing Based EEG System. In: Proceedings of the 17th International Conference on Ubiquitous Wireless Broadband. 2017 Presented at: ICUWB'17; September 12-15, 2017; Salamanca, Spain p. 1-5. [doi: 10.1109/icuwb.2017.8251001]
128. Santos DF, Perkusich A, Almeida HO. Standard-Based and Distributed Health Information Sharing for mHealth IoT Systems. In: Proceedings of the 16th International Conference on e-Health Networking, Applications and Services. 2014 Presented at: Healthcom'14; October 15-18, 2014; Natal, Brazil p. 94-98. [doi: 10.1109/healthcom.2014.7001820]
129. Bellagente P, Depari A, Ferrari P, Flammini A, Sisinni E, Rinaldi S. M3IoT — Message-Oriented Middleware for M-health Internet of Things: Design and Validation. In: Proceedings of the International Instrumentation and Measurement Technology Conference. 2018 Presented at: I2MTC'18; May 14-17, 2018; Houston, TX, USA p. 1-6. [doi: 10.1109/i2mtc.2018.8409656]
130. Neggazi M, Amira A, Hamami L. A Wireless Reconfigurable System for Falls Detection. In: Proceedings of the 11th International Conference on Information Science, Signal Processing and their Applications. 2012 Presented at: ISSPA'12; July 2-5, 2012; Montreal, QC, Canada p. 77-82. [doi: 10.1109/isspa.2012.6310658]
131. Karamitsios K, Orphanoudakis T. Efficient IoT Data Aggregation for Connected Health Applications. In: Proceedings of the Symposium on Computers and Communications (ISCC). 2017 Presented at: ISCC'17; July 3-6, 2017; Heraklion, Greece p. 1182-1185. [doi: 10.1109/iscc.2017.8024685]
132. Sinharay A, Ghosh D, Deshpande P, Alam S, Banerjee R, Pal A. Smartphone Based Digital Stethoscope for Connected Health--A Direct Acoustic Coupling Technique. In: Proceedings of the First International Conference on Connected Health: Applications, Systems and Engineering Technologies. 2016 Presented at: CHASE'16; June 27-29, 2016; Washington, DC, USA p. 193-198. [doi: 10.1109/chase.2016.23]

133. Aldeer M, Martin RP, Howard RE. PillSense: Designing a Medication Adherence Monitoring System Using Pill Bottle-Mounted Wireless Sensors. In: Proceedings of the International Conference on Communications Workshops (ICC Workshops). 2018 Presented at: ICC Workshops'18; May 20-24, 2018; Kansas City, MO, USA p. 1-6. [doi: 10.1109/iccw.2018.8403547]
134. Daponte P, Lamonaca F, Picariello F, de Vito L, Mazzilli G, Tudosa I. A Survey of Measurement Applications Based on IoT. In: Proceedings of the Workshop on Metrology for Industry 4.0 and IoT. 2018 Presented at: IEEE'18; April 16-18, 2018; Brescia, Italy p. 1-6. [doi: 10.1109/metroi4.2018.8428335]
135. Starič KD, Bukovec P, Caulfield B, Bogdanova AM, Belani H. ICT-Assisted Personalized Therapy and Rehabilitation in Urinary Incontinence. In: Proceedings of the 14th International Conference on Telecommunications. 2017 Presented at: ConTEL'17; June 28-30, 2017; Zagreb, Croatia p. 69-74. [doi: 10.23919/contel.2017.8000041]
136. Chaher Y, Belaud JP, Pingaud H. Managing Open Innovation in Connected Health Through a Living Lab. In: Proceedings of the International Conference on Engineering, Technology and Innovation (ICE/ITMC). 2017 Presented at: ICE/ITMC'17; June 27-29, 2017; Funchal, Portugal p. 577-583. [doi: 10.1109/ice.2017.8279937]
137. da Silva JM, Oliveira C, Mendes B, Dias R, Marques T. Design for Dependability and Autonomy of a Wearable Cardiac and Coronary Monitor. In: Proceedings of the Euromicro Conference on Digital System Design. 2015 Presented at: DSD'15; August 26-28, 2015; Funchal, Portugal p. 567-570. [doi: 10.1109/dsd.2015.109]
138. Barbosa P, Queiroz J, Santos D, Figueiredo A, Leite F, Galdino K. RE4CH: Requirements Engineering for Connected Health. In: Proceedings of the 31st International Symposium on Computer-Based Medical Systems. 2018 Presented at: CBMS'18; June 18-21, 2018; Karlstad, Sweden p. 292-297. [doi: 10.1109/cbms.2018.00058]
139. Hassanalieragh M, Page A, Soyata T, Sharma G, Aktas M, Mateos G, et al. Health Monitoring and Management Using Internet-of-Things (IoT) Sensing with Cloud-Based Processing: Opportunities and Challenges. In: Proceedings of the International Conference on Services Computing. 2015 Presented at: SCC'15; June 27- July 2, 2015; New York, NY, USA p. 285-292. [doi: 10.1109/scc.2015.47]
140. Carroll N. Key success factors for smart and connected health software solutions. Computer 2016 Nov;49(11):22-28. [doi: 10.1109/mc.2016.340]
141. Liu J, Zhang C, Fang Y. EPIC: a differential privacy framework to defend smart homes against internet traffic analysis. IEEE Internet Things J 2018 Apr;5(2):1206-1217. [doi: 10.1109/jiot.2018.2799820]
142. Djelouat H, Baali H, Amira A, Bensaali F. CS-Based Fall Detection for Connected Health Applications. In: Proceedings of the Fourth International Conference on Advances in Biomedical Engineering. 2017 Presented at: ICABME'17; October 19-21, 2017; Beirut, Lebanon p. 1-4. [doi: 10.1109/icabme.2017.8167540]
143. Stephanie L. Singapore’s NEHR: Challenges on the Path to Connected Health. In: Proceedings of the International Conference on Industrial Engineering and Engineering Management. 2017 Presented at: IEEM'17; December 10-13, 2017; Singapore, Malaysia p. 1128-1132. [doi: 10.1109/ieem.2017.8290068]
144. Giunti G, Guisado-Fernandez E, Caulfield B. Connected Health in Multiple Sclerosis a Mobile Applications Review. In: Proceedings of the 30th International Symposium on Computer-Based Medical Systems. 2017 Presented at: CBMS'17; June 22-24, 2017; Thessaloniki, Greece p. 660-665. [doi: 10.1109/cbms.2017.27]
145. Kilintzis V, Maramis C, Maglaveras N. Wrist Sensors — An Application to Acquire Sensory Data From Android Wear® Smartwatches for Connected Health. In: Proceedings of the International Conference on Biomedical & Health Informatics. 2017 Presented at: BHI'17; February 16-19, 2017; Orlando, FL, USA p. 125-128. [doi: 10.1109/bhi.2017.7897221]
146. Lim YY, Garcia MS, Cuneen M, Thompson G, Assem H, Kenny L, et al. OP 4 learning to evaluate and manage antenatal blood pressure at home (LEANBH). Pregnancy Hypertens 2017 Jul;9:10-11. [doi: 10.1016/j.preghy.2017.07.027]
147. Gao J, Yi P, Chi Z, Zhu T. A smart medical system for dynamic closed-loop blood glucose-insulin control. Smart Health 2017 Jun;1-2:18-33. [doi: 10.1016/j.smhl.2017.04.001]
148. Zhai X, Ali AA, Amira A, Bensaali F. ECG encryption and identification based security solution on the Zynq SoC for connected health systems. J Parallel Distr Com 2017 Aug;106:143-152. [doi: 10.1016/j.jpdc.2016.12.016]
149. Santos DF, Almeida HO, Perkusich A. A personal connected health system for the internet of things based on the constrained application protocol. Comput Electr Eng 2015 May;44:122-136. [doi: 10.1016/j.compeleceng.2015.02.020]
150. Charlon Y, Campo E, Brulin D. Design and evaluation of a smart insole: application for continuous monitoring of frail people at home. Expert Syst Appl 2018 Apr;95:57-71. [doi: 10.1016/j.eswa.2017.11.024]
151. Barik RK, Dubey AC, Tripathi A, Pratik T, Sasane S, Lenka RK, et al. Mist data: leveraging mist computing for secure and scalable architecture for smart and connected health. Procedia Comput Sci 2018;125:647-653. [doi: 10.1016/j.procs.2017.12.083]
152. Lim YY, Garcia MS, Cuneen M, Mattson M, Thompson G, Assem H, et al. P 47 learning to evaluate and manage antenatal blood pressure at home (LEANBH). Pregnancy Hypertens 2017 Jul;9:58-59. [doi: 10.1016/j.preghy.2017.07.125]
153. Dowd WN, Cowell AJ, Regan D, Moran K, Slevin P, Doyle G, et al. An exploratory cost-effectiveness analysis of the connected health intervention to improve care for people with dementia: a simulation analysis. Health Serv Outcomes Res Method 2017 Dec 15;18(1):47-62. [doi: 10.1007/s10742-017-0175-y]

154. Fitzgerald S, Kachersky L, Saldanha N, Chung E, Farrell L, Laud G, et al. The Dark Side of Connected Health Technology: How Connectivity Creates Exercise Addiction Among Conscientious Users. In: Proceedings of The Association for Consumer Research Conference. 2016 Presented at: ACR'16; October 27-30, 2016; Berlin,Germany.
155. Battistella R, Burchfield D. The future of employment-based health insurance. J Healthc Manag 2000;45(1):46-56; discussion 56. [Medline: 11066952]
156. Gross N, Connolly N, McNamara P. The Case Centre. 2015. Connected Health Technology: Private Pharmacies Competing Innovatively in Ireland URL: https://www.thecasecentre.org/main/products/view?id=131345 [accessed 2019-09-02]
157. Mountford N, Kessie T. Towards a more holistic understanding of whole organizational networks: anthropological approaches in evolving markets. Electron J Bus Res Methods 2017;15(2):74-84 [FREE Full text]
158. Alchihabi A, Dervis A, Ever E, Al-Turjman F. A generic framework for optimizing performance metrics by tuning parameters of clustering protocols in WSNs. Wireless Netw 2018 Jan 25;25(3):1031-1046. [doi: 10.1007/s11276-018-1665-8]
159. Bucci S, Barrowclough C, Ainsworth J, Morris R, Berry K, Machin M, et al. Using mobile technology to deliver a cognitive behaviour therapy-informed intervention in early psychosis (Actissist): study protocol for a randomised controlled trial. Trials 2015 Sep 10;16:404 [FREE Full text] [doi: 10.1186/s13063-015-0943-3] [Medline: 26357943]
160. Burmaoglu S, Saritas O, Kıdak LB, Berber İ. Evolution of connected health: a network perspective. Scientometrics 2017 Jun 15;112(3):1419-1438. [doi: 10.1007/s11192-017-2431-x]
161. Hassan L, Swarbrick C, Sanders C, Parker A, Machin M, Tully MP, et al. Tea, talk and technology: patient and public involvement to improve connected health 'wearables' research in dementia. Res Involv Engagem 2017;3:12 [FREE Full text] [doi: 10.1186/s40900-017-0063-1] [Medline: 29062537]
162. McGrath D, Greene BR, Sheehan K, Walsh L, Kenny RA, Caulfield B. Stability of daily home-based measures of postural control over an 8-week period in highly functioning older adults. Eur J Appl Physiol 2015 Feb;115(2):437-449. [doi: 10.1007/s00421-014-3034-3] [Medline: 25344800]
163. O’Neill SA, Mason S, Parente G, Donnelly MP, Nugent CD, McClean S, et al. Video reminders as cognitive prosthetics for people with dementia. Ageing Int 2010 Nov 23;36(2):267-282. [doi: 10.1007/s12126-010-9089-5]
164. Carroll N, Travers M, Richardson I. Connecting Multistakeholder Analysis Across Connected Health Solutions. In: Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies. 2016 Presented at: BIOSTEC'16; February 21-23, 2016; Rome, Italy p. 319-339. [doi: 10.1007/978-3-319-54717-6_18]
165. O'Leary P, Buckley P, Richardson I. Modelling Care Pathways in a Connected Health Setting. In: Proceedings of the International Symposium on Foundations of Health Informatics Engineering and Systems. 2013 Presented at: FHIES'13; August 21-23, 2013; Macau, China p. 32-40. [doi: 10.1007/978-3-642-53956-5_3]
166. Myneni S, Rogith D, Franklin A. Digilego: A Standardized Analytics-Driven Consumer-Oriented Connected Health Framework. In: Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. 2018 Presented at: SBP-BRiMS'18; July 10-13, 2018; Washington DC, USA p. 263-273. [doi: 10.1007/978-3-319-93372-6_30]
167. Synnott J, McComb S, Nugent C, McLaughlin J. NI-CHIC: A Model for Academic Engagement with Industry. In: Proceedings of the Conference on the Ambient Intelligence for Health. 2015 Presented at: AmIHEALTH'15; December 1-4, 2015; Puerto Varas, Chile p. 258-263. [doi: 10.1007/978-3-319-26508-7_25]
168. Weitzel M, Smith A, Lee D, de Deugd S, Helal S. Participatory Medicine: Leveraging Social Networks in Telehealth Solutions. In: Proceedings of the International Conference on Smart Homes and Health Telematics. 2009 Presented at: ICOST'09; July 1-3, 2009; Tours, France p. 40-47. [doi: 10.1007/978-3-642-02868-7_6]
169. Galway L, Zhang S, Nugent C, McClean S, Finlay D, Scotney B. Utilizing Wearable Sensors to Investigate the Impact of Everyday Activities on Heart Rate. In: Proceedings of the International Conference on Smart Homes and Health Telematics. 2011 Presented at: ICOST'11; June 20-22, 2011; Montreal, QC, Canada p. 184-191. [doi: 10.1007/978-3-642-21535-3_24]
170. Leijdekkers P, Gay V. Improving User Engagement by Aggregating and Analysing Health and Fitness Data on a Mobile App. In: Proceedings of the International Conference on Smart Homes and Health Telematics. 2015 Presented at: ICOST'15; June 10-12, 2015; Geneva, Switzerland p. 325-330. [doi: 10.1007/978-3-319-19312-0_30]
171. Alkobaisi S, Bae WD, Narayanappa S. SCHAS: A Visual Evaluation Framework for Mobile Data Analysis of Individual Exposure to Environmental Risk Factors. In: Proceedings of the International Symposium on Spatial and Temporal Databases. 2015 Presented at: SSTD'15; August 26-28, 2015; Hong Kong, China p. 484-490. [doi: 10.1007/978-3-319-22363-6_27]
172. Carroll N, Richardson I. Aligning Healthcare Innovation and Software Requirements Through Design Thinking. In: Proceedings of the International Workshop on Software Engineering in Healthcare Systems. 2016 Presented at: SEHS'16; May 14-22, 2016; Austin, Texas p. 1-7. [doi: 10.1145/2897683.2897687]
173. Moon J, Yang H, Lee Y, Won D. Improvement of User Authentication Scheme Preserving Uniqueness and Anonymity for Connected Health Care. In: Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication. 2017 Presented at: IMCOM'17; January 5-7, 2017; Beppu, Japan. [doi: 10.1145/3022227.3022237]
174. Richardson I, O'Mahony J, Howarth P, O'Connor P, Glenny L. Connected Health: From Rural Ireland to Rural India. In: Proceedings of the International Workshop on Software Engineering in Healthcare Systems. 2016 Presented at: SEHS"16; May 14-22, 2016; Austin, Texas p. 39-42. [doi: 10.1145/2897683.2897689]

175. O'Sullivan P, Connolly A, Carroll N, Richardson I. IBM's Smarter Care: Challenges and Strategies. In: Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering. 2016 Presented at: EASE'16; June 1-3, 2016; Limerick, Ireland. [doi: 10.1145/2915970.2915980]
176. Kuziemsky C, Abbas RM, Carroll N. Toward a Connected Health Delivery Framework. In: Proceedings of the International Workshop on Software Engineering in Healthcare Systems. 2018 Presented at: SEHS'18; May 28, 2018; Gothenburg, Sweden p. 46-49. [doi: 10.1145/3194696.3194703]
177. Aldeer M, Martin RP, Howard RE. Tackling the Fidelity-Energy Trade-Off in Wireless Body Sensor Networks. In: Proceedings of the International Conference on Connected Health: Applications, Systems and Engineering Technologies. 2017 Presented at: CHASE'17; July 17-19, 2017; Philadelphia, PA, USA p. 1-7. [doi: 10.1109/chase.2017.52]
178. Rodolfo I, Laranjo L, Correia N, Duarte C. Design Strategy for a National Integrated Personal Health Record. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational. 2014 Presented at: NordiCHI'14; October 26-30, 2014; Helsinki, Finland p. 411-420. [doi: 10.1145/2639189.2641205]
179. Yuan C, Kropczynski J, Wirth R, Rosson MB, Carroll JM. Investigating Older Adults' Social Networks and Coproduction Activities for Health. In: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2017 Presented at: PervasiveHealth'17; May 23-26, 2017; Barcelona, Spain p. 68-77. [doi: 10.1145/3154862.3154876]
180. Karamitsios K, Orphanoudakis T, Dagiuklas T. Evaluation of IoT-Based Distributed Health Management Systems. In: Proceedings of the 20th Pan-Hellenic Conference on Informatics. 2016 Presented at: PCI'16; November 10-12, 2016; Patras, Greece. [doi: 10.1145/3003733.3003751]
181. Veiga JJ, Ward TE. Data Collection Requirements for Mobile Connected Health an End User Development Approach. In: Proceedings of the 1st International Workshop on Mobile Development. 2016 Presented at: Mobile!'16; October 31, 2016; Amsterdam, Netherlands p. 23-30. [doi: 10.1145/3001854.3001856]
182. Srinivasan K, Currim F, Ram S, Lindberg C, Sternberg E, Skeath P, et al. Feature Importance and Predictive Modeling for Multi-Source Healthcare Data With Missing Values. In: Proceedings of the 6th International Conference on Digital Health Conference. 2016 Presented at: DH'16; April 11-13, 2016; Montréal, Québec, Canada p. 47-54. [doi: 10.1145/2896338.2896347]
183. Abdullah NN, Clancey WJ, Raj A, Zain AZ, Khalid KF, Ooi A. Application of a Double-Loop Learning Approach for Healthcare Systems Design in an Emerging Market. In: Proceedings of the International Workshop on Software Engineering in Healthcare Systems. 2018 Presented at: SEHS'18; May 28, 2018; Gothenburg, Sweden p. 10-13. [doi: 10.1145/3194696.3194698]
184. Cao Y, Hou P, Brown D, Wang J, Chen S. Distributed Analytics and Edge Intelligence: Pervasive Health Monitoring at the Era of Fog Computing. In: Proceedings of the 2015 Workshop on Mobile Big Data. 2015 Presented at: Mobidata'15; June 21, 2015; Hangzhou, China p. 43-48. [doi: 10.1145/2757384.2757398]
185. O'Leary P, Carroll N, Clarke P, Richardson I. Untangling the Complexity of Connected Health Evaluations. In: Proceedings of the 2015 International Conference on Healthcare Informatics. 2015 Presented at: ICHI'15; October 21-23, 2015; Dallas, TX, USA p. 272-281. [doi: 10.1109/ichi.2015.39]
186. Davey SM, Brennan M, Meenan BJ, McAdam R. A Framework to Manage the Innovation Strategies of New Technology Based Firms. In: First International Technology Management Conference. 2011 Presented at: ICE'11; June 27-30, 2011; San Jose, CA, USA p. 1007-1013. [doi: 10.1109/itmc.2011.5995997]
187. Al Disi M, Djelouat H, Amira A, Bensaali F. The Accuracy and Efficacy of Real-Time Compressed ECG Signal Reconstruction on a Heterogeneous Multicore Edge-Device. In: Proceedings of the 21st Euromicro Conference on Digital System Design. 2018 Presented at: DSD'18; August 29-31, 2018; Prague, Czech Republic p. 458-463. [doi: 10.1109/dsd.2018.00082]
188. Nyberg T, Xiong G, Luostarinen J. Connected Health Services Internet, Mobile and Wireless Technologies in Healthcare. In: Proceedings of 2011 IEEE International Conference on Service Operations, Logistics and Informatics. 2011 Presented at: SOLI'11; July 10-12, 2011; Beijing, China p. 220-224. [doi: 10.1109/soli.2011.5986559]
189. Barbosa P, Leite F, Santos D, Figueiredo A, Galdino K. Introducing Traceability Information Models in Connected Health Projects. In: Proceedings of the 31st International Symposium on Computer-Based Medical Systems. 2018 Presented at: CNMS'18; July 23, 2018; Karlstad, Sweden p. 18-23. [doi: 10.1109/cbms.2018.00011]
190. Saranummi N. In the spotlight: health information systems--PHR and value based healthcare. IEEE Rev Biomed Eng 2009;2:15-17. [doi: 10.1109/RBME.2009.2034699] [Medline: 22275039]
191. Parvin S, Gawanmeh A, Venkatraman S, Alwadi A, Al-Karaki JN. Efficient Lightweight Mechanism for Node Authentication in WBSN. In: Advances in Science and Engineering Technology International Conferences (ASET). 2018 Presented at: ASET'18; February 6-April 5, 2018; Abu Dhabi, United Arab Emirates p. 1-6. [doi: 10.1109/icaset.2018.8376827]
192. Martins AF, Santos DF, Perkusich A, Almeida HO. IEEE 11073 And Connected Health: Preparing Personal Health Devices for the Internet. In: Proceedings of the International Conference on Consumer Electronics. 2014 Presented at: ICCE'14; January 10-13, 2014; Las Vegas, NV, USA p. 274-275. [doi: 10.1109/icce.2014.6776001]

193. Nambiar AR, Reddy N, Dutta D. Connected Health: Opportunities and Challenges. In: Proceedings of the International Conference on Big Data. 2017 Presented at: Big Data'17; December 11-14, 2017; Boston, MA, USA p. 1658-1662. [doi: 10.1109/bigdata.2017.8258102]
194. Gawanmeh A. Open Issues in Reliability, Safety, and Efficiency of Connected Health. In: Proceedings of the First International Conference on Connected Health: Applications, Systems and Engineering Technologies. 2016 Presented at: CHASE'16; June 27-29, 2016; Washington, DC, USA p. 1-6. [doi: 10.1109/chase.2016.60]
195. Stenhaug M, Johansen HD, Johansen D. Transforming Healthcare Through Life-Long Personal Digital Footprints. In: Proceedings of the First International Conference on Connected Health: Applications, Systems and Engineering Technologies. 2016 Presented at: CHASE'16; June 27-29, 2016; Washington, DC, USA p. 36-41. [doi: 10.1109/chase.2016.54]
196. Ali AA, Zhai X, Amira A, Bensaali F, Ramzan N. Heterogeneous Implementation of ECG Encryption and Identification on the Zynq SoC. In: Proceedings of the 24th Annual International Symposium on Field-Programmable Custom Computing Machines. 2016 Presented at: FCCM'16; May 1-3, 2016; Washington, DC, USA p. 139. [doi: 10.1109/fccm.2016.44]
197. Gomes YF, Santos DF, Almeida HO, Perkusich A. Integrating MQTT and ISO/IEEE 11073 For Health Information Sharing in the Internet of Things. In: Proceedings of the International Conference on Consumer Electronics. 2015 Presented at: ICCE'15; January 9-12, 2015; Las Vegas, NV, USA p. 200-201. [doi: 10.1109/icce.2015.7066380]
198. Bellekens X, Hamilton A, Seeam P, Nieradzinska K, Franssen Q, Seeam A. Pervasive eHealth Services a Security and Privacy Risk Awareness Survey. In: Proceedings of the International Conference On Cyber Situational Awareness, Data Analytics And Assessment. 2016 Presented at: CyberSA'16; June 13-14, 2016; London, UK p. 1-4. [doi: 10.1109/cybersa.2016.7503293]
199. Anumala H, Busetty SM. Distributed Device Health Platform Using Internet of Things Devices. In: Proceedings of the International Conference on Data Science and Data Intensive Systems. 2015 Presented at: DSDIS'15; December 11-13, 2015; Sydney, NSW, Australia p. 525-531. [doi: 10.1109/dsdis.2015.110]
200. O'Quigley C, Sabourin M, Coyle S, Connolly J, Condall J, Curran K, et al. Characteristics of a Piezo-Resistive Fabric Stretch Sensor Glove for Home-Monitoring of Rheumatoid Arthritis. In: Proceedings of the 11th International Conference on Wearable and Implantable Body Sensor Networks Workshops. 2014 Presented at: BSN'14; June 16-19, 2014; Zurich, Switzerland p. 23-26. [doi: 10.1109/bsn.workshops.2014.15]
201. Baali H, Zhai X, Djelouat H, Amira A, Bensaali F. Inequality indexes as sparsity measures applied to ventricular ectopic beats detection and its efficient hardware implementation. IEEE Access 2018;6:9464-9472. [doi: 10.1109/access.2017.2780190]
202. Antoniou PE, Rivera-Romero O, Karagianni M, Bamidis PD. Towards Evidence Based M-Health Application Design in Cancer Patient Healthy Lifestyle Interventions. In: Proceedings of the 30th International Symposium on Computer-Based Medical Systems. 2017 Presented at: CBMS'17; June 22-24, 2017; Thessaloniki, Greece p. 690-695. [doi: 10.1109/cbms.2017.117]
203. Asim M, Petkovi M, Qu M, Wang C. An Interoperable Security Framework for Connected Healthcare. In: Proceedings of the Consumer Communications and Networking Conference. 2011 Presented at: CCNC'11; January 9-12, 2011; Las Vegas, NV, USA p. 116-120. [doi: 10.1109/ccnc.2011.5766361]
204. Maramis C, Chouvarda I, Maglaveras N, Isomursu M. Introducing a Framework for Reporting Behavioral Informatics Interventions. In: Proceedings of the International Conference on Telecommunications. 2017 Presented at: ConTEL'17; June 28-30, 2017; Zagreb, Croatia p. 63-68. [doi: 10.23919/contel.2017.8000040]
205. Cao Y, Chen S, Hou P, Brown D. FAST: A Fog Computing Assisted Distributed Analytics System to Monitor Fall for Stroke Mitigation. In: Proceedings of the International Conference on Networking, Architecture and Storage. 2015 Presented at: NAS'15; August 6-7, 2015; Boston, MA, USA p. 2-11. [doi: 10.1109/nas.2015.7255196]
206. Silvello A. How connected insurance is reshaping the health insurance industry. Stud Health Technol Inform 2018;251:179-182. [Medline: 29968632]
207. Tonheim AN, Babic A. Assessing information needs for a personal multiple sclerosis application. Stud Health Technol Inform 2018;247:486-490. [Medline: 29678008]
208. Vaughan PW. Quality and innovation: redesigning a coordinated and connected health system. Healthc Pap 2017;16(3):35-39. [doi: 10.12927/hcpap.2017.25082] [Medline: 28671543]
209. Van Alstin CM. Virtual reality: coming to a hospital near you. VR may take connected health to a whole new level. Health Manag Technol 2016 Oct;37(3):6-9. [Medline: 29480678]
210. Skiba DJ. Connected health 2015: the year of virtual patient visits. Nurs Educ Perspect 2015;36(2):131-133. [Medline: 29194142]
211. Leskinen H, Korpelainen J, Mäkiniemi M. The university hospital as a hub for connected health: Oulu Healthworld, Northern Scandinavia. World Hosp Health Serv 2012;48(4):20-24. [Medline: 23484430]
212. Glaser J. Tipping point. The arrival of connected health. J Healthc Inf Manag 2008;22(4):21-23. [Medline: 19267015]
213. Mathur A, Kvedar JC, Watson AJ. Connected health: a new framework for evaluation of communication technology use in care improvement strategies for type 2 diabetes. Curr Diabetes Rev 2007 Nov;3(4):229-234. [doi: 10.2174/157339907782330003] [Medline: 18220678]

214. Whitlinger D, Ayyagari D, McClure D, Fisher J, Lopez F. Straight talk. Collaboration fosters connected health: a new paradigm of proactive heathcare. Mod Healthc 2007 May 21;37(21):47-50. [Medline: 17824192]
215. Gerdsen F, Müeller S, Jablonski S, Prokosch HU. Standardized exchange of medical data between a research database, an electronic patient record and an electronic health record using CDA/SCIPHOX. AMIA Annu Symp Proc 2005:963 [FREE Full text] [Medline: 16779250]
216. Noroozi B, Morshed BI. Coil Distance and Angle Misalignment Effects on the Mutual Inductance for 13.56 MHz WRAP Sensors. In: Proceedings of the United States National Committee of URSI National Radio Science Meeting. 2018 Presented at: USNC-URSI NRSM'18; January 4-7, 2018; Boulder, CO, USA p. 1-2. [doi: 10.1109/usnc-ursi-nrsm.2017.7878304]
217. Rahimi N, Ibarra M. A Review of Multiple User Center Design Methods for New Product Development in Smart and Connected Health Applications. In: Proceedings of the Conference on Portland International Center for Management of Engineering and Technology; Infrastructure and Service Integration. 2014 Presented at: PICMET'14; July 27-31, 2014; Kanazawa, Japan p. 3498-3510.
218. Pandas. Python Data Analysis Library URL: https://pandas.pydata.org/ [accessed 2019-07-16]
219. Matplotlib: Python Plotting. URL: https://matplotlib.org/ [accessed 2019-07-16]
220. Montesi M, Owen JM. From conference to journal publication: how conference papers in software engineering are extended for publication in journals. J Am Soc Inf Sci 2008 Mar;59(5):816-829. [doi: 10.1002/asi.20805]
221. Knight LV, Steinbach TA. Selecting an appropriate publication outlet: a comprehensive model of journal selection criteria for researchers in a broad range of academic disciplines. Int J Doctoral Stud 2008;3:59-79. [doi: 10.28945/3289]
222. Donovan SK. The importance of resubmitting rejected papers. J Sch Publ 2007;38(3):151-155. [doi: 10.1353/scp.2007.0013]
223. CATCH ITN - eHealth for Cancer Patients. URL: https://www.catchitn.eu [accessed 2019-07-16]
224. Community Research and Development Information Service. Connected Health Early Stage Researcher Support System URL: https://cordis.europa.eu/project/rcn/198323/factsheet/en [accessed 2019-07-16]
225. European Connected Health Alliance | (ECHAlliance). URL: https://echalliance.com/ [accessed 2019-07-16]
226. World Health Organization. 2015. Ageing and Life-Course: World Report on Ageing and Health 2015 URL: http://www.who.int/ageing/events/world-report-2015-launch/en/ [accessed 2019-07-16]
227. Bloom DE, Canning D, Fink G. Implications of population ageing for economic growth. Oxf Rev Econ Policy 2011 May 13;26(4):583-612. [doi: 10.1093/oxrep/grq038]
228. Farahani B, Firouzi F, Chang V, Badaroglu M, Constant N, Mankodiya K. Towards fog-driven IoT ehealth: promises and challenges of IoT in medicine and healthcare. Future Gener Comp Sy 2018 Jan;78:659-676. [doi: 10.1016/j.future.2017.04.036]
229. Statista. Number of Smartphones Sold to End Users Worldwide From 2007 to 2020 (in Million Units) URL: https://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/ [accessed 2019-07-16]
230. Statista. Forecast Wearables Unit Shipments Worldwide From 2014 to 2023 (in Millions) URL: https://www.statista.com/statistics/437871/wearables-worldwide-shipments/ [accessed 2019-07-16]
231. Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I. Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. Future Gener Comp Sy 2009 Jun;25(6):599-616. [doi: 10.1016/j.future.2008.12.001]
232. World Health Organization. Global Diffusion of eHealth: Making Universal Health Coverage Achievable Report of the Third Global Survey on eHealth. Geneva, Switzerland: World Health Organization; 2016.
233. Rizzatti L. EE Times | Electronic Engineering Times. 2016. Digital Data Storage is Undergoing Mind-Boggling Growth URL: https://www.eetimes.com/author.asp?section_id=36&doc_id=1330462 [accessed 2019-07-16]
234. Eisenhardt KM. Building theories from case study research. Acad Manage Rev 1989 Oct 1;14(4):532-550. [doi: 10.5465/AMR.1989.4308385]
235. Neergaard H, Ulhøi JP. Handbook of Qualitative Research Methods in Entrepreneurship. Cheltenham, United Kingdom: Edward Elgar Publishing; 2007.
236. Yin RK. Case Study Research: Design and Methods. Thousand Oaks, CA: Sage Publications; 2009.
237. Kerlinger FN. Foundations of Behavioural Research: Educational, Psychological and Sociological Enquiry. New York, United States: Holt, Rinehart and Winston Publishers; 1986.
238. Klein JT. Interdisciplinarity: History, Theory, and Practice. Detroit, Michigan, United States: Wayne State University Press; 1990.
239. Bradbeer J. Barriers to interdisciplinarity: disciplinary discourses and student learning. J Geogr High Educ 1999 Nov;23(3):381-396. [doi: 10.1080/03098269985326]
240. Wagner KH, Brath H. A global view on the development of non communicable diseases. Prev Med 2012 May(54 Suppl):S38-S41. [doi: 10.1016/j.ypmed.2011.11.012] [Medline: 22178469]
241. Freyne J, Coyle L, Smyth B, Cunningham P. Relative status of journal and conference publications in computer science. Commun ACM 2010 Nov 1;53(11):124-132. [doi: 10.1145/1839676.1839701]
242. Garousi V, Mesbah A, Betin-Can A, Mirshokraie S. A systematic mapping study of web application testing. Inf Softw Technol 2013 Aug;55(8):1374-1396. [doi: 10.1016/j.infsof.2013.02.006]

243. Giunti G, Guisado-Fernandez E, Belani H, Lacalle-Remigio JR. Mapping the access of future doctors to health information technologies training in the European union: cross-sectional descriptive study. J Med Internet Res 2019 Aug 12;21(8):e14086 [FREE Full text] [doi: 10.2196/14086] [Medline: 31407668]

Abbreviations

BMC: BioMed Central
CATCH: cancer: activating technology for connected health
CH: connected health
CHESS: connected health early stage researcher support system
COST: Cooperation in Science and Technology
DOI: digital object identifier
EC: exclusion criteria
eHealth: electronic health
JMIR: Journal of Medical Internet Research
MQ: mapping questions
PHR: personal health record
QJM: Quarterly Journal of Medicine.

Edited by M Tilney; submitted 15.04.19; peer-reviewed by I Mircheva, B Chaudhry, C Perrin, MS Aslam, M Reynolds; comments to author 20.06.19; revised version received 27.07.19; accepted 18.08.19; published 25.09.19.

Please cite as:
Karampela M, Isomursu M, Porat T, Maramis C, Mountford N, Giunti G, Chouvarda I, Lehocki F
The Extent and Coverage of Current Knowledge of Connected Health: Systematic Mapping Study
J Med Internet Res 2019;21(9):e14394
URL: http://www.jmir.org/2019/9/e14394/
doi:10.2196/14394
PMID:31573915

Review

Health Professions Digital Education on Antibiotic Management: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration

Bhone Myint Kyaw[1*], MBBS, MSc, PhD; Lorainne Tudor Car[2,3*], MD, MSc, PhD; Louise Sandra van Galen[1,4], MD, PhD; Michiel A van Agtmael[4,5], MD, PhD; Céire E Costelloe[3], BA, MSc, PhD; Onyema Ajuebor[6], MBBS, MSc; James Campbell[6], MPH, MSc; Josip Car[1,3], MD, PhD, FRCP, FFPH

[1]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[2]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[3]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom
[4]Department of Internal Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
[5]Research & Expertise Center In Pharmacotherapy Education, Amsterdam, Netherlands
[6]Health Workforce Department, World Health Organization, Geneva, Switzerland
[*]these authors contributed equally

Corresponding Author:
Lorainne Tudor Car, MD, MSc, PhD
Family Medicine and Primary Care
Lee Kong Chian School of Medicine
Nanyang Technological University
11 Mandalay Raod
Singapore, 308232
Singapore
Phone: 65 69041258
Email: lorainne.tudor.car@ntu.edu.sg

Abstract

Background: Inappropriate antibiotic prescription is one of the key contributors to antibiotic resistance, which is managed with a range of interventions including education.

Objective: We aimed to summarize evidence on the effectiveness of digital education of antibiotic management compared to traditional education for improving health care professionals' knowledge, skills, attitudes, and clinical practice.

Methods: Seven electronic databases and two trial registries were searched for randomized controlled trials (RCTs) and cluster RCTs published between January 1, 1990, and September 20, 2018. There were no language restrictions. We also searched the International Clinical Trials Registry Platform Search Portal and metaRegister of Controlled Trials to identify unpublished trials and checked the reference lists of included studies and relevant systematic reviews for study eligibility. We followed Cochrane methods to select studies, extract data, and appraise and synthesize eligible studies. We used random-effect models for the pooled analysis and assessed statistical heterogeneity by visual inspection of a forest plot and calculation of the I^2 statistic.

Results: Six cluster RCTs and two RCTs with 655 primary care practices, 1392 primary care physicians, and 485,632 patients were included. The interventions included personal digital assistants; short text messages; online digital education including emails and websites; and online blended education, which used a combination of online digital education and traditional education materials. The control groups received traditional education. Six studies assessed postintervention change in clinical practice. The majority of the studies (4/6) reported greater reduction in antibiotic prescription or dispensing rate with digital education than with traditional education. Two studies showed significant differences in postintervention knowledge scores in favor of mobile education over traditional education (standardized mean difference=1.09, 95% CI 0.90-1.28; I^2=0%; large effect size; 491 participants [2 studies]). The findings for health care professionals' attitudes and patient-related outcomes were mixed or inconclusive. Three studies found digital education to be more cost-effective than traditional education. None of the included studies reported on skills, satisfaction, or potential adverse effects.

Conclusions: Findings from studies deploying mobile or online modalities of digital education on antibiotic management were complementary and found to be more cost-effective than traditional education in improving clinical practice and postintervention knowledge, particularly in postregistration settings. There is a lack of evidence on the effectiveness of other digital education modalities such as virtual reality or serious games. Future studies should also include health care professionals working in settings other than primary care and low- and middle-income countries.

Clinical Trial: PROSPERO CRD42018109742; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=109742

(J Med Internet Res 2019;21(9):e14984) doi:10.2196/14984

KEYWORDS

digital education; antibiotic management; systematic review; meta-analysis; randomized controlled trial

Introduction

Antibiotic resistance is one of the most important public health concerns globally, and the magnitude of the problem is increasing [1]. The World Health Organization (WHO) warns that “without urgent action, we are headed for a post-antibiotic era, in which common infections and minor injuries can once again kill” [2]. Each year, in the United States, at least 2 million people become infected with bacteria that are resistant to antibiotics, and annually, 23,000 people die as a direct result of these infections [1,3]. Indirectly, antibiotic resistance contributes to the increasing costs of health care worldwide. Failure to address the issue of antibiotic resistance is estimated to result in excess of 10 million deaths worldwide and cost up to US $100 trillion annually by 2050 [4].

The main causes of antibiotic resistance include overprescription of antibiotics, overuse of antibiotics in livestock and fish farming, poor infection control in hospitals and clinics, patients not finishing their treatment, a lack of hygiene, and poor sanitation [5-7]. Adding to the problem is a shortage of health manpower and training facilities coupled with a lack of effective educational programs for health care professionals [8]. In line with the aforementioned challenges, the WHO supports countries through its competency framework on antimicrobial resistance (AMR) by providing guidance on the requisite competencies for health workforce learners at preregistration and postregistration levels in order to help address AMR in policy and practice settings [9].

The use of information and communication technology in education, ie, digital education, could improve antibiotic management education [10]. Digital education offers several potential advantages over traditional education methods, such as easier access to learning materials and facilitating the learning process without time or location constraints. Through its range of different modalities, digital education enables different educational experiences with varying forms and levels of interactivity, immersion, duration of the intervention, feedback, etc. Digital education also allows for uncomplicated and scalable dissemination of the latest evidence. For antimicrobial resistance, digital platforms are increasingly being used to deliver education and training, and there is potential to increase their uptake by, for example, ensuring that such programs are accredited and health care professionals are incentivized through various means [11,12]. As antibiotic resistance is the most prevalent form of antimicrobial resistance [13] and can potentially be prevented through health professions education [14,15], this paper focuses on the use different forms of digital education for antibiotic management by pre- and postregistration health care professionals.

Digital education (also known as electronic learning or digital learning) is the act of teaching and learning by means of digital technologies. It is an overarching term for an evolving multitude of educational approaches, concepts, methods, and technologies. Digital education can be further characterized by specific pedagogies and instructional methods, contexts of provision, and technical affordances of hardware and software [16]. It includes, but is not limited to, offline and online digital education [17-19], serious gaming and gamification [20], massive open online courses, virtual reality [21], virtual patient simulations [22], and mobile digital education [23] (Multimedia Appendix 1) [24-28].

Past reviews [29,30] on the effectiveness of educational interventions for the management of antibiotics focused mainly on traditional modes of intervention such as the use of text-based or paper-based interventions. To our knowledge, there are no reviews assessing the effectiveness of digital education for health professions on antibiotic management. To address this gap, we performed a systematic review on the effectiveness of digital technology either alone (single modality), combined with other digital technology (multimodalities), or combined with traditional education (blended digital education) to deliver education on antibiotic management for both pre- and postregistration health care professionals.

Methods

Search Strategy and Selection Criteria

We carried out a systematic review and meta-analysis by following Preferred Reporting Items for Systematic reviews and Meta-Analysis guidelines [31]. The protocol for this review was registered in the International Prospective Register of Systematic Reviews (PROSPERO, CRD42018109742) [32]. Randomized controlled trials (RCTs) and cluster RCTs (cRCTs) of pre- and postregistration health care professionals using any type of digital education (either standalone or blended) with any type of control (including traditional education and other forms of digital education) measuring change in clinical practice, knowledge, skills, attitudes, satisfaction, patient-centered outcomes (as primary outcomes); adverse-effects; or economic outcomes (as secondary outcomes) were eligible for inclusion in this review. We included participants and holders of the qualifications listed in the Health Field of Education and

Training (section 091) of the International Standard Classification of Education: Fields of Education and Training (United Nations Educational, Scientific and Cultural Organization) and excluded studies with participants from the field of traditional, alternative, and complementary medicine. We excluded crossover trials due to the high likelihood of carry-over effect. A detailed description of the methodology is provided in a previous study by Car et al [16].

We developed a comprehensive search strategy for Medline (Ovid), Embase (Ovid), Cochrane Central Register of Controlled Trials (CENTRAL, Wiley), PsycINFO (Ebsco), Educational Research Information Centre (Ebsco), Cumulative Index to Nursing and Allied Health Literature (Ebsco), and Web of Science Core Collection (Thomson Reuters; see Multimedia Appendix 2 for the Medline [Ovid] search strategy). Databases were searched from January 1, 1990, to September 20, 2018. We chose 1990 as the starting year for our search, as the use of computers was limited to basic tasks prior to this period [16]. We also searched the International Clinical Trials Registry Platform Search Portal and metaRegister of Controlled Trials to identify unpublished trials (see Multimedia Appendix 3 for the lists of Clinical Trial Registries).

The search results from each source were combined in a single library, and duplicate records were removed using reference management software [33]. Three reviewers (BK, GD, and LG) independently screened titles and abstracts of all records to identify potentially relevant studies. We retrieved full-text copies of articles deemed potentially relevant and independently assessed the retrieved articles against the eligibility criteria. Finally, we searched reference lists of all the studies that we deemed eligible for inclusion in our review and relevant systematic reviews.

Three reviewers (BK, GD, and LG) independently extracted relevant characteristics related to participants, intervention, comparators, outcome measures, and results from all the included studies using a standard data collection form. Any disagreements were resolved by consensus through discussion.

Data Analysis

For continuous outcomes, we reported mean values and SDs of the outcomes in each intervention group along with the number of participants and *P* values. For dichotomous outcomes, we reported risk ratio (RR) with 95% CIs. We were unable to identify a clinically meaningful interpretation of effect size in the literature for digital education interventions. Therefore, in line with other studies in the field, we presented outcomes using postintervention standardized mean difference (SMD) and interpreted the effect size using the Cohen rule [34,35].

Where studies reported more than one measure for each outcome, the primary measure, as defined by the study authors, was used in the analysis. Where no primary measure was reported, the measure that was most consistent with the outcomes presented in other included studies and the first measurement after the intervention were reported.

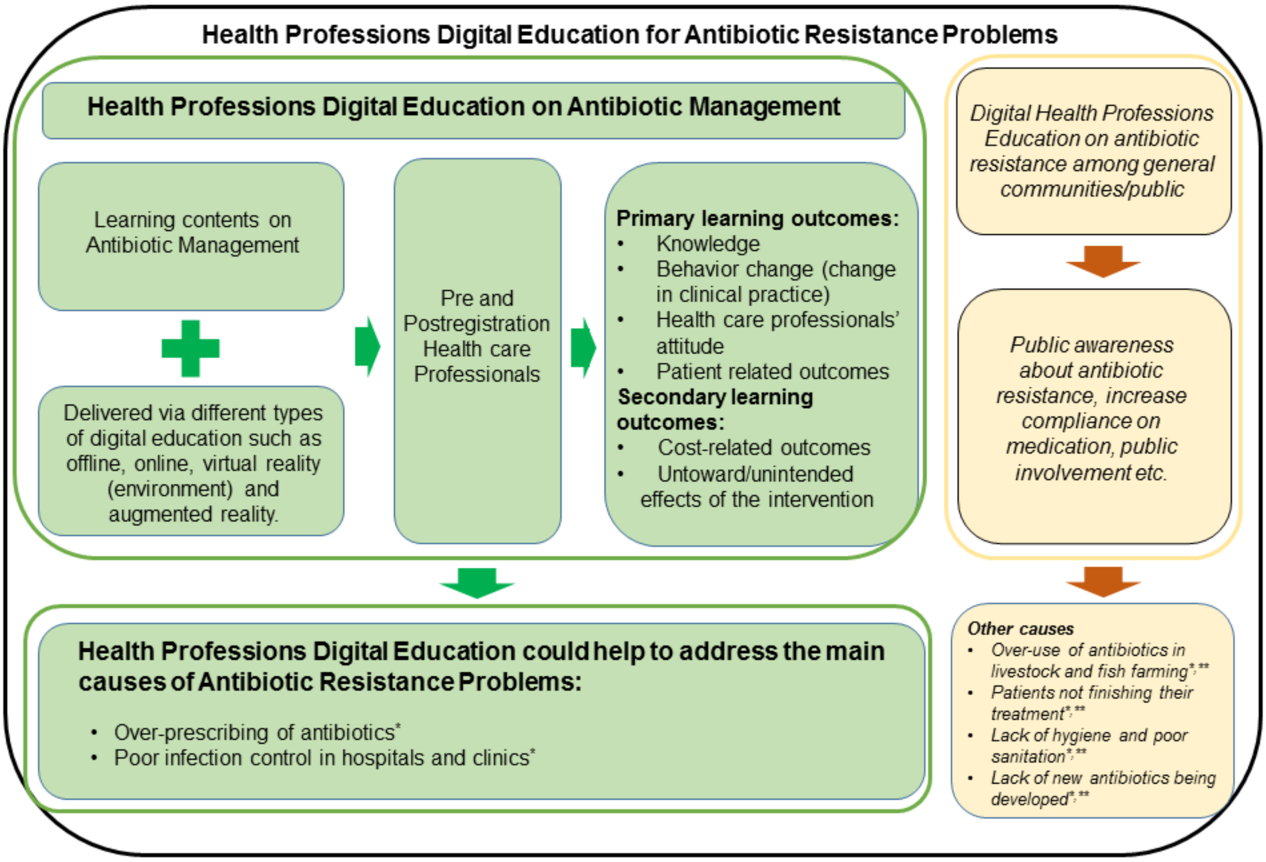
For cRCTs, we attempted to obtain data at the individual level. In cases where the statistical analysis of the cRCT had already been adjusted for data clustering, we extracted the reported effect estimates and used them directly for our analyses. If a meta-analysis of the included studies was indicated, we assessed statistical heterogeneity by visual inspection of a forest plot and by examining the I^2 statistic, with $I^2<25\%$, 25%-75%, and >75% representing a low, moderate, and high degree of heterogeneity, respectively [35]. We used a narrative approach to data synthesis instead of reporting pooled results from meta-analysis if we detected substantial clinical, methodological, or statistical heterogeneity across included studies.

For meta-analysis, we used a random-effects model. Continuous outcome data were presented in the form of standardized mean difference (SMD) (due to the use of different scales), along with 95% CIs. In the analysis of continuous outcomes, we used the inverse variance method. We displayed the result of the meta-analysis in a forest plot that provided effect estimates and 95% CIs for each individual study as well as a pooled effect estimate and 95% CI. We performed meta-analysis using Review Manager Software 5.3 [36]. We adhered to the statistical guidelines described by Higgins et al [35].

Three reviewers (BK, GD, and LG) independently assessed the methodological risk of bias of included studies in accordance with the Cochrane methodology [35]. The following individual elements of RCTs were assessed: random sequence generation, allocation concealment, blinding (outcome assessment), completeness of outcome data (attrition bias), selective outcome reporting (relevant outcomes reported), and other sources of bias (baseline imbalances). For cRCTs, we assessed the risk of the following additional domains: recruitment bias, baseline imbalance, loss of clusters, incorrect analysis, and comparability with individually randomized trials as recommended by Puffer et al [37]. Judgments concerning risk of bias for each study were scored as high, low, or unclear. Studies were judged to be at high risk of bias if there was a high risk of bias for one or more domains or unclear risk of bias in three or more domains. Studies were judged to be at unclear risk of bias if there was an unclear risk of bias for two domains. We incorporated the results of the risk of bias assessment into the review using risk of bias tables, a graph, and a narrative summary.

We developed a conceptual framework on how different digital education interventions focused on antibiotic management could contribute to an improvement in antibiotic-prescribing practices and a reduction in antibiotic resistance (Figure 1). The aim of the framework was to enable an improvement of health care professionals' knowledge and skills (as well as other outcomes including patient-related outcomes) by empowering health professions education via digital education. We hope that by facilitating the delivery methods in health professions education, learning outcomes will be improved.

Figure 1. A simplified conceptual framework to address antibiotic resistance by empowering health professions education through digital education. *Causes of antibiotic resistance as per the World Health Organization (2018) data [2]. **Knight Gwenan et al (2018) [38].

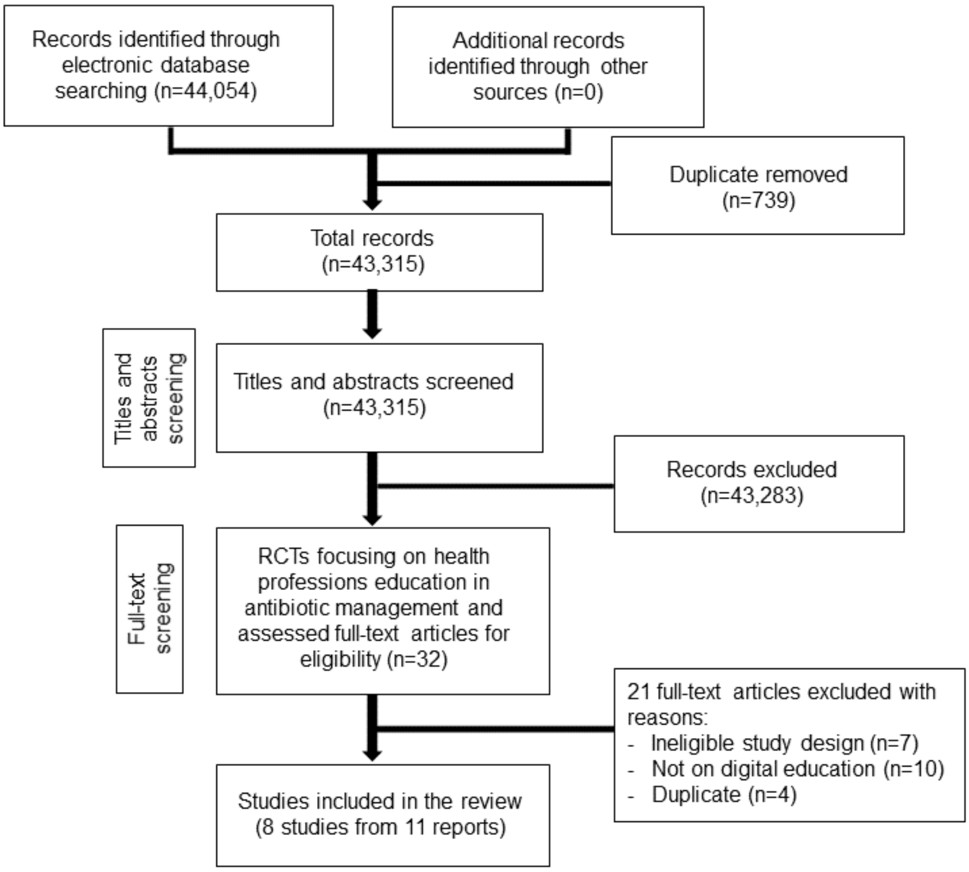


Results

Overview

Our search strategy retrieved 44,054 unique references (Figure 2). We included 8 studies from 11 reports involving 655 health care practices, 1392 primary care physicians, and 485,632 patients [39-49]. Six studies were cRCTs [41-45,47], and two were RCTs [48,49]. Six studies randomized the participants into two groups, and two studies randomized participants into four groups. All participants were postregistration primary care physicians. The number of participants across the studies ranged from 12 [49] to 479 [47]. Two studies provided additional information for patients and caregivers in the form of leaflets or booklets as a part of the intervention [43,45]. Dekker et al [45] assessed the effects of an intervention that consisted of online digital education for primary care physicians and an information booklet for patients or caregivers. Little et al [43] delivered online digital education targeting primary care physicians and provided interactive booklets as additional resources for physicians to use during consultations with patients.

Figure 2. Study selection process. RCT: randomized controlled trial.



The content of the interventions focused on different aspects of antibiotic management ranging from a shared decision-making training program for antibiotic prescription in acute respiratory infection [44], management of antibiotic prescription in acute respiratory infections [43,45], inappropriate prescribing in primary care [41], a multifaceted education program to reduce antibiotic dispensing in primary care [42,48], and an antibiotic decision management guide [49].

All studies were conducted in high-income countries, except for one study that was conducted in an upper–middle-income country (China) [47]. Of the seven studies conducted in high-income countries, two were performed in the United States [41,49], two were solely performed in the United Kingdom [42,48], one was performed in six European countries (England, Wales, Belgium, the Netherlands, Spain, and Poland) [43], one was performed in Canada [44], and one was performed in the Netherlands [45]. All studies were conducted in clinical settings: either primary care or hospital/clinics. All studies received funding support from nonindustrial sponsorship, except one study where the authors received funding support from companies (Johnson & Johnson and "Research-In-Motion corporations") [49].

For the intervention groups, three studies used blended digital education (ie, online plus traditional education such as face-to-face interactive workshops; group discussions and reflections, printed leaflets, posters) [42,44,48], three studies used online digital education, which involved online training materials as well as emails with feedback and suggestions [41,43,45]; and two studies used mobile digital education devices, namely, a mobile phone or personal digital assistant [47,49]. The duration of the interventions ranged from 1 hour [42] to 18 months [41]. Three studies did not provide information about the duration of the interventions [43,45,47].

For the control groups, one study used a traditional continuing medical education program, while the other seven studies employed usual practice (ie, participants received no additional intervention other than usual on the job training). Additionally, one study compared different forms of online digital education programs to traditional education [43].

The included studies reported outcomes such as clinical practice (n=6 studies), patient-related outcomes (n=4 studies), knowledge (n=2 studies), health care professionals' attitude toward the intervention (n=2 studies), and economic outcomes (n=3 studies). None of the studies reported adverse or unintended effects of interventions. Table 1 presents a detailed description of the characteristics of the included studies.

Table 1. Characteristics of included studies.

Study author, year, study design, country	Population (n)	Setting and source of funding	Field of study	Intervention type	Duration of the intervention	Control	Outcome(s) assessed
Bochicchio, 2006, RCT[a], US [49]	12 primary care physicians	Primary care setting, industrial funding (Johnson & Johnson, Research-In-Motion Corporation)	Antibiotic decision management guide	Mobile digital education	3 months	Traditional education (usual practice)	Knowledge
Butler, 2012, RCT, UK [48]	68 practices, 263 primary care physicians, and 480,000 patients (approximate number)	Primary care setting, nonindustrial funding	Multifaceted educational program to reduce antibiotic dispensing in primary care	Online blended education (online digital education plus traditional education)	1.5 hours	Traditional education (usual practice)	Clinical practice improvement; patient-related outcomes; economic outcome(s)
Chen, 2014, cRCT[b], China [47]	100 practices and 479 primary care physicians	Primary care setting, nonindustrial funding	Management of upper respiratory infection	Mobile digital education	N/A[c]	Traditional education (traditional CME[d] program)	Knowledge; clinical practice improvement; health care professionals' attitude; economic outcome(s)
Dekker, 2018, cRCT, the Netherlands [45]	35 practices, 75 primary care physicians, and 1009 patients	Primary care setting, nonindustrial funding	Antibiotic prescription in acute respiratory infection	Online digital education	N/A	Traditional education (usual practice)	Clinical practice improvement; patient-related outcomes
Legare, 2012, cRCT, Canada [44]	9 practices, 149 primary care physicians, and 359 patients	Primary care setting, nonindustrial funding	Antibiotic prescription in acute respiratory infection	Online blended education (online digital education plus traditional education)	4 hours	Traditional education (usual practice)	Patient-related outcomes
Meeker, 2016, cRCT, USA [41]	47 practices and 248 primary care physicians	Primary care setting, nonindustrial funding	Antibiotic prescription among primary care practices	Online digital education	>18 months	Traditional education (usual practice)	Clinical practice improvement
McNulty, 2018, cRCT, UK [42]	150 practices and 166 primary care physicians	Primary care setting, nonindustrial funding	Antibiotic dispensing in primary care	Online blended education (online digital education plus traditional education)	1 hour	Traditional education (usual practice)	Clinical practice improvement
Little, 2013, cRCT, six European countries (England, Wales, Belgium, the Netherlands, Spain, and Poland) [43]	246 practices and 4264 patients	Primary care setting, nonindustrial funding	Antibiotic prescription in acute respiratory infection	Online digital education (CRP[e] training)	N/A	Traditional education (usual practice)	Clinical practice improvement; health care professionals' attitude; patient-related outcomes; economic outcome(s)
Yardley 2013, cRCT, six European countries (as mentioned above) [39]	246 practices and 4264 patients	Primary care setting, nonindustrial funding	Antibiotic prescription in acute respiratory infection	Online digital education (enhanced-communication training)	N/A	Traditional education (usual practice)	Clinical practice improvement; health care professionals' attitude; patient-related outcomes; economic outcome(s)

Study author, year, study design, country	Population (n)	Setting and source of funding	Field of study	Intervention type	Duration of the intervention	Control	Outcome(s) assessed
Oppong 2018, cRCT, six European countries (as mentioned above) [40]	246 practices and 4264 patients	Primary care setting, nonindustrial funding	Antibiotic prescription in acute respiratory infection	Online digital education (combined CRP and enhanced-communication training)	N/A	Traditional education (usual practice)	Clinical practice improvement; health care professionals' attitude; patient-related outcomes; economic outcome(s

[a] RCT: randomized controlled trial.

[b] cRCT: cluster randomized controlled trial.

[c] N/A: not applicable.

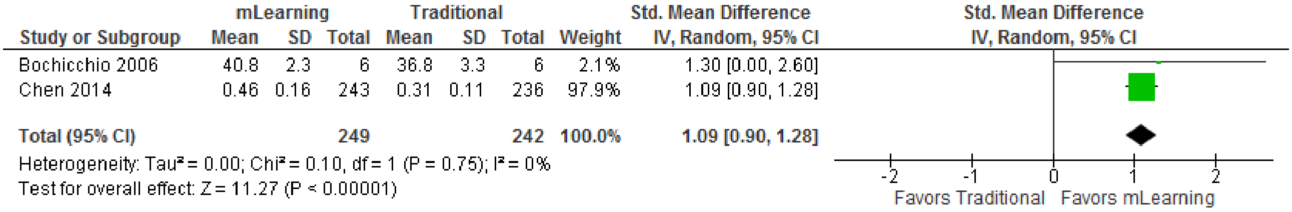
[d] CME: continuing medical education.

[e] CRP: C-reactive protein.

Knowledge Outcome (Postintervention)

Two studies (491 participants) assessed knowledge as the primary outcome and evaluated it with multiple choice questionnaires. One study assessed short-term postintervention knowledge [47], and another study assessed both postintervention knowledge as well as knowledge retention at 3 months postintervention [49]. The pooled estimate suggests that the intervention (mobile phone or personal digital assistant device [mobile digital education]) may improve postintervention knowledge scores as compared to traditional education (SMD=1.09, 95% CI 0.90-1.28; I^2=0%; 491 participants [2 studies], large effect size; Figure 3).

Figure 3. Forest plot for postintervention knowledge outcome (mobile digital education vs traditional education). mLearning: mobile learning; IV: inverse variance.



Clinical Practice Improvement

Six studies assessed the effectiveness of online digital education (either alone or in a blended format) and mobile digital education modalities in reducing antibiotic prescribing or dispensing rate [41-43,45,47,48]. Out of six studies, four studies [41,43,45,48] favored digital education and two studies [42,47] reported no difference in the effectiveness between digital education and traditional education. Because of high heterogeneity across the studies, the results of these studies are presented narratively.

Little et al [43] assessed the effectiveness of different forms of online digital education in comparison to traditional education and reported that the combined intervention (online training on point of care C-reactive protein [CRP] test and enhanced-communication skills) was associated with the greatest reduction in prescribing rate (combined intervention: RR=0.38, 95% CI 0.25-0.55, *P*<.001; online-based CRP training: RR=0.53, 95% CI 0.36-0.74, *P*<.001; enhanced communication skills training: RR=0.68, 95% CI 0.5-0.89, *P*=.003) compared to traditional education. Meeker et al [41] assessed the use of online digital education (peer comparison via email with feedback and suggestions for improving performance) in comparison to traditional education and reported a greater reduction in antibiotic-prescribing rates in the intervention group compared to the control (difference in mean change score: –5.2%, 95% CI −6.9 to −1.6]; *P*<.001) [41]. Similarly, Dekker et al [45] reported that the antibiotic prescribing rate was lower in the online digital education group (received online learning resources) compared to the traditional education group (RR=0.65, 95% CI 0.46-0.91) [45]. Butler et al [48] reported a reduction in the total oral antibiotic dispensing rate for the study year in online blended education (4.2%, 95% CI 0.6-7.7; *P*=.02) compared to traditional education [48]. Chen et al [47] assessed the effectiveness of mobile phone text messaging in comparison to traditional education and reported no significant difference in antibiotic-prescribing rates postintervention between the groups (RR=1.02, 95% CI 0.94-1.1; *P*=.63) [47]. McNulty et al [42] compared online blended education to usual practice and found no significant difference in the antibiotic dispensing rates between the groups (dispensing rate ratio 0.973, 95% CI 0.945-1.001; *P*=.06) [42].

In general, the findings suggest that, compared to traditional education, a variety of different forms of online digital education (eg, email with feedback and suggestions, studying online learning resources, and online training on point-of-care CRP test) may improve primary care physicians’ clinical practice and reduce antibiotic prescriptions or dispensing rates. The combined training of online-based point-of-care CRP test and communication skills was found to be more effective than other forms of online digital education and traditional education in reducing the antibiotic-prescribing rate among primary care practice.

Primary Care Physicians' Attitude Toward the Intervention

Two studies assessed postintervention attitudes toward different forms of digital education on the topic of antibiotic prescriptions. Chen et al [47] assessed only the attitude of primary care physicians in the intervention group toward the intervention (ie, text messages containing evidence-based recommendations). One-third of the participants in the intervention group reported that they frequently adopted the recommendations in their clinical decision making, and 95% wanted to continue receiving the text messages. Little et al [43] compared three different forms of online training on antibiotic prescription and assessed primary care physicians' attitudes towards the intervention. The study reported no difference between the participants of different online training groups regarding their perceptions of any potential damage that the intervention could have had on their relationships with patients and their confidence in reducing prescription Multimedia Appendix 4) [43].

Patient-Related Outcomes

Four studies assessed diverse patient-related outcomes [43-45,48] and reported mixed or inconclusive findings. Butler et al [48] reported there was no difference in the postintervention hospital admission rate (–1.9% reduction in the intervention group, 95% CI –13.2 to 8.2; *P*=.72) and reconsultation rate in the 7-day postintervention median scores (MD=–0.65, 95% CI –1.69 to 0.55; *P*=.45) between the intervention (ie, online blended digital education) and the traditional education (ie, usual practice) groups.

Dekker et al [45] reported that the reconsultation rate for children within the same disease episode was lower in the intervention group receiving online digital education than the group receiving traditional education (RR=0.66, 95% CI 0.38-1.16). However, the same study reported that the probability of consultation for a new respiratory tract infection within 6 months (RR=1.06, 95% CI 0.72-1.58) and hospital referral (RR=0.66, 95% CI 0.31-1.40) did not differ significantly.

Legare et al [44] reported that an online-based shared decision-making program enhanced patient participation in decision making and led to fewer patients deciding to use antibiotics for acute respiratory infections compared to traditional education (RR=0.48, 95% CI 0.34-0.68). However, there was no difference in the effects between the two groups regarding other patient-related outcomes such as adherence to the decision (SMD=–0.82, 95% CI –2.23 to 0.59), repeated consultations (SMD=0.80, 95% CI –0.60 to 2.20), decisional regret (SMD=0.25, 95% CI –1.07 to 1.57), quality of life (SMD=0.04, 95% CI –1.27 to 1.36), and intention to engage in shared decision making in future consultations regarding the use of antibiotics for acute respiratory infections (SMD= 0.16, 95% CI –1.16 to 1.47).

Little et al [43] reported that compared to usual practice, online digital education about CRP training had little or no difference in postintervention patient-related outcomes such as patient enablement (SMD=–0.11, 95% CI –0.24 to 0.01), patients' satisfaction with consultation (SMD=–0.09, 95% CI –0.22 to 0.03), and antibiotic usage (SMD=0.13, 95% CI 0.01-0.26). Similarly, the study reported no difference in postintervention patient outcomes with the use of online training on enhanced communication skills compared to traditional education (Multimedia Appendix 4).

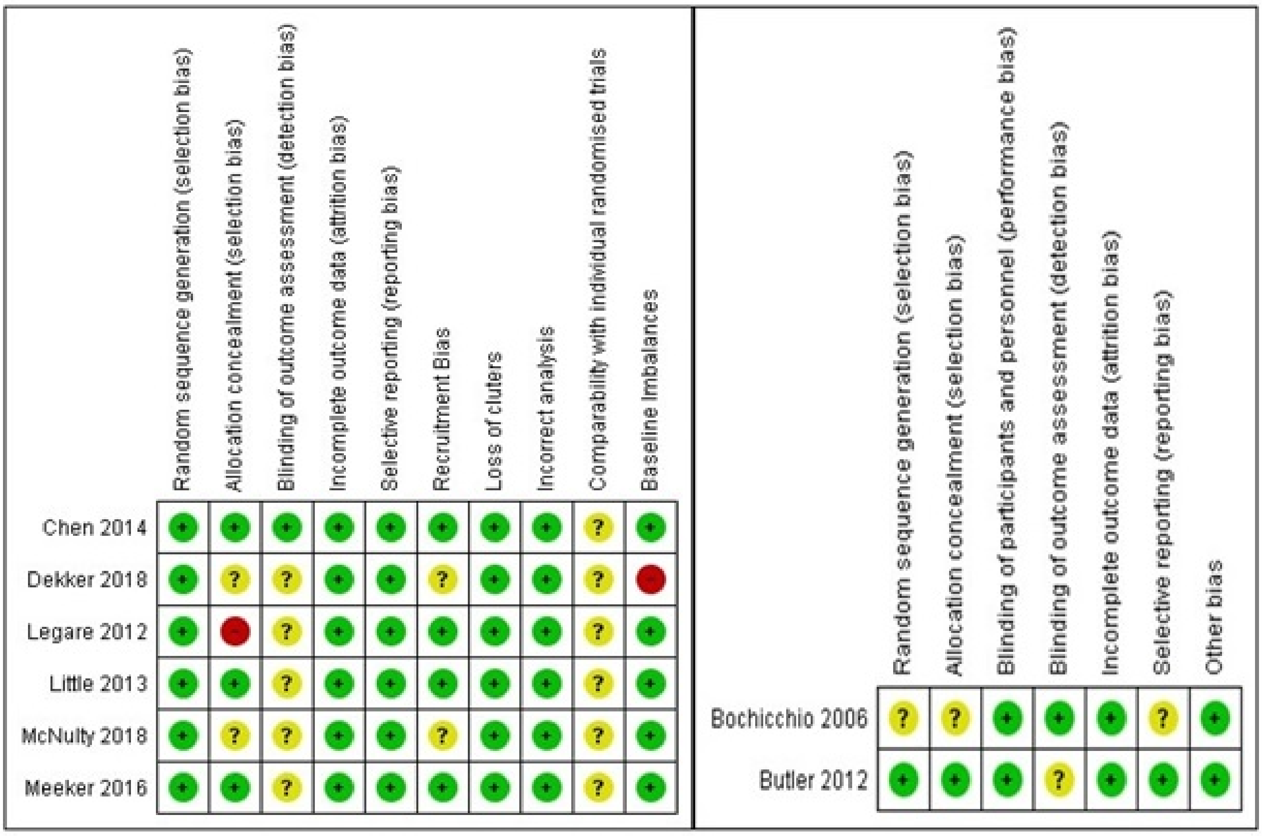
Economic Outcomes

Three studies reported that the intervention costs were lower than those of traditional education [43,47,48]. Butler et al [48] assessed the effects of online blended education compared to traditional education and reported that the mean cost of the program was £2923 (~US $4559) per practice (SD £1187 [~US $1852]), and there was a 5.5% reduction in the cost of dispensed antibiotics in the intervention group compared with the control group (–0.4% to 11.4%), which was equivalent to a reduction of about £830 (~US $1295) a year for an average intervention practice. Chen et al [47] also reported that total expenditure on text messages for each health worker in the intervention group was less than 2 ¥ (US $0.32) and for the control group with traditional education, it cost 560 ¥ (US $89.96) per health worker, for printed materials, accommodation, and transportation costs. This amounts to a 280-fold difference per person. Little et al [43] reported that online-based communication skills training was more cost-effective than traditional education (10% cost reduction) if the cost of antibiotic resistance was accounted for (€83.21 [~US $110.67] vs €92.46 [US $122.97]).

None of the included studies reported health care professionals' post-intervention skills, satisfaction, and adverse or untoward effects of digital education.

The overall risk of bias was high or unclear for most of the studies. We judged high or unclear risk of bias for six studies where risk of bias was mainly unclear in selection, detection, recruitment, reporting, and comparability with individual RCT (for cRCTs) domains due to unclear/lack of information (in four studies), and high risk of allocation concealment and other bias due to baseline differences (in two studies). In two studies, we judged the risk of bias as low, as the studies provided detailed information on random sequence generation, allocation concealment, blinding of outcome assessment, attrition rate, reporting of outcomes of interest, and baseline differences (Figure 4).

Figure 4. Risk of bias summary. Green: high risk of bias; red: low risk of bias; yellow: unclear risk of bias.



	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Recruitment Bias	Loss of cluters	Incorrect analysis	Comparability with individual randomised trials	Baseline Imbalances
Chen 2014	+	+	+	+	+	+	+	+	?	+
Dekker 2018	+	?	?	+	+	?	+	+	?	−
Legare 2012	+	−	?	+	+	+	+	+	?	+
Little 2013	+	+	?	+	+	+	+	+	?	+
McNulty 2018	+	?	?	+	+	?	+	+	?	+
Meeker 2016	+	+	?	+	+	+	+	+	?	+

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Bochicchio 2006	?	?	+	+	+	?	+
Butler 2012	+	+	+	?	+	+	+

Discussion

Principal Findings

Our review evaluated the effectiveness of digital education for health professions on antibiotic management compared with traditional education. Our results suggest that the use of online and mobile digital education in the management of antibiotics for postregistration health care professionals is associated with increased professional knowledge compared with traditional education. Health care professionals' clinical practice improvement (reduction in antibiotics prescribing or dispensing) was evaluated in six of eight included studies, which showed that online digital education may be more effective than traditional education. We found that certain types of mobile digital education modes were associated with significantly lower costs compared to traditional education. Overall, the risk of bias was high or unclear due to a lack of information mainly in relation to selection, detection, recruitment, reporting, and comparability with individual RCT (for cRCTs) domains.

Literature has shown that primary care settings are the most significant contributor to avenues for inappropriate antibiotic prescribing. There is therefore an emphasis on interventions promoting prudent prescribing in primary care, as detailed globally in national action plans on AMR [50]. Correspondingly, studies included in our review had a primary care focus—to compare the effectiveness of different educational approaches such as nondigital interventions (ie, delayed prescriptions or face-to-face interactive sessions) with digital education (ie, delivered via mobile digital education or online digital education) to improve antibiotic management in these settings [39-49].

Past reviews [29,30] highlighted that antibiotic management can be improved by educational interventions. However, these reviews focused on the effectiveness of nondigital education (ie, traditional education such as the use of text-based or paper-based interventions and quality improvement strategies). To the best of our knowledge, no reviews evaluated the effectiveness of digital education on the management of antibiotics for health care professionals. Our review provides the most comprehensive and up-to-date evaluation of digital education on antibiotic prescription in health care professionals. It shows the potential effectiveness for the use of certain types of digital technologies.

Strengths and Limitations

Our review has a number of strengths. We used robust methodology to synthesize the evidence by following gold-standard Cochrane guidelines while conducting this systematic review. We used a comprehensive search strategy and searched all major bibliographic databases for eligible studies.

However, our review has some limitations, which may affect the generalizability of the evaluated evidence. First, all included studies were conducted in primary care settings, and all participants were primary care physicians. Therefore, there is a need for studies focusing on other types of health care professionals such as dentists and allied health care professionals and on other settings such as pharmacies and other hospital practices. Antibiotic stewardship activity is increasingly focusing on the role of the community pharmacist in antibiotic prescription [51]. Second, all included studies focused on mobile digital education and online digital education. The effectiveness of other digital education modalities such as virtual reality, virtual patient scenarios, and offline digital education are underrepresented in the literature. Third, there is a lack of data reported in the included studies on health care professionals' postintervention skills, attitude, satisfaction, and economic and patient-related outcomes. None of the included studies evaluated untoward or adverse effects of the intervention for patients (eg, patient mortality, morbidity, and medical errors) and for learners (eg, dizziness, vertigo, and loss of attention). Due to limited data from the included studies and heterogeneity across the studies, we could not perform any sensitivity analysis or subgroup analyses that we had initially planned. Moreover, we were unable to evaluate some of the risk of bias domains due to poor description of the studies and lack of information within the study, which may affect the comprehensiveness of the risk of bias assessment. We analyzed postintervention data for two studies that reported knowledge outcome, as we could not calculate mean change score for the comparisons. Fourth, most of the included studies (7/8) were from high-income countries and no or limited data were from low- and middle-income countries, indicating that the depth of evidence related to any given outcome is mostly limited to high-income countries. Fifth, our review focuses only on the effectiveness of the digital education for antibiotic management training and does not address other important aspects such as potential barriers and facilitators or attrition rates. Lastly, only three studies assessed the cost of intervention and none reported on health care professionals' postintervention skills or adverse/untoward effects of digital education, thereby limiting the overall completeness and applicability of evidence.

Future Research

While the evidence included in our review is limited, it shows the potential effectiveness and applicability of certain digital education modalities such as mobile learning and online digital education in postregistration health professions education for antibiotic management. Future research should explore the effectiveness of other technologies such as virtual reality, virtual patient simulations, or serious gaming for training of health care professionals in diverse settings and measure outcomes such as change in skills and satisfaction.

Conclusions

Our findings suggest the potential effectiveness of mobile learning and online digital education for health professions education in antibiotic management. There is a lack of studies evaluating the effectiveness in low- and middle-income countries and of other forms of digital education such as virtual reality, virtual patient education, and offline digital education on antibiotic management. Future research should focus on assessing the effectiveness of different digital education modalities on health care professionals' antibiotic management across different settings. This should also be extended to the role of digital education in the management of antimicrobials, in general. There is also a need to integrate the available evidence into health professions education programs or clinical practice, especially on the use of online and mobile digital education modalities for postregistration health professions education on antibiotic management.

Acknowledgments

We would like to thank Mr Gerard Dunleavy for the help with article screening, data extraction, and suggestions on the review. We gratefully acknowledge funding support from Lee Kong Chian School of Medicine Start-Up Grant, Nanyang Technological University Singapore, Singapore. LC and BK have full access to the data and take responsibility of the data. LC had final responsibility for the decision to submit for publication.

This study obtained funding from the Lee Kong Chian School of Medicine Start-Up Grant, Nanyang Technological University Singapore, Singapore. The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all data in the study and had final responsibility of the decision to submit for publication.

Authors' Contributions

LC and BK conceived the idea for the review. BK, GD, and LG screened the studies and analyzed the data. BK and LC wrote the review. LC provided methodological guidance and provided critical inputs on the review. LG, MA, CC, OA, JC, and JoC provided insightful comments on and edited the review.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of the digital education modalities.
[PDF File (Adobe PDF File)31 KB - jmir_v21i9e14984_app1.pdf]

Multimedia Appendix 2
MEDLINE (OVID) search strategy.

[PDF File (Adobe PDF File)31 KB - jmir_v21i9e14984_app2.pdf]

Multimedia Appendix 3
Lists of clinical trial registries.
[PDF File (Adobe PDF File)25 KB - jmir_v21i9e14984_app3.pdf]

Multimedia Appendix 4
Results of included studies.
[PDF File (Adobe PDF File)53 KB - jmir_v21i9e14984_app4.pdf]

References

1. Toner E, Adalja A, Gronvall GK, Cicero A, Inglesby TV. Antimicrobial resistance is a global health emergency. Health Secur 2015;13(3):153-155 [FREE Full text] [doi: 10.1089/hs.2014.0088] [Medline: 26042858]
2. World Health Organization. 2018 Feb 05. Antibiotic resistance (fact sheet) URL: http://www.who.int/mediacentre/factsheets/antibiotic-resistance/en/ [accessed 2019-08-20]
3. Centres for Disease Control and Prevention. 2013. Biggest Threats and Data URL: https://www.cdc.gov/drugresistance/biggest_threats.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fdrugresistance%2Fthreat-report-2013%2Findex.html [accessed 2019-08-20]
4. de Kraker ME, Stewardson AJ, Harbarth S. Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050? PLoS Med 2016 Nov;13(11):e1002184 [FREE Full text] [doi: 10.1371/journal.pmed.1002184] [Medline: 27898664]
5. WHO. World Health Organization. 2018 Feb 20. Antimicrobial resistance URL: http://www.who.int/features/factfiles/antimicrobial_resistance/en/ [accessed 2019-08-20]
6. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. BMJ 2010 May 18;340:c2096. [doi: 10.1136/bmj.c2096] [Medline: 20483949]
7. Hansen MP, Hoffmann TC, McCullough AR, van Driel ML, Del Mar CB. Antibiotic Resistance: What are the Opportunities for Primary Care in Alleviating the Crisis? Front Public Health 2015;3:35 [FREE Full text] [doi: 10.3389/fpubh.2015.00035] [Medline: 25759809]
8. WHO. World Health Organization. 2017. Working for Health and Growth: Investing in the Health Workforce URL: https://www.who.int/hrh/com-heeg/reports/en/ [accessed 2019-08-20]
9. World Health Organization. 2018. WHO competency framework for health workers' education and training on antimicrobial resistance URL: http://apps.who.int/medicinedocs/documents/s23443en/s23443en.pdf [accessed 2019-08-30]
10. Krasnova TN, Samokhodskaya LM, Ivanitsky LV, Korogodina AD, Borisov EN, Nikiforova NV, et al. [Impact of interleukin-10 and interleukin-28 gene polymorphisms on the development and course of lupus nephritis]. Ter Arkh 2015;87(6):40-44. [doi: 10.17116/terarkh201587640-44] [Medline: 26281194]
11. Rogers Van Katwyk S, Jones SL, Hoffman SJ. Mapping educational opportunities for healthcare workers on antimicrobial resistance and stewardship around the world. Hum Resour Health 2018 Feb 05;16(1):9 [FREE Full text] [doi: 10.1186/s12960-018-0270-3] [Medline: 29402327]
12. Nathwani D, Guise T, Gilchrist M, BSAC. e-learning for global antimicrobial stewardship. Lancet Infect Dis 2017 Jun;17(6):579. [doi: 10.1016/S1473-3099(17)30275-X] [Medline: 28555577]
13. Fair RJ, Tor Y. Antibiotics and bacterial resistance in the 21st century. Perspect Medicin Chem 2014;6:25-64 [FREE Full text] [doi: 10.4137/PMC.S14459] [Medline: 25232278]
14. McNulty CA, Cookson BD, Lewis MA. Education of healthcare professionals and the public. J Antimicrob Chemother 2012 Jul;67 Suppl 1:i11-i18. [doi: 10.1093/jac/dks199] [Medline: 22855873]
15. Pulcini C, Gyssens IC. How to educate prescribers in antimicrobial stewardship practices. Virulence 2013 Feb 15;4(2):192-202 [FREE Full text] [doi: 10.4161/viru.23706] [Medline: 23361336]
16. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, Digital Health Education Collaboration. Digital Education in Health Professions: The Need for Overarching Evidence Synthesis. J Med Internet Res 2019 Feb 14;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
17. George PP, Zhabenko O, Kyaw BM, Antoniou P, Posadzki P, Saxena N, et al. Online Digital Education for Postregistration Training of Medical Doctors: Systematic Review by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 25;21(2):e13269 [FREE Full text] [doi: 10.2196/13269] [Medline: 30801252]
18. Kyaw BM, Posadzki P, Dunleavy G, Semwal M, Divakar U, Hervatis V, et al. Offline Digital Education for Medical Students: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Mar 25;21(3):e13165 [FREE Full text] [doi: 10.2196/13165] [Medline: 30907731]
19. Posadzki P, Bala MM, Kyaw BM, Semwal M, Divakar U, Koperny M, et al. Offline Digital Education for Postregistration Health Professions: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Apr 24;21(4):e12968 [FREE Full text] [doi: 10.2196/12968] [Medline: 31017584]

20. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious Gaming and Gamification Education in Health Professions: Systematic Review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]
21. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual Reality for Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
22. Kononowicz AA, Woodham L, Georg C, Edelbring S, Stathakarou N, Davies D, et al. PROSPERO CRD42016045912. 2016 Aug 15. Virtual patient simulations for health professional education [Cochrane protocol] URL: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=45912 [accessed 2019-08-20]
23. Dunleavy G, Nikolaou CK, Nifakos S, Atun R, Law GCY, Tudor Car L. Mobile Digital Education for Health Professions: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 12;21(2):e12937 [FREE Full text] [doi: 10.2196/12937] [Medline: 30747711]
24. Erhel S, Jamet E. Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. Comput Educ 2013 Sep:156-167 [FREE Full text] [doi: 10.1016/j.compedu.2013.02.019]
25. Stuchlíková L, Kósa A. Massive open online courses - Challenges and solutions in engineering education. 2013 Presented at: IEEE 11th International Conference on Emerging eLearning Technologies and Applications (ICETA); Oct; October 24-25, 2013; Stara Lesna, Slovakia URL: https://ieeexplore.ieee.org/document/6674459/authors#authors [doi: 10.1109/iceta.2013.6674459]
26. Ellaway R, Candler C, Greene P, Smothers V. MedBiquitous Virtual Patients. 2006 Sep 11. An Architectural Model for MedBiquitous Virtual Patients: MedBiquitous White Paper URL: https://tinyurl.com/y4yfrtts [accessed 2019-08-28]
27. Foley RP, Smilansky J. Teaching techniques: A handbook for health professionals. J Oral Maxillofac Surg 1982 Jun 06;40(6):394. [doi: 10.1016/0278-2391(82)90272-5]
28. Crompton H. A historical overview of mobile learning: Toward learner-centered education. In: Berge Z, Muilenburg L, editors. Handbook of mobile learning. Florence, KA: Routledge; 2013:e2013-e2014.
29. Ranji SR, Steinman MA, Shojania KG, Gonzales R. Interventions to reduce unnecessary antibiotic prescribing: a systematic review and quantitative analysis. Med Care 2008 Aug;46(8):847-862. [doi: 10.1097/MLR.0b013e318178eabd] [Medline: 18665065]
30. Roque F, Herdeiro MT, Soares S, Teixeira Rodrigues A, Breitenfeld L, Figueiras A. Educational interventions to improve prescription and dispensing of antibiotics: a systematic review. BMC Public Health 2014 Dec 15;14:1276 [FREE Full text] [doi: 10.1186/1471-2458-14-1276] [Medline: 25511932]
31. 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]
32. Kyaw BM, Dunleavy G, Tudor Car L. PROSPERO CRD42018109742. 2018 Oct 04. Digital education in health professions in antibiotic management: a systematic review protocol by the digital health education collaboration [Cochrane Protocol] URL: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=109742 [accessed 2019-08-20]
33. Clarivate Analytics. Philadelphia, PA: Thomson Reuters; 2015. EndNote Version X7.3 (for Windows) URL: https://support.clarivate.com/Endnote/s/article/EndNote-X7-Updates?language=en_US [accessed 2019-08-28]
34. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 07;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
35. Higgins JG, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. 2011. URL: http://handbook-5-1.cochrane.org/ [accessed 2019-08-28]
36. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Cochrane Collaboration; 2014.
37. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. BMJ 2003 Oct 04;327(7418):785-789 [FREE Full text] [doi: 10.1136/bmj.327.7418.785] [Medline: 14525877]
38. Knight GM, Costelloe C, Murray KA, Robotham JV, Atun R, Holmes AH. Addressing the Unknowns of Antimicrobial Resistance: Quantifying and Mapping the Drivers of Burden. Clin Infect Dis 2018 Feb 01;66(4):612-616 [FREE Full text] [doi: 10.1093/cid/cix765] [Medline: 29020246]
39. Yardley L, Douglas E, Anthierens S, Tonkin-Crine S, O'Reilly G, Stuart B, GRACE consortium. Evaluation of a web-based intervention to reduce antibiotic prescribing for LRTI in six European countries: quantitative process analysis of the GRACE/INTRO randomised controlled trial. Implement Sci 2013 Nov 15;8:134 [FREE Full text] [doi: 10.1186/1748-5908-8-134] [Medline: 24238118]
40. Oppong R, Smith RD, Little P, Verheij T, Butler CC, Goossens H, et al. Cost-effectiveness of internet-based training for primary care clinicians on antibiotic prescribing for acute respiratory tract infections in Europe. J Antimicrob Chemother 2018 Nov 01;73(11):3189-3198. [doi: 10.1093/jac/dky309] [Medline: 30165522]

41. Meeker D, Linder JA, Fox CR, Friedberg MW, Persell SD, Goldstein NJ, et al. Effect of Behavioral Interventions on Inappropriate Antibiotic Prescribing Among Primary Care Practices: A Randomized Clinical Trial. JAMA 2016 Feb 09;315(6):562-570 [FREE Full text] [doi: 10.1001/jama.2016.0275] [Medline: 26864410]
42. McNulty C, Hawking M, Lecky D, Jones L, Owens R, Charlett A, et al. Effects of primary care antimicrobial stewardship outreach on antibiotic use by general practice staff: pragmatic randomized controlled trial of the TARGET antibiotics workshop. J Antimicrob Chemother 2018 May 01;73(5):1423-1432 [FREE Full text] [doi: 10.1093/jac/dky004] [Medline: 29514268]
43. Little P, Stuart B, Francis N, Douglas E, Tonkin-Crine S, Anthierens S, GRACE consortium. Effects of internet-based training on antibiotic prescribing rates for acute respiratory-tract infections: a multinational, cluster, randomised, factorial, controlled trial. Lancet 2013 Oct 05;382(9899):1175-1182 [FREE Full text] [doi: 10.1016/S0140-6736(13)60994-0] [Medline: 23915885]
44. Légaré F, Labrecque M, Cauchon M, Castel J, Turcotte S, Grimshaw J. Training family physicians in shared decision-making to reduce the overuse of antibiotics in acute respiratory infections: a cluster randomized trial. CMAJ 2012 Sep 18;184(13):E726-E734 [FREE Full text] [doi: 10.1503/cmaj.120568] [Medline: 22847969]
45. Dekker ARJ, Verheij TJM, Broekhuizen BDL, Butler CC, Cals JWL, Francis NA, et al. Effectiveness of general practitioner online training and an information booklet for parents on antibiotic prescribing for children with respiratory tract infection in primary care: a cluster randomized controlled trial. J Antimicrob Chemother 2018 May 01;73(5):1416-1422. [doi: 10.1093/jac/dkx542] [Medline: 29438547]
46. Couët N, Labrecque M, Robitaille H, Turcotte S, Légaré F. The impact of DECISION+2 on patient intention to engage in shared decision making: secondary analysis of a multicentre clustered randomized trial. Health Expect 2015 Dec;18(6):2629-2637 [FREE Full text] [doi: 10.1111/hex.12235] [Medline: 25041071]
47. Chen Y, Yang K, Jing T, Tian J, Shen X, Xie C, et al. Use of text messages to communicate clinical recommendations to health workers in rural China: a cluster-randomized trial. Bull World Health Organ 2014 Jul 01;92(7):474-481 [FREE Full text] [doi: 10.2471/BLT.13.127076] [Medline: 25110372]
48. Butler CC, Simpson SA, Dunstan F, Rollnick S, Cohen D, Gillespie D, et al. Effectiveness of multifaceted educational programme to reduce antibiotic dispensing in primary care: practice based randomised controlled trial. BMJ 2012 Feb 02;344:d8173 [FREE Full text] [doi: 10.1136/bmj.d8173] [Medline: 22302780]
49. Bochicchio GV, Smit PA, Moore R, Bochicchio K, Auwaerter P, Johnson SB, POC-IT Group. Pilot study of a web-based antibiotic decision management guide. J Am Coll Surg 2006 Mar;202(3):459-467. [doi: 10.1016/j.jamcollsurg.2005.11.010] [Medline: 16500251]
50. World Health Organization. 2015 May. Antimicrobial resistance: National action plans URL: https://www.who.int/antimicrobial-resistance/national-action-plans/en/ [accessed 2019-08-20]
51. Jamshed S, Padzil F, Shamsudin SH, Bux SH, Jamaluddin AA, Bhagavathula AS, et al. Antibiotic Stewardship in Community Pharmacies: A Scoping Review. Pharmacy (Basel) 2018 Aug 23;6(3):E92 [FREE Full text] [doi: 10.3390/pharmacy6030092] [Medline: 30142902]

Abbreviations

AMR: antimicrobial resistance
CME: continuing medical education
cRCT: cluster randomized controlled trial
CRP: C-reactive protein
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
RR: risk ratio
SMD: standardized mean difference
WHO: World Health Organization

Edited by A Marusic, G Eysenbach; submitted 10.06.19; peer-reviewed by H Mufti, M Behzadifar, M Vidak; comments to author 04.07.19; revised version received 07.08.19; accepted 17.08.19; published 12.09.19.

Please cite as:
Kyaw BM, Tudor Car L, van Galen LS, van Agtmael MA, Costelloe CE, Ajuebor O, Campbell J, Car J
Health Professions Digital Education on Antibiotic Management: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration
J Med Internet Res 2019;21(9):e14984
URL: http://www.jmir.org/2019/9/e14984/
doi:10.2196/14984
PMID:31516125

Review

Digital Health Professions Education in the Field of Pediatrics: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration

Serena Brusamento[1], MPH, MD; Bhone Myint Kyaw[2], MSc, MBBS, PhD; Penny Whiting[3], BA, MSc, PhD; Li Li[4], MSc, MPH, MD; Lorainne Tudor Car[1,5], MSc, MD, PhD

[1]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom
[2]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[3]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom
[4]Office of Medical Education, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[5]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

Corresponding Author:
Lorainne Tudor Car, MSc, MD, PhD
Family Medicine and Primary Care
Lee Kong Chian School of Medicine
Nanyang Technological University
11 Mandalay Road
Singapore
Singapore
Phone: 65 69041258
Email: lorainne.tudor.car@ntu.edu.sg

Abstract

Background: Reducing childhood morbidity and mortality is challenging, particularly in countries with a shortage of qualified health care workers. Lack of trainers makes it difficult to provide the necessary continuing education in pediatrics for postregistration health professionals. Digital education, teaching and learning by means of digital technologies, has the potential to deliver medical education to a large audience while limiting the number of trainers needed.

Objective: The goal of the research was to evaluate whether digital education can replace traditional learning to improve postregistration health professionals' knowledge, skills, attitudes, and satisfaction and foster behavior change in the field of pediatrics.

Methods: We completed a systematic review of the literature by following the Cochrane methodology. We searched 7 major electronic databases for articles published from January 1990 to August 2017. No language restrictions were applied. We independently selected studies, extracted data, and assessed risk of bias, and pairs of authors compared information. We contacted authors of studies for additional information if necessary. All pooled analyses were based on random effects models. We included individually or cluster randomized controlled trials that compared digital education with traditional learning, no intervention, or other forms of digital education. We assessed the quality of evidence using the Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) criteria.

Results: Twenty studies (1382 participants) were included. Participants included pediatricians, physicians, nurses, and midwives. Digital education technologies were assessed including high-fidelity mannequins (6 studies), computer-based education (12 studies), mobile learning (1 study), and virtual reality (1 study). Most studies reported that digital education was either as effective as or more effective than the control intervention for outcomes including skill, knowledge, attitude, and satisfaction. High-fidelity mannequins were associated with higher postintervention skill scores compared with low-fidelity mannequins (standardized mean difference 0.62; 95% CI 0.17-1.06; moderate effect size, low-quality evidence). One study reported physician change in practicing behavior and found similar effects between offline plus online digital education and no intervention. The only study that assessed impact on patient outcome found no difference between intervention and control groups. None of the included studies reported adverse or untoward effects or economic outcomes of the digital education interventions. The risk of bias was mainly unclear or high. The quality of evidence was low due to study inconsistencies, limitations, or imprecision across the studies.

Conclusions: Digital education for postregistration health professions education in pediatrics is at least as effective as traditional learning and more effective than no learning. High-fidelity mannequins were found to be more effective at improving skills than traditional learning with low-fidelity mannequins. Computer-based offline/online digital education was better than no intervention for knowledge and skill outcomes and as good as traditional face-to-face learning. This review highlights evidence gaps calling for more methodologically rigorous randomized controlled trials on the topic.

Trial Registration: PROSPERO CRD42017057793; https://tinyurl.com/y5q9q5o6

(J Med Internet Res 2019;21(9):e14231) doi:10.2196/14231

KEYWORDS

digital education; randomized controlled trials; pediatrics; systematic review; meta-analysis; traditional learning; high-fidelity mannequins

Introduction

Reducing childhood morbidity and mortality is a global health priority. Mortality remains high in many low- and middle-income countries (LMIC), despite improvements achieved as a result of the Millennium Development Goals [1,2]. Reducing childhood mortality and ensuring global access to health care through health workforce development is one of the 17 United Nations Sustainable Development Goals [3-5]. A major factor in reducing childhood morbidity and mortality is the quality of pediatric health care. This is influenced by the skills of the health professionals—physicians, nonphysician clinicians, nurses, and midwives. To provide optimal care, health professionals need continuous, high-quality, and up-to-date education [6]. Lack of access to learning resources, coupled with remote locations, limited health professionals, and a need for the ongoing provision of health services represent significant barriers to health professions education in many settings.

Pediatric health professions education is particularly important due to the unique nature of the diseases, need for timely and appropriate treatment, and the narrow margin of treatment error compared with the adult patients [7]. Additionally, there is a growing demand to educate pediatrics health professionals in certain topics that need timely incorporation/implementation of evidence-based recommendations and guidelines such as delivering updated guidelines on immunizations [8], chemotherapy [9], respiratory infections [10], and neonatal management [11]. Traditional forms of education such as face-to-face or didactic lectures or workshops might not be adequate to address these demands in a timely manner, and digital education can be an alternative option for educating pediatric health professionals as it provides an efficient, timely, and convenient mode for the learners which further helps to improve outcomes [12].

Digital education helps overcome resource, geographical, and time barriers. Computer-based and mobile learning allow learners to access educational materials without time or place restrictions, allowing them to work at their own pace and time from any location [13]. A further advantage of this type of learning is that it generally requires fewer tutors. Unlike face-to-face lectures or tutorials, the number of learners who can participate in this type of education is far greater. Computer-based offline digital education provides increased access to learning materials with limited internet connectivity [14,15]. Mobile learning or mLearning supports learning in a similar way by enhancing the delivery of learning materials without time and place limitation through a handheld mobile device. There are a number of mobile device–based functions such as short message service (SMS or texts), multimedia message service, podcasts, and mobile apps that support the delivery of educational materials based on the needs of learners and learning processes [16,17].

Simulation-based medical education such as training in virtual reality environments and virtual patient scenarios supports creation of 3D virtual world or patient case-based scenarios that are similar to real-life clinical scenarios, designed specifically for health professional training [18-20]. Similarly, training via psychomotor skills trainers such as high- or middle-fidelity mannequins allows for training of different types of psychomotor or technical skills acquisitions such as resuscitation and suturing skills [21-23].

The potential benefits of digital education for health professions education have been evaluated in previous reviews and acknowledged by the World Health Organization [12-14,24-27]. While there are reviews on the use of digital education in specific pediatric fields (eg, emergency or rehabilitation care) and for undergraduate education [28-30], we are unaware of any systematic review assessing the effectiveness of digital education in the field of pediatrics for postregistration health professionals. This review evaluates the effectiveness of different modalities of digital education for postregistration pediatrics health professionals in comparison with traditional learning or other forms of digital education. We assessed the impact of digital education on participants' knowledge, skills, attitudes, clinical practice, and satisfaction compared with other forms of learning.

Methods

Systematic Review

The protocol for the systematic review was registered with PROSPERO [CRD42017057793] [31]. For a detailed description of the methodology, please refer to the methods as described by the Digital Health Education Collaboration [32], a global initiative focused on evaluating the effectiveness of digital health professions education through a series of methodologically robust systematic reviews.

For the purpose of this review, digital education can be defined as "an approach to teaching and learning, representing all or part of the educational model applied, that is based on the use of electronic media and devices as tools for improving access to training, communication and interaction, that facilitates the adaptation of new ways of understanding and developing learning" [33]. Digital education encompasses a variety of learning modalities including computer-based online/offline digital education (online/offline digital education), high-fidelity mannequins, virtual reality environments, virtual patient scenarios, serious gaming and gamification, and mobile learning, etc [15,20,27,34-36]. It is often combined with traditional nondigital learning, known as blended learning. Traditional learning means learning via traditional forms of education such as paper- or textbook–based learning, didactic or face-to-face lectures, tutorials, box trainers, or low-fidelity mannequins.

Inclusion Criteria

We included studies involving learners who were enrolled in any postregistration health professional or continuing medical education (CME) programs. For this review, postregistration health professional programs can be defined as any type of study after a qualification which is recognized by the relevant governmental or professional bodies that enables the qualification holder entry into or continuation of work in the health care workforce in a more independent or senior role. We also included studies focusing on CME programs that involved the use of digital education to deliver the learning contents. We included all postregistration health professionals listed in the Health Field of Education and Training (091) of the International Standard Classification of Education except professionals from traditional, alternative, and complementary medicine [31].

We included individually or cluster randomized controlled trials (RCTs) that compared digital education interventions on any pediatric-related topic for postregistration health professions with traditional learning, no intervention, or other forms of digital education [37]. Eligible studies had to report at least one of the specified primary or secondary outcomes. Primary outcomes (measured using any validated or nonvalidated instruments) were (1) participants' knowledge scores, (2) participants' skills, and (3) participants' attitudes toward the interventions or toward new clinical knowledge. Secondary outcomes included participants' satisfaction with the intervention, participants' change in clinical practice, the economic impact of digital education (eg, cost and cost effectiveness), patient-related outcomes, and any adverse or unintended effects of digital education.

Search Strategy and Data Sources

We developed a comprehensive search strategy for 7 electronic databases: MEDLINE (Ovid), Embase (Elsevier), Cochrane Central Register of Controlled Trials (Wiley), PsycINFO (Ovid), Educational Research Information Center (Ovid), Cumulative Index to Nursing and Allied Health Literature (Ebsco), and Web of Science Core Collection (Thomson Reuters). The databases were searched for articles published from January 1990 to August 2017 without language or publication restrictions; studies published prior to this were not considered due to technological advances (Multimedia Appendix 1). We searched reference lists of all the studies that we deemed eligible for inclusion in our review and relevant systematic reviews. We also searched the International Clinical Trials Registry Platform Search Portal and metaRegister of Controlled Trials to identify unpublished trials.

Data Collection and Analysis

The search results from different electronic databases were combined in a single Endnote library, and duplicate records were removed. Three review authors (SB, BMK, and LL) independently screened titles and abstracts of all the records to identify potentially eligible studies. We retrieved full-text copies of the articles deemed potentially relevant. Finally, three reviewers (SB, BMK, and LL) independently assessed the full-text versions of the retrieved articles against the eligibility criteria. Any disagreements were resolved through discussion between the reviewers.

Data Extraction and Management

Each manuscript was independently extracted by two reviewers from a team of three (SB, BMK and LL). We extracted relevant characteristics related to types of participants (ie, doctors, nurses, or midwives), interventions used, comparators or control groups, outcome measures including details of assessment methods, and results from all included studies using a standard data collection form built on an Excel (Microsoft Corp) template (Multimedia Appendix 2). Any disagreements between the reviewers were resolved by discussion. We contacted study authors for any missing information, particularly information required to judge the risk of bias.

Measures of Treatment Effect

For continuous outcomes, we extracted mean postintervention scores, standard deviations (SD), and number of participants for each intervention group and control groups. We used these data to calculate standardized mean differences (SMD) with associated 95% confidence intervals using random effect models. For studies that reported only median and ranges, we converted these to mean and SD [38]. Dichotomous data were extracted as number of events and number of participants in each intervention group. These data were used to calculated odds ratios (ORs) with associated 95% confidence intervals using random effect models. We were unable to identify a clinically meaningful interpretation of effect size in the literature for digital education interventions. Therefore, in line with other studies in the field, we presented outcomes using postintervention SMD and interpreted the effect size using the Cohen rule [39,40].

Assessment of Risk of Bias in Included Studies

Three reviewers (SB, BMK, and LL) independently assessed the methodological risk of bias of included studies in line with the Cochrane methodology [40]. This includes domains covering random sequence generation, allocation concealment, blinding (outcome assessment), completeness of outcome data (attrition bias), selective outcome reporting (relevant outcomes reported) and other sources of bias such as baseline imbalances, inappropriate administration of an intervention, and contamination.

For cluster RCTs (cRCTs), we also assessed the risk of the following additional domains: recruitment bias, baseline imbalance, loss of clusters, incorrect analysis, and comparability with individual RCTs as recommended by Puffer et al [41]. Judgments concerning the risk of bias for each study were scored as high, low, or unclear. Disagreements were resolved by consensus between the two authors or through discussion with a third author (Multimedia Appendix 3) [41].

Data Synthesis

We grouped studies by type of digital education, comparator group, and outcome. Where sufficient data were available, we used random effect meta-analysis to estimate summary effect estimates. Heterogeneity was assessed visually using forest plots and statistically using the I^2 statistic [42]. Where meta-analysis was not possible (except for skill outcome), a narrative synthesis was presented. We aimed to carry out prespecified subgroup analysis including the analysis based on country income such as LMIC versus HIC (high-income countries). However, due to the limited primary data, we were unable to conduct prespecified subgroup analyses.

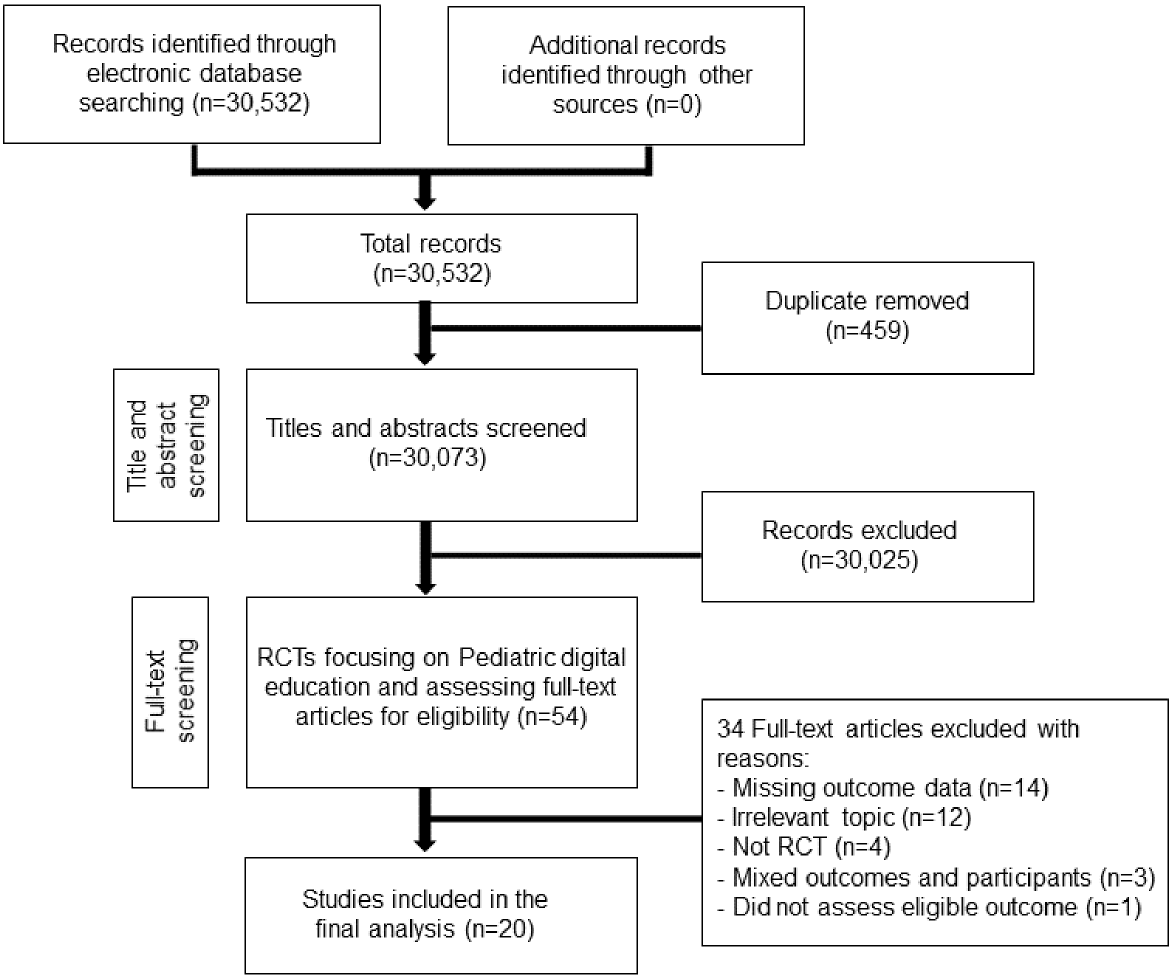
We also prepared summary of findings tables for the major comparisons focusing on high-fidelity mannequins and computer-based education and assessed the overall quality of evidence by using GRADEprofiler (GRADEproGDT Web-based version, McMaster University) [43]. We presented the findings of the review based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses statements, and details of the statements are presented in Multimedia Appendix 4. Descriptions of the terminologies used in this review are presented in Multimedia Appendix 5.

Results

Search Results

We used a common search strategy for a series of systematic reviews evaluating the effectiveness of digital education on different areas of health professions education. The overall searches identified 30,532 references. After reading titles and abstracts, 3466 references with different digital education interventions focusing on different areas of health professions education were identified, of which 54 studies focusing on pediatric education were selected for full-text review, and 20 trials fulfilled our inclusion criteria (1382 participants): 18 RCTs and two cRCTs (Figure 1).

Figure 1. Flow of studies through the review process.



Three studies were performed in the LMICs India [44], Ethiopia [45], and Lebanon [23]. The remaining were performed in HICs: North America (13 studies), Europe (3 studies), and Asia (1 study). Digital education was used for education in a variety of pediatric fields (Multimedia Appendix 6): neonatal/pediatric resuscitation and intubation (8 studies) [21-23,44-48], childhood obesity (2 studies) [49,50], pediatric emergency (1 study) [51], firearm injury prevention (1 study) [52], detection of childhood abuse (1 study) [53], drug prescriptions (1 study) [54], pediatric sedation (1 study) [55], child emergency nursing care (1 study) [56], counseling for parent's smoking cessation (1 study) [57], pediatric orthopedic surgery (1 study) [58], and asthma (1 study) [59].

Participants in included studies were pediatricians (9 studies) [23,46,48,50-52,55,59,60]; postregistration nurses (4 studies) [22,44,53,56]; midwives and health extension workers (1 study) [45]; childcare health consultants (1 study) [49]; orthopedic surgery residents (1 study) [58]; family medicine residents (1 study) [47]; junior doctors (1 study) [54]; mixed participants including doctors, nurses, emergency technicians, and paramedics (1 study) [21]; and pediatric respiratory therapists and nurses (1 study) [57].

Seventeen studies were 2-arm studies: 9 compared digital education with traditional learning [21-23,44,45,47,48,56,58], 7 compared digital education with no intervention [50-54,57,59], and one compared 2 digital education methods [55]. Three studies included 3 intervention arms: 2 compared digital education with traditional learning and blended learning [46,60], and one compared digital education with traditional learning and no intervention [49]. Digital education technologies evaluated included high-fidelity mannequins (6 studies), computer-based education (12 studies), mLearning (1 study), and a virtual reality environment (1 study).

Risk of Bias Assessment and Quality of Evidence

The main limitation with the included studies was incomplete outcome data—5 studies were judged as high risk of bias for this domain [45,50,55,59,60], and for 3 studies, there was insufficient information on missing data to make a judgement [22,48,51]. Many studies were poorly reported, making it difficult to judge risk of bias. Randomization, concealment of allocation, and blinding of outcome assessors were poorly reported, with 8, 13, and 12 studies, respectively, judged at unclear risk of bias for these domains. However, only single studies were judged as high risk of bias for randomization [22], blinding of outcome assessment domains [23], selective reporting [57], and other bias [44]. Additional domains assessed for the 2-cluster trials were all at low risk of bias (Figure 2). The quality of evidence was low due to study inconsistencies, limitations, and/or imprecision across the studies (Multimedia Appendix 7).

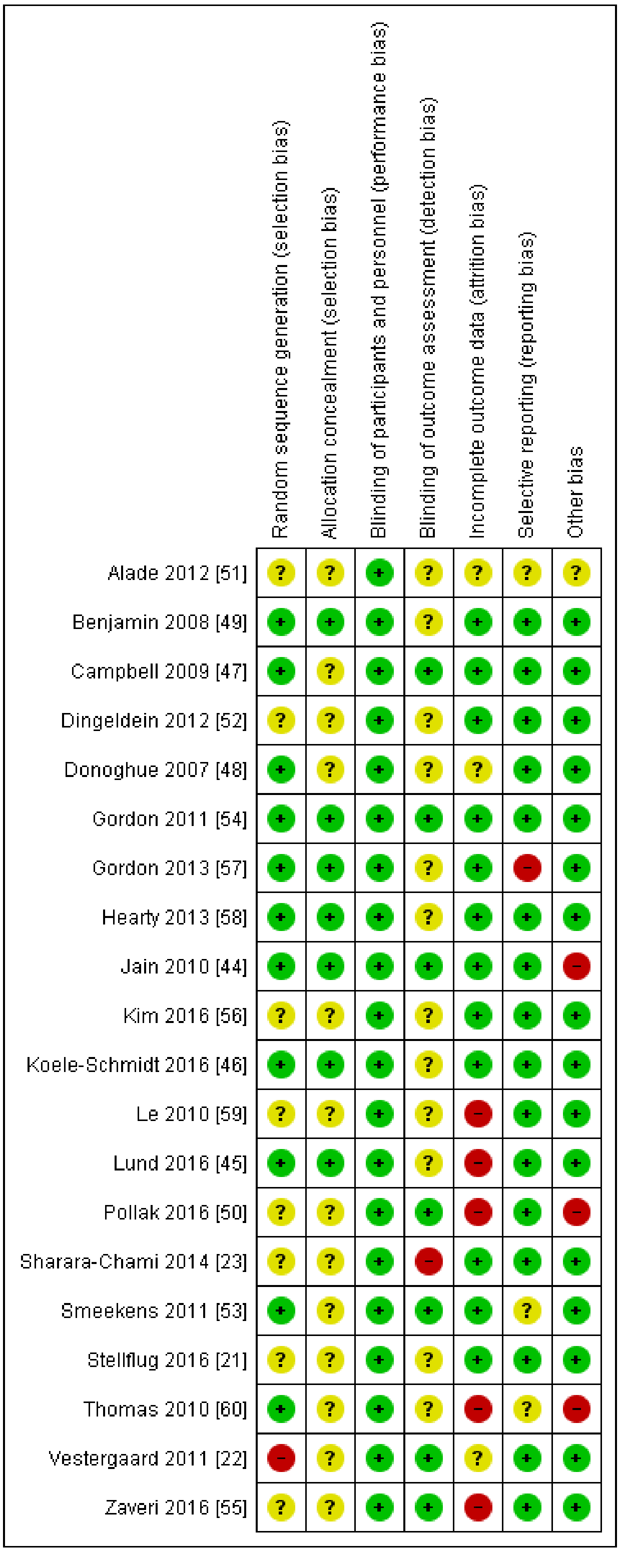
The detailed results of included studies are presented in Multimedia Appendix 8. Where available, standardized mean differences are presented in Figures 3-5.

High-Fidelity Mannequins

Six studies (320 participants) assessed the use of high-fidelity mannequins to provide training in neonatal or pediatric resuscitation, 5 compared with low-fidelity mannequins [21,23,47,48,60] and one compared with traditional learning with checklist procedure training [56]. One study included an additional intervention group consisting of blended learning where participants receiving training used high-fidelity mannequins combined with team training [60].

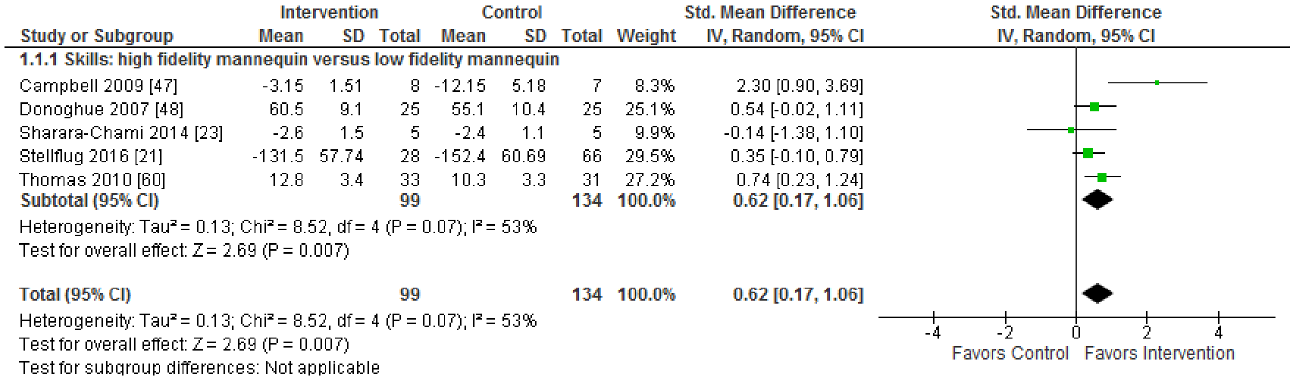
All studies assessed participants' psychomotor skill scores, although different measures of skill were used including time to complete different steps of the intubation and resuscitation, number of redirections provided during the procedure, and performance checklists for different tasks. Overall, high-fidelity mannequins were associated with greater postintervention skill scores compared with low-fidelity mannequins (SMD 0.62; 95% CI 0.17 to 1.06; I^2=53%, 5 studies; Figure 3). One study compared high-fidelity mannequins with traditional learning with checklist procedure training and reported higher postintervention skill scores in the intervention group (SMD 0.86; 95% CI 0.27 to 1.44; Figure 4) [56]. The study that included an additional blended learning group reported greater improvement in skill in the blended learning group compared with low-fidelity mannequins alone (SMD 1.34; 95% CI 0.82 to 1.87) [60].

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item across all included studies.



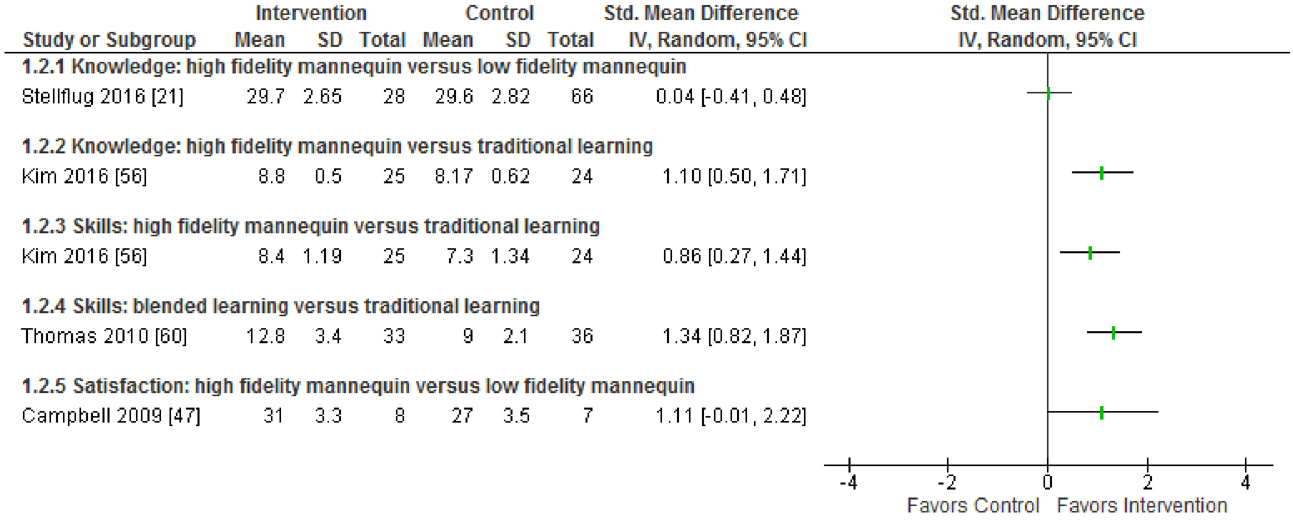
	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Alade 2012 [51]	?	?	+	?	?	?	?
Benjamin 2008 [49]	+	+	+	?	+	+	+
Campbell 2009 [47]	+	?	+	+	+	+	+
Dingeldein 2012 [52]	?	?	+	?	+	+	+
Donoghue 2007 [48]	+	?	+	?	?	+	+
Gordon 2011 [54]	+	+	+	+	+	+	+
Gordon 2013 [57]	+	+	+	?	+	−	+
Hearty 2013 [58]	+	+	+	?	+	+	+
Jain 2010 [44]	+	+	+	+	+	+	−
Kim 2016 [56]	?	?	+	?	+	+	+
Koele-Schmidt 2016 [46]	+	+	+	?	+	+	+
Le 2010 [59]	?	?	+	?	−	+	+
Lund 2016 [45]	+	+	+	?	−	+	+
Pollak 2016 [50]	?	?	+	+	−	+	−
Sharara-Chami 2014 [23]	?	?	+	−	+	+	+
Smeekens 2011 [53]	+	?	+	+	+	?	+
Stellflug 2016 [21]	?	?	+	?	+	+	+
Thomas 2010 [60]	+	?	+	?	−	?	−
Vestergaard 2011 [22]	−	?	+	+	?	+	+
Zaveri 2016 [55]	?	?	+	+	−	+	+

Figure 3. Difference in postintervention skill between interventions using high-fidelity mannequins and low-fidelity mannequins.



Study or Subgroup	Intervention Mean	Intervention SD	Intervention Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Skills: high fidelity mannequin versus low fidelity mannequin								
Campbell 2009 [47]	-3.15	1.51	8	-12.15	5.18	7	8.3%	2.30 [0.90, 3.69]
Donoghue 2007 [48]	60.5	9.1	25	55.1	10.4	25	25.1%	0.54 [-0.02, 1.11]
Sharara-Chami 2014 [23]	-2.6	1.5	5	-2.4	1.1	5	9.9%	-0.14 [-1.38, 1.10]
Stellflug 2016 [21]	-131.5	57.74	28	-152.4	60.69	66	29.5%	0.35 [-0.10, 0.79]
Thomas 2010 [60]	12.8	3.4	33	10.3	3.3	31	27.2%	0.74 [0.23, 1.24]
Subtotal (95% CI)			**99**			**134**	**100.0%**	**0.62 [0.17, 1.06]**
Heterogeneity: Tau² = 0.13; Chi² = 8.52, df = 4 (P = 0.07); I² = 53%								
Test for overall effect: Z = 2.69 (P = 0.007)								
Total (95% CI)			**99**			**134**	**100.0%**	**0.62 [0.17, 1.06]**
Heterogeneity: Tau² = 0.13; Chi² = 8.52, df = 4 (P = 0.07); I² = 53%								
Test for overall effect: Z = 2.69 (P = 0.007)								
Test for subgroup differences: Not applicable								

Figure 4. Difference in knowledge, skill, and satisfaction between interventions using high-fidelity mannequins and control groups.



Study or Subgroup	Intervention Mean	Intervention SD	Intervention Total	Control Mean	Control SD	Control Total	Std. Mean Difference IV, Random, 95% CI
1.2.1 Knowledge: high fidelity mannequin versus low fidelity mannequin							
Stellflug 2016 [21]	29.7	2.65	28	29.6	2.82	66	0.04 [-0.41, 0.48]
1.2.2 Knowledge: high fidelity mannequin versus traditional learning							
Kim 2016 [56]	8.8	0.5	25	8.17	0.62	24	1.10 [0.50, 1.71]
1.2.3 Skills: high fidelity mannequin versus traditional learning							
Kim 2016 [56]	8.4	1.19	25	7.3	1.34	24	0.86 [0.27, 1.44]
1.2.4 Skills: blended learning versus traditional learning							
Thomas 2010 [60]	12.8	3.4	33	9	2.1	36	1.34 [0.82, 1.87]
1.2.5 Satisfaction: high fidelity mannequin versus low fidelity mannequin							
Campbell 2009 [47]	31	3.3	8	27	3.5	7	1.11 [-0.01, 2.22]

Two studies assessed knowledge [21,56]. One found that knowledge gain was greater with high-fidelity mannequins than with traditional learning (SMD 1.10; 95% CI 0.50 to 1.71). The other found no difference between groups (SMD 0.04; 95% CI –0.41 to 0.48). One study assessed participant satisfaction and found weak evidence for greater satisfaction with high-fidelity mannequins compared with low-fidelity mannequins (SMD 1.11; 95% CI –0.01 to 2.22).

Computer-Based Education

Twelve studies (854 participants) assessed computer-based education, 7 compared with no intervention, 3 compared with traditional learning, one compared with traditional learning and no intervention, and one compared with blended learning and traditional learning. Nine studies evaluated computer-based online education including Web-based modules, tele-education, PowerPoint presentations, and online modules with audio-recorded patient cases. Two studies assessed offline computer-based education using PowerPoint presentations and computer-based video teaching. One assessed both online and offline digital education (Web- or CD-ROM–based learning plus 2 conference calls).

Seven studies (442 participants) assessed participant psychomotor skills postintervention [22,44,46,50,51,53,54]. The studies used different types of outcome measures such as time to complete different steps of the intubation and resuscitation, number of redirections provided during the procedure, and performance checklist for different tasks. Three studies compared computer-based education (online, tele-education, online video learning, and offline computer module with pictures and videos) with traditional learning [22,44,46]; one study included an additional blended learning group [46]. None of the comparisons suggested a difference in postintervention skill between learning groups (Figure 5). Two studies [53,54] compared computer-based learning (PowerPoint presentation, software, flash program, videos, animations, and online webpage) with no intervention, all reported greater improvement in skill following computer-based learning (Figure 5). Two studies did not report data in an appropriate form to include on the forest plot. One compared digital education (bimonthly, brief Web-based computer modules) with traditional learning to improve ultrasonography skills for pediatric emergencies and reported higher scores in the online learning group compared with traditional learning (*P*=.02) [51]. One study assessed primary care physicians' adolescent weight management skills and favored digital education (online learning containing patients' audio clips) over traditional learning (*P*=.001) [50].

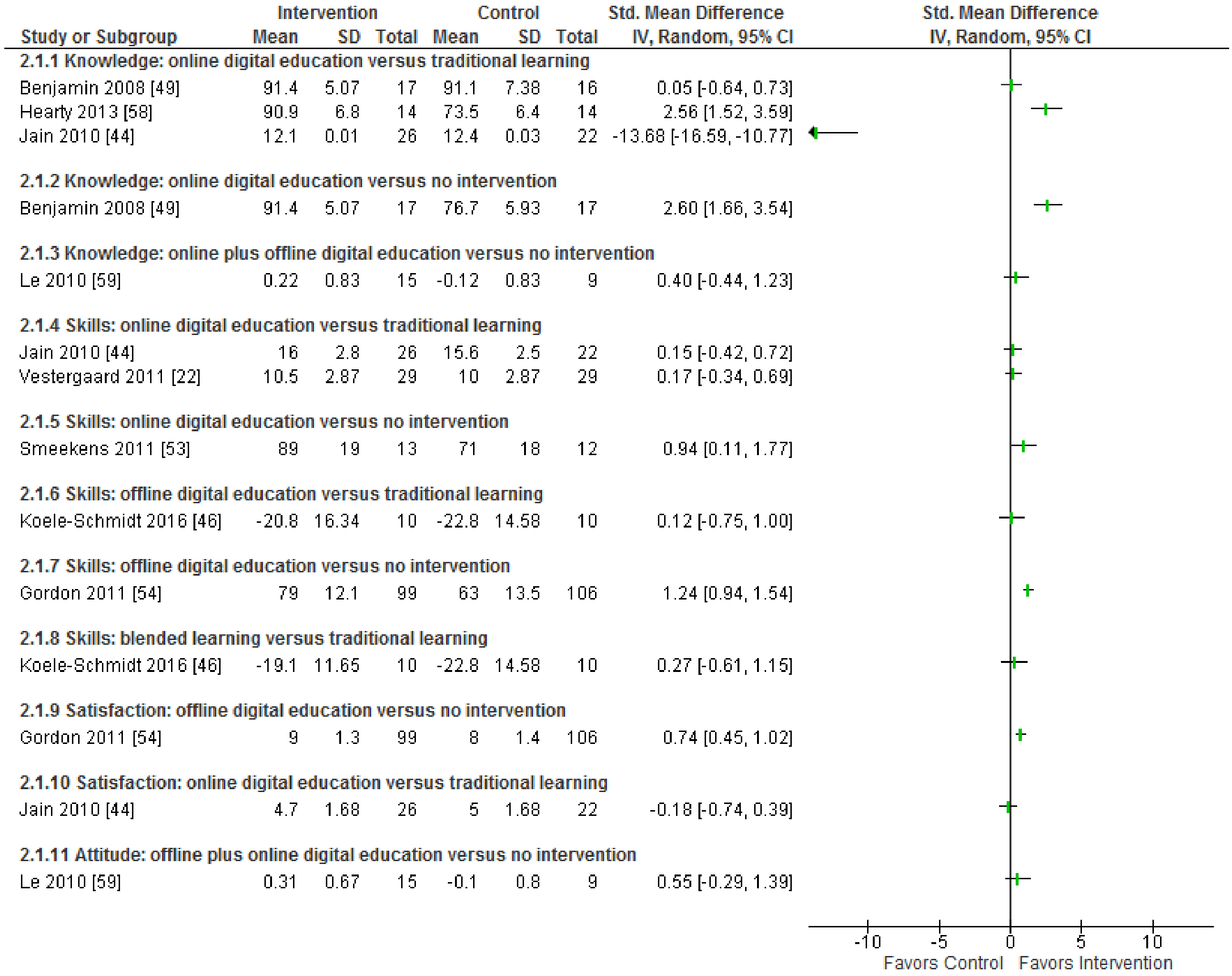


Figure 5. Difference in knowledge, skill, attitude, and satisfaction between interventions using computer-based education and control groups.

Study or Subgroup	Intervention Mean	Intervention SD	Intervention Total	Control Mean	Control SD	Control Total	Std. Mean Difference IV, Random, 95% CI
2.1.1 Knowledge: online digital education versus traditional learning							
Benjamin 2008 [49]	91.4	5.07	17	91.1	7.38	16	0.05 [-0.64, 0.73]
Hearty 2013 [58]	90.9	6.8	14	73.5	6.4	14	2.56 [1.52, 3.59]
Jain 2010 [44]	12.1	0.01	26	12.4	0.03	22	-13.68 [-16.59, -10.77]
2.1.2 Knowledge: online digital education versus no intervention							
Benjamin 2008 [49]	91.4	5.07	17	76.7	5.93	17	2.60 [1.66, 3.54]
2.1.3 Knowledge: online plus offline digital education versus no intervention							
Le 2010 [59]	0.22	0.83	15	-0.12	0.83	9	0.40 [-0.44, 1.23]
2.1.4 Skills: online digital education versus traditional learning							
Jain 2010 [44]	16	2.8	26	15.6	2.5	22	0.15 [-0.42, 0.72]
Vestergaard 2011 [22]	10.5	2.87	29	10	2.87	29	0.17 [-0.34, 0.69]
2.1.5 Skills: online digital education versus no intervention							
Smeekens 2011 [53]	89	19	13	71	18	12	0.94 [0.11, 1.77]
2.1.6 Skills: offline digital education versus traditional learning							
Koele-Schmidt 2016 [46]	-20.8	16.34	10	-22.8	14.58	10	0.12 [-0.75, 1.00]
2.1.7 Skills: offline digital education versus no intervention							
Gordon 2011 [54]	79	12.1	99	63	13.5	106	1.24 [0.94, 1.54]
2.1.8 Skills: blended learning versus traditional learning							
Koele-Schmidt 2016 [46]	-19.1	11.65	10	-22.8	14.58	10	0.27 [-0.61, 1.15]
2.1.9 Satisfaction: offline digital education versus no intervention							
Gordon 2011 [54]	9	1.3	99	8	1.4	106	0.74 [0.45, 1.02]
2.1.10 Satisfaction: online digital education versus traditional learning							
Jain 2010 [44]	4.7	1.68	26	5	1.68	22	-0.18 [-0.74, 0.39]
2.1.11 Attitude: offline plus online digital education versus no intervention							
Le 2010 [59]	0.31	0.67	15	-0.1	0.8	9	0.55 [-0.29, 1.39]

Five studies (244 participants) assessed the effect of computer-based learning on knowledge [44,49,52,58,59]. One 3-arm study compared online education that included Web-based training on childhood weight management with traditional learning and no intervention. This study found no difference between online digital education and traditional learning (SMD 0.05; 95% CI –0.64 to 0.73) but reported greater knowledge scores following online digital education compared with no intervention (SMD 2.60; 95% CI 1.66 to 3.54) [49]. A further study that compared online and offline digital education with no intervention did not find any difference in knowledge between groups (SMD 0.40; 95% CI –0.44 to 1.23) [59]. Two studies compared online learning with traditional learning [44,58]. One found greater knowledge scores with online learning (SMD 2.56; 95% CI 1.52 to 3.59) [58] and the other found better outcome in the control group at postintervention with adjusted pretest values (SMD –13.68; 95% CI –16.59 to –10.77). However, it showed no difference between groups with unadjusted pretest values (SMD 0.17; 95% CI –0.40 to 0.74) [44]. One study that compared online learning with no intervention only reported knowledge outcomes for the intervention group [52].

Three studies (333 participants) provided information on attitude [52,57,59]. One study provided sufficient data to calculate SMDs; this study found no difference in attitude toward the intervention between digital education and no intervention groups (SMD 0.55; 95% CI –0.29 to 1.39) [59]. Two studies only reported information on attitude toward the intervention in the online learning group, with both suggesting positive attitudes toward online digital education [52,57].

Six studies (615 participants) provided information on satisfaction, although only 2 reported data in a suitable format to allow inclusion in the forest plot (Figure 5). One study reported greater satisfaction with offline digital education compared with no intervention (SMD 0.74; 95% CI 0.45 to 1.02) [54], and one study reported that participant satisfaction was similar for online digital education and traditional learning (SMD –0.18; 95% CI –0.74 to 0.39) [44]. Four studies only reported data for the intervention groups and reported that participants were satisfied with online and offline digital learning in terms of duration of the intervention, usability, format, and design [52,57-59].

Mobile Learning

One study (176 participants) compared mLearning using a smartphone-based mobile app software with traditional learning to provide training in neonatal resuscitation to midwives and health extension workers. The study reported greater knowledge (SMD 0.91; 95% CI 0.60 to 1.22) and skill (SMD 0.99; 95%

CI 0.68 to 1.30) scores in the mLearning group [45]. The study also assessed the impact on patient outcomes and found no difference in perinatal mortality rate between learning groups (OR 0.60; 95% CI 0.35 to 1.03).

Virtual Reality Environments

One small study (32 participants) compared a virtual reality environment (ie, commercial virtual learning platform, SecondLife, which provides computer-generated virtual patient scenarios for clinical cases management) with online digital education to provide training in pediatric sedation to pediatric residents. This study found no difference in skill (SMD 0.24; 95% CI –0.47 to 0.94) [55] or knowledge (*P*=.14) between groups. None of the included studies reported adverse or unintended effects of digital education interventions or economic outcomes.

Discussion

Principal Findings

We identified 20 studies assessing the effects of digital interventions for education in the field of pediatrics for postregistration health professionals. Included studies assessed a broad range of interventions, comparisons, and outcomes. All studies reported that digital education was either as effective as or more effective than the control intervention for outcomes including skill, knowledge, attitude, and satisfaction. Only one study with high risk of other bias due to baseline imbalances suggested that outcomes were worse with digital education compared with standard learning. All studies reported data on either skill or knowledge. One study reported physicians' change in practicing behavior and found similar effects between offline plus online digital education and no intervention. The only study to assess impact on patient outcome found no difference between intervention and control groups. The risk of bias was mainly unclear or high and the quality of evidence was low due to study inconsistencies, limitations, and/or imprecision across the included studies.

Studies assessed the use of different forms of digital education technologies such as high-fidelity mannequins (30%), computer-based education (60%), mLearning (5%), and virtual reality environment (5%). The majority of participants included in the studies were pediatric residents and junior doctors. Only four studies focused on nurses and one on midwives.

Strengths and Limitations

Our systematic review has a number of strengths. We followed Cochrane guidance to minimize the risk of bias in the review process [40]. We used the Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) criteria and assessed the quality of evidence for each primary and secondary outcome for the comparison with high-fidelity mannequin and low-fidelity mannequin. We conducted a comprehensive search across a broad range of databases to identify relevant studies. We searched for studies going back to 1990 as we considered that studies published before this were unlikely to be applicable to current practice due to technological advances. We included any postregistration health professionals working in the field of pediatrics to cover all available evidence on different types of participants on the topic. We also covered studies of any type of digital education interventions, which are primarily designed to deliver learning contents for pediatric education to synthesize the most robust evidence on the use of digital education for pediatric education. We believe that covering different types of health professions using different digital education technologies for pediatric education would provide the most comprehensive evidence on the topic. We conducted a formal risk of bias assessment to identify potential sources of bias in the primary studies. Two independent reviewers were involved in all stages of the review process to minimize the risk of bias and errors.

The small number of included studies meant that it was not possible to carry out any subgroup analyses or assess the risk of publication bias. Therefore, there is likelihood of publication bias, and the chances of publication bias cannot be ruled out in this case. The only comparison for which sufficient data were available to estimate summary effect sizes was for the impact of high-fidelity mannequins compared with low-fidelity mannequins on skill. Differences in interventions evaluated, populations targeted, and outcomes assessed also precluded meta-analysis for other outcomes and types of digital education.

The main methodological limitation of the included studies identified by our risk of bias assessment was the large number of withdrawals in 5 of the included studies. Almost half of the studies (9 out of 20) did not report on methods of randomization or allocation concealment and so it was not possible to judge whether appropriate steps were taken to minimize the risk of bias for these domains. Details of the intervention were often poorly reported, and most studies used a nonvalidated instrument to assess outcomes. Some studies did not assess skill level before the intervention so an imbalance across groups cannot be excluded. None of the studies reported following the Consolidated Standards on Reporting Trials (CONSORT) statement [61] or any other reporting guidelines. Twelve out of 20 studies included fewer than 50 participants, which meant they were unlikely to have sufficient power to identify differences between intervention groups. Only 3 of the 20 included studies were conducted in LMICs, thus the results of our review may have limited applicability for policymakers in in these countries. There was limited information on outcomes such as attitude, satisfaction, patient outcomes, costs, and adverse or untoward effects of digital education.

We are not aware of any other systematic reviews that focus on the effectiveness of digital education interventions for health professions in the field of pediatrics. Our review highlights the most up-to-date and comprehensive evidence regarding the effectiveness of digital education on the topic.

Further primary studies are needed to assess the impact of digital education for continuation of education for health professionals in the field of pediatrics. Studies should compare digital education with traditional face-to-face learning rather than with no intervention. Possibilities to continue education of health care workers in LMICs are limited, and it is known that the level of knowledge and skills is lower than the level of postgraduate health care workers in HICs [62,63]. Therefore, it could be argued that digital education interventions could be advantageous in these settings. However, only 3 out of the 20 studies were conducted in these settings. With limited studies from LMIC and poor quality of evidence for reported outcomes, this means that the applicability of the evidence from our review might be limited for policymakers implementing health policies in these countries. Moreover, there is limited information on other outcomes such as attitude, satisfaction, costs, and adverse or untoward effect of digital education interventions.

Conclusion

Digital education for postregistration health professionals in pediatrics is at least as effective as traditional learning and more effective than no learning. High-fidelity mannequins were found to be more effective at improving psychomotor skills than traditional learning with low-fidelity mannequins. Computer-based offline/online digital education was better than no intervention for knowledge and skill outcomes and as good as traditional face-to-face learning. The evidence on other outcomes and other digital education modalities was limited. This review highlights evidence gaps in the field of digital education for health professions calling for more methodologically rigorous RCTs on the effectiveness of other forms of digital education such as mLearning, virtual reality environments, virtual patient scenarios, serious gaming and gamification, and massive open online courses for education of pediatric health professionals.

Acknowledgments

We gratefully acknowledge funding support from Nanyang Technological University Singapore (eLearning for Health Professional Education grant).

Authors' Contributions

SB and BMK 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. All authors were involved in study concept and design and acquisition, analysis, or interpretation of the data. SB, BMK, PW, LL, and LTC were involved in drafting the manuscript; all authors were involved in revision. LTC supervised the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
MEDLINE (Ovid) search strategy.
[PDF File (Adobe PDF File)31 KB - jmir_v21i9e14231_app1.pdf]

Multimedia Appendix 2
Data extraction form.
[PDF File (Adobe PDF File)62 KB - jmir_v21i9e14231_app2.pdf]

Multimedia Appendix 3
Risk of bias assessment for cluster randomized controlled trials.
[PDF File (Adobe PDF File)19 KB - jmir_v21i9e14231_app3.pdf]

Multimedia Appendix 4
Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist.
[PDF File (Adobe PDF File)67 KB - jmir_v21i9e14231_app4.pdf]

Multimedia Appendix 5
Glossary.
[PDF File (Adobe PDF File)33 KB - jmir_v21i9e14231_app5.pdf]

Multimedia Appendix 6
Characteristics of studies.
[PDF File (Adobe PDF File)56 KB - jmir_v21i9e14231_app6.pdf]

Multimedia Appendix 7
Summary of findings table: effects of high fidelity mannequin on knowledge, skills, attitude, satisfaction, and behavior change outcomes.
[PDF File (Adobe PDF File)41 KB - jmir_v21i9e14231_app7.pdf]

Multimedia Appendix 8
Results of the included studies.
[PDF File (Adobe PDF File)42 KB - jmir_v21i9e14231_app8.pdf]

References

1. World Health Organization. World Health Statistics 2015 URL: http://apps.who.int/iris/bitstream/10665/170250/1/9789240694439_eng.pdf?ua=1&ua=1 [accessed 2019-09-04]
2. World Health Organization. 2015. Progress Report on the Health-Related Millennium Development Goals URL: https://www.who.int/en/news-room/fact-sheets/detail/millennium-development-goals-(mdgs) [accessed 2019-09-04] [WebCite Cache ID 77CjoJf1L]
3. Lassi ZS, Mallick D, Das JK, Mal L, Salam RA, Bhutta ZA. Essential interventions for child health. Reprod Health 2014;11 Suppl 1:S4 [FREE Full text] [doi: 10.1186/1742-4755-11-S1-S4] [Medline: 25177974]
4. Bale Jr SB, Lucas AO, editors. Reducing neonatal mortality and morbidity. In: Improving Birth Outcomes: Meeting the Challenge in the Developing World. Washington: Institute of Medicine (US) Committee on Improving Birth Outcomes; 2003:90-133.
5. United Nations (UN) General Assembly. 2015. Draft outcome document of the United Nations summit for the adoption of the post-2015 development agenda URL: https://digitallibrary.un.org/record/800852 [accessed 2019-09-04]
6. World Health Organization. 2015. Progress Report on Health Workforce Education, 2013-2015 URL: https://www.who.int/hrh/documents/education_training_report13-15.pdf?ua=1 [accessed 2019-09-04] [WebCite Cache ID 77CjtzVAO]
7. Goldman R, Koren G. Keeping it safe in the paediatric emergency department - drug errors and ways to prevent them. Paediatr Child Health 2007 Jul;12(6):461-464 [FREE Full text] [Medline: 19030408]
8. Committee on Infectious Diseases. Recommended childhood and adolescent immunization schedules: United States, 2019. Pediatrics 2019 Mar;143(3):e20190065. [doi: 10.1542/peds.2019-0065] [Medline: 30824606]
9. Unguru Y, Bernhardt MB, Berg SL, Johnson LM, Pyke-Grimm K, Woodman C, et al. Chemotherapy and supportive care agents as essential medicines for children with cancer. JAMA Pediatr 2019 May 01;173(5):477-484. [doi: 10.1001/jamapediatrics.2019.0070] [Medline: 30830204]
10. Miyakawa R, Barreto NB, Kato RM, Neely MN, Russell CJ. Early use of anti-influenza medications in hospitalized children with tracheostomy. Pediatrics 2019 Mar;143(3):e20182608. [doi: 10.1542/peds.2018-2608] [Medline: 30814271]
11. Vento M, Saugstad OD. Oxygen supplementation in the delivery room: updated information. J Pediatr 2011 Feb;158(2 Suppl):e5-e7. [doi: 10.1016/j.jpeds.2010.11.004] [Medline: 21238706]
12. Xu X, Posadzki PP, Lee GE, Car J, Smith HE. Digital education for health professions in the field of dermatology: a systematic review by Digital Health Education Collaboration. Acta Derm Venereol 2019 Feb 01;99(2):133-138 [FREE Full text] [doi: 10.2340/00015555-3068] [Medline: 30320871]
13. George PP, Papachristou N, Belisario JM, Wang W, Wark PA, Cotic Z, et al. Online eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health 2014 Jun;4(1):010406 [FREE Full text] [doi: 10.7189/jogh.04.010406] [Medline: 24976965]
14. Rasmussen K, Belisario JM, Wark PA, Molina JA, Loong SL, Cotic Z, et al. Offline eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health 2014 Jun;4(1):010405 [FREE Full text] [doi: 10.7189/jogh.04.010405] [Medline: 24976964]
15. Kyaw BM, Posadzki P, Dunleavy G, Semwal M, Divakar U, Hervatis V, et al. Offline digital education for medical students: systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Mar 25;21(3):e13165 [FREE Full text] [doi: 10.2196/13165] [Medline: 30907731]
16. Tudor Car L, Marcano BJ, Nikolaou C, Majeed A, Zary N, Car J. Mobile learning for delivering health professional education. Cochrane Database Syst Rev 2015;9:CD011861. [doi: 10.1002/14651858.CD011861.pub2]
17. Dunleavy G, Nikolaou C, Nifakos S, Atun R, Law G, Tudor Car L. Mobile digital education for health professions: systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 12;21(2):e12937 [FREE Full text] [doi: 10.2196/12937] [Medline: 30747711]
18. Saxena NK, Vseteckova J, Dev P, Paul P, Lim K, Kononowicz A, et al. Virtual reality environments for health professional education (Protocol). Cochrane Database Syst Rev 2016:CD012090 [FREE Full text] [doi: 10.1002/14651858.cd012090]
19. Vash JH, Yunesian M, Shariati M, Keshvari A, Harirchi I. Virtual patients in undergraduate surgery education: a randomized controlled study. ANZ J Surg 2007;77(1-2):54-59. [doi: 10.1111/j.1445-2197.2006.03978.x] [Medline: 17295822]
20. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou C, George P, et al. Virtual reality for health professions education: systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
21. Stellflug SM. The Effect of Simulation Training on Knowledge Retention and Skill Self Efficacy in Pedatric Advanced Life Support Courses [thesis]. Boulder: University of Colorado; May 2015.
22. Vestergaard L, Loofgren B, Jessen C, Boorlum C, Wolff A, Nielsen H. Pediatric basic life support self-training is comparable to instructor-led training: a randomized manikin study. Circulation 2011;124(21):A156.

23. Sharara-Chami R, Taher S, Kaddoum R, Tamim H, Charafeddine L. Simulation training in endotracheal intubation in a pediatric residency. Middle East J Anaesthesiol 2014;22(5):477-485. [doi: 10.5005/jp/books/12116_19]
24. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. Acad Med 2010 May;85(5):909-922. [doi: 10.1097/ACM.0b013e3181d6c319] [Medline: 20520049]
25. Lahti M, Hätönen H, Välimäki M. Impact of e-learning on nurses' and student nurses knowledge, skills, and satisfaction: a systematic review and meta-analysis. Int J Nurs Stud 2014 Jan;51(1):136-149. [doi: 10.1016/j.ijnurstu.2012.12.017] [Medline: 23384695]
26. World Health Organization. 2015. eLearning for undergraduate health professional education: a systematic review informing a radical transformation of health workforce development URL: http://whoeducationguidelines.org/sites/default/files/uploads/eLearning?healthprof?report.pdf [accessed 2017-12-11] [WebCite Cache ID 77CjyoX0e]
27. Semwal M, Whiting P, Bajpai R, Bajpai S, Kyaw B, Tudor Car L. Digital education for health professions on smoking cessation management: systematic review by the Digital Health Education Collaboration. J Med Internet Res 2019 Mar 04;21(3):e13000 [FREE Full text] [doi: 10.2196/13000] [Medline: 30829576]
28. Glegg SMN, Livingstone R, Montgomery I. Facilitating interprofessional evidence-based practice in paediatric rehabilitation: development, implementation and evaluation of an online toolkit for health professionals. Disabil Rehabil 2016;38(4):391-399. [doi: 10.3109/09638288.2015.1041616] [Medline: 25924019]
29. Lewis KO, Cidon MJ, Seto TL, Chen H, Mahan JD. Leveraging e-learning in medical education. Curr Probl Pediatr Adolesc Health Care 2014 Jul;44(6):150-163. [doi: 10.1016/j.cppeds.2014.01.004] [Medline: 24981664]
30. Rao A, O'Leary F. Training clinicians to care for children in emergency departments. J Paediatr Child Health 2016 Feb;52(2):126-130. [doi: 10.1111/jpc.13108] [Medline: 27062614]
31. PROSPERO. 2017. eLearning interventions for continuous medical education of health care professionals on pediatrics URL: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017057793 [accessed 2019-03-27] [WebCite Cache ID 77BNZXkpv]
32. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, Digital Health Education Collaboration. Digital education in health professions: the need for overarching evidence synthesis. J Med Internet Res 2019 Feb 14;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
33. Sangrà A, Vlachopoulos D, Cabrera N. Building an inclusive definition of e-learning: an approach to the conceptual framework. IRRODL 2012 Apr 13;13(2):145. [doi: 10.19173/irrodl.v13i2.1161]
34. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: systematic review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]
35. George PP, Zhabenko O, Kyaw BM, Antoniou P, Posadzki P, Saxena N, et al. Online digital education for postregistration training of medical doctors: systematic review by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 25;21(2):e13269 [FREE Full text] [doi: 10.2196/13269] [Medline: 30801252]
36. Posadzki P, Bala M, Kyaw BM, Semwal M, Divakar U, Koperny M, et al. Offline digital education for postregistration health professions: systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Apr 24;21(4):e12968 [FREE Full text] [doi: 10.2196/12968] [Medline: 31017584]
37. International Standard Classification of Occupations. 2008. Options for the classification of health occupations in the Updated International Standard Classification of Occupations (ISCO-08) URL: http://www.ilo.org/public/english/bureau/stat/isco/docs/healthocc.pdf [accessed 2019-08-21] [WebCite Cache ID 77BOwmvDC]
38. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol 2014 Dec 19;14:135 [FREE Full text] [doi: 10.1186/1471-2288-14-135] [Medline: 25524443]
39. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 07;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
40. Higgins JG. The Cochrane Collaboration. 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1 URL: http://handbook-5-1.cochrane.org/ [accessed 2019-08-21]
41. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. BMJ 2003 Oct 04;327(7418):785-789 [FREE Full text] [doi: 10.1136/bmj.327.7418.785] [Medline: 14525877]
42. Ryan R. Cochrane Consumers and Communications Review Group. 2016 Dec. Heterogeneity and subgroup analyses in Cochrane Consumers and Communication Group reviews: planning the analysis at protocol stage URL: https://cccrg.cochrane.org/sites/cccrg.cochrane.org/files/public/uploads/heterogeneity_subgroup_analyses_revising_december_1st_2016.pdf [accessed 2019-08-21] [WebCite Cache ID 77BRawcWD]
43. Schünemann H, Oxman A, Higgins J, Vist G, Glasziou P, Guyatt G. Presenting results and “summary of findings” tables. Cochrane Handbook for Systematic Reviews of Interventions 2008:335-357. [doi: 10.1002/9780470712184.ch11]

44. Jain A, Agarwal R, Chawla D, Paul V, Deorari A. Tele-education vs classroom training of neonatal resuscitation: a randomized trial. J Perinatol 2010 Dec;30(12):773-779. [doi: 10.1038/jp.2010.42] [Medline: 20357810]
45. Lund S, Boas IM, Bedesa T, Fekede W, Nielsen HS, Sørensen BL. Association between the safe delivery app and quality of care and perinatal survival in ethiopia: a randomized clinical trial. JAMA Pediatr 2016 Aug 01;170(8):765-771. [doi: 10.1001/jamapediatrics.2016.0687] [Medline: 27322089]
46. Koele-Schmidt L, Vasquez MM. NewB for newbies: a randomized control trial training housestaff to perform neonatal intubation with direct and videolaryngoscopy. Paediatr Anaesth 2016 Apr;26(4):392-398. [doi: 10.1111/pan.12832] [Medline: 26714736]
47. Campbell DM, Barozzino T, Farrugia M, Sgro M. High-fidelity simulation in neonatal resuscitation. Paediatr Child Health 2009 Jan;14(1):19-23 [FREE Full text] [doi: 10.1093/pch/14.1.19] [Medline: 19436459]
48. Donoghue A, Durbin D, Nadel F, Stryjewski GR, Kost SI, Nadkarni VM. Effect of high-fidelity simulation on Pediatric Advanced Life Support training in pediatric house staff: a randomized trial. Pediatr Emerg Care 2009 Mar;25(3):139-144. [doi: 10.1097/PEC.0b013e31819a7f90] [Medline: 19262421]
49. Benjamin SE, Tate DF, Bangdiwala SI, Neelon BH, Ammerman AS, Dodds JM, et al. Preparing child care health consultants to address childhood overweight: a randomized controlled trial comparing web to in-person training. Matern Child Health J 2008 Sep;12(5):662-669. [doi: 10.1007/s10995-007-0277-1] [Medline: 17713847]
50. Pollak KI, Tulsky JA, Bravender T, Østbye T, Lyna P, Dolor RJ, et al. Teaching primary care physicians the 5 A's for discussing weight with overweight and obese adolescents. Patient Educ Couns 2016 Oct;99(10):1620-1625 [FREE Full text] [doi: 10.1016/j.pec.2016.05.007] [Medline: 27228899]
51. Alade K, Caviness A, Miller S, Press G. Evaluation of a pediatric emergency ultrasonography course. Ann Emerg Med 2012 Oct;60(4):S68. [doi: 10.1016/j.annemergmed.2012.06.167]
52. Dingeldein L, Sheehan K, Krcmarik M, Dowd MD. Evaluation of a firearm injury prevention web-based curriculum. Teach Learn Med 2012;24(4):327-333. [doi: 10.1080/10401334.2012.715254] [Medline: 23036000]
53. Smeekens AEFN, Broekhuijsen-van Henten DM, Sittig JS, Russel IMB, ten Cate OTJ, Turner NM, et al. Successful e-learning programme on the detection of child abuse in emergency departments: a randomised controlled trial. Arch Dis Child 2011 Apr;96(4):330-334. [doi: 10.1136/adc.2010.190801] [Medline: 21233084]
54. Gordon M, Chandratilake M, Baker P. Improved junior paediatric prescribing skills after a short e-learning intervention: a randomised controlled trial. Arch Dis Child 2011 Dec;96(12):1191-1194. [doi: 10.1136/archdischild-2011-300577] [Medline: 22080459]
55. Zaveri PP, Davis AB, O'Connell KJ, Willner E, Aronson Schinasi DA, Ottolini M. Virtual reality for pediatric sedation: a randomized controlled trial using simulation. Cureus 2016 Feb 09;8(2):e486 [FREE Full text] [doi: 10.7759/cureus.486] [Medline: 27014520]
56. Kim S, Shin G. Effects of nursing process-based simulation for maternal child emergency nursing care on knowledge, attitude, and skills in clinical nurses. Nurse Educ Today 2016 Feb;37:59-65. [doi: 10.1016/j.nedt.2015.11.016] [Medline: 26655857]
57. Gordon JS, Mahabee-Gittens EM, Andrews JA, Christiansen SM, Byron DJ. A randomized clinical trial of a web-based tobacco cessation education program. Pediatrics 2013 Feb;131(2):e455-e462 [FREE Full text] [doi: 10.1542/peds.2012-0611] [Medline: 23319529]
58. Hearty T, Maizels M, Pring M, Mazur J, Liu R, Sarwark J, et al. Orthopaedic resident preparedness for closed reduction and pinning of pediatric supracondylar fractures is improved by e-learning: a multisite randomized controlled study. J Bone Joint Surg Am 2013 Sep 04;95(17):e1261-e1267. [doi: 10.2106/JBJS.L.01065] [Medline: 24005210]
59. Le TT, Rait MA, Jarlsberg LG, Eid NS, Cabana MD. A randomized controlled trial to evaluate the effectiveness of a distance asthma learning program for pediatricians. J Asthma 2010 Apr;47(3):245-250. [doi: 10.3109/02770900903560209] [Medline: 20394510]
60. Thomas EJ, Williams AL, Reichman EF, Lasky RE, Crandell S, Taggart WR. Team training in the neonatal resuscitation program for interns: teamwork and quality of resuscitations. Pediatrics 2010 Mar;125(3):539-546. [doi: 10.1542/peds.2009-1635] [Medline: 20156896]
61. CONSORT. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2011 Sep 29;343(sep29 2):d6131-d6131. [doi: 10.1136/bmj.d6131]
62. Pakenham-Walsh N, Bukachi F. Information needs of health care workers in developing countries: a literature review with a focus on Africa. Hum Resour Health 2009 Apr 08;7:30 [FREE Full text] [doi: 10.1186/1478-4491-7-30] [Medline: 19356239]
63. Ameh CA, Kerr R, Madaj B, Mdegela M, Kana T, Jones S, et al. Knowledge and skills of healthcare providers in sub-Saharan Africa and Asia before and after competency-based training in emergency obstetric and early newborn care. PLoS One 2016;11(12):e0167270 [FREE Full text] [doi: 10.1371/journal.pone.0167270] [Medline: 28005984]

Abbreviations

CME: continuing medical education

CONSORT: Consolidated Standards of Reporting Trials
cRCT: cluster randomized controlled trial
GRADE: Grading of Recommendations, Assessment, Development, and Evaluations
HIC: high-income countries
LMIC: low- and middle-income countries
OR: odds ratio
RCT: randomized controlled trial
SMD: standardized mean difference

Edited by G Eysenbach; submitted 16.04.19; peer-reviewed by IN Gomez, J Derzon; comments to author 15.06.19; revised version received 17.07.19; accepted 18.08.19; published 25.09.19.

Please cite as:
Brusamento S, Kyaw BM, Whiting P, Li L, Tudor Car L
Digital Health Professions Education in the Field of Pediatrics: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration
J Med Internet Res 2019;21(9):e14231
URL: https://www.jmir.org/2019/9/e14231
doi:10.2196/14231
PMID:31573906

Corrigenda and Addenda

Multimedia Appendix Correction: Patient Health Record Systems Scope and Functionalities: Literature Review and Future Directions

Lina Bouayad[1,2*], PhD; Anna Ialynytchev[2*], PhD; Balaji Padmanabhan[3*], PhD

[1]Department of Information Systems and Business Analytics, Florida International University, Miami, FL, United States
[2]Health Services Research and Development Service, Center of Innovation on Disability and Rehabilitation Research, Tampa, FL, United States
[3]Department of Information Systems and Decision Sciences, University of South Florida, Tampa, FL, United States
*all authors contributed equally

Corresponding Author:
Lina Bouayad, PhD
Department of Information Systems and Business Analytics
Florida International University
11200 SW 8th Street
Miami, FL, 33199
United States
Phone: 1 305 348 5437
Email: lbouayad@fiu.edu

Related Article:

Correction of: https://www.jmir.org/2017/11/e388

(J Med Internet Res 2019;21(9):e15796) doi:10.2196/15796

The authors of “Patient Health Record Systems Scope and Functionalities: Literature Review and Future Directions” (J Med Internet Res 2017;19(11):e388) have replaced Multimedia Appendix 1 with a new file. Because of formatting issues, the original appendix was missing data elements in the table. Published here is a new version of Multimedia Appendix 1 with all data elements.

The corrections will appear in the online version of the paper on the JMIR website on September 5, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

Multimedia Appendix 1
Patient data elements reported in the literature.
[PDF File (Adobe PDF File), 90 KB - jmir_v21i9e15796_app1.pdf]

Submitted 07.08.19; this is a non–peer-reviewed article;accepted 12.08.19; published 05.09.19.

Please cite as:
Bouayad L, Ialynytchev A, Padmanabhan B
Multimedia Appendix Correction: Patient Health Record Systems Scope and Functionalities: Literature Review and Future Directions
J Med Internet Res 2019;21(9):e15796
URL: http://www.jmir.org/2019/9/e15796/
doi:10.2196/15796
PMID:31489840

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org